STRUCTURAL STUDIES OF WILD-TYPE AND VARIANT YEAST ISO-1-CYTOCHROMES *c*

by

Gordon V. Louie

B.Sc., The University of British Columbia, 1984

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

November 1990

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.

Department of Biochemistry

The University of British Columbia
Vancouver, Canada

Date November 13, 1990

ABSTRACT

The crystal structure of yeast (*Saccharomyces cerevisiae*) iso-1-cytochrome *c* has been determined through molecular replacement techniques, and refined against X-ray diffraction data in the resolution range 6.0-1.23 Å to a crystallographic R-factor of 0.192. The yeast iso-1-cytochrome *c* molecule has the typical cytochrome *c* fold, with the polypeptide chain organized into five α-helices and a series of loops which serve to enclose almost completely the heme prosthetic group within a hydrophobic pocket. Comparison of the structures of yeast iso-1-, tuna and rice cytochromes *c* shows that the polypeptide backbone fold, intramolecular hydrogen bonding, conformation of side chains and particularly packing within the heme crevice of protein groups against the heme moiety are very similar in the three proteins. Significant structural differences among the three cytochromes *c* can be explained by differences in amino acid sequence.

X-ray crystallographic techniques have also been used to study the effect of single-site amino acid substitutions at Phe82 and at Arg38 in iso-1-cytochrome *c*. The structures of the various variant iso-1-cytochromes *c* have been determined at nominal resolutions in the range 2.8 to 1.76 Å. Conspicuous structural perturbations in the neighborhood of the substituted side chain are evident in all of the variant proteins. In wild-type iso-1-cytochrome *c*, the phenyl ring of Phe82 is positioned adjacent and approximately parallel to the heme group, and occupies a non-polar cavity within the heme crevice. In the Ser82 variant, a channel extending from the surface of the molecule down into the heme crevice is created. In the Gly82 variant, the polypeptide backbone has refolded into the space formerly occupied by the phenyl ring of Phe82. Steric conflicts prevent both the phenolic ring of Tyr82 and the side chain of Ile82 from being completely accommodated within the pocket normally occupied by a phenyl ring. Substitution of alanine at position 38 causes a slight reorganization of the hydrogen bonding network in which Arg38 normally participates, and also exposes to external solvent a normally buried propionic acid group of the heme.

The altered functional properties of the position 82 variant proteins have been interpreted with respect to the observed structural perturbations. The drop in reduction potential, most notably

for the Ser82 and Gly82 variants, can be explained by the elevated heme environment polarity arising from the increased access of solvent or polar protein groups to the heme pocket. The reduced stability of the heme crevice, as indicated by lowered pK_a's for alkaline isomerization, is likely due to the disruption of stabilizing packing forces formed by the Phe82 phenyl ring within its hydrophobic cavity. The lowered activity, in comparison to the wild-type protein and the Tyr82 variant, for electron transfer with Zn^+-cytochrome *c* peroxidase is attributed to the loss of an aromatic group positioned adjacent to the heme group. The altered surface topography of the variant proteins (particularly the Gly82, Tyr82 and Ile82 variants) may further hinder productive complex formation between cytochrome *c* and its redox partners. These results suggest that the invariant Phe82 contributes in at least three ways to the proper functioning of cytochrome *c*. It has an important structural role in maintaining the integrity of the heme crevice and in establishing the appropriate heme environment. The phenyl ring of Phe82 may also be required for efficient movement of an electron to and from the heme of cytochrome *c*. Finally, Phe82 may have a role in forming intermolecular interactions with enzymic redox partners of cytochrome *c*.

TABLE OF CONTENTS

LIST OF TABLES

IX. ALANINE 38 VARIANT OF YEAST ISO-1-CYTOCHROME *c*

LIST OF FIGURES

ABBREVIATIONS AND SYMBOLS USED

a, b, c	Crystallographic unit cell axes, or axis lengths
a^*, b^*, c^*	Axes of the crystallographic reciprocal lattice
α, β, γ	Euler angles specifying orientation of molecule
α	Phase associated with a structure factor magnitude (sometimes written with a subscript 'c' to emphasize that the phase is calculated)
ATP	Adenosine 5'-triphosphate
CCP	Yeast cytochrome *c* peroxidase
CD	Circular dichroism
DTT	Dithiothreitol
E_m	Midpoint reduction potential (usually at pH 7)
EDTA	Ethylene diamine tetraacetic acid
F	Structure factor magnitude (F_o: observed; F_c: calculated; F_c': calculated without the contribution of a fragment of the atomic model)
h, k, l	Miller indices specifying indices of a X-ray reflection
I	Observed diffraction intensity
NMR	Nuclear magnetic resonance
NOE	Nuclear Overhauser effect
ORD	Optical rotatory dispersion
r.m.s.	Root mean squared
σ_F	Estimated standard deviation in the magnitude of an observed structure factor
SDS	Sodium dodecyl sulfate
Tml	ϵ-N-trimethylated lysine
TMPD	N,N,N',N'-tetramethyl-p-phenylenediamine

UV	Ultraviolet
u, *v*, *w*	Fractional coordinates specifying position in Patterson space
x, *y*, *z*	Fractional coordinates specifying position in the crystallographic unit cell

The conventions of the IUPAC–IUB Combined Commission 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 of the polypeptide chain [*J. Biol. Chem.* **245**, 6489–6497 (1970)]; and for describing the conformational torsion angles of main chain and side chain groups of the polypeptide chain [*J. Biol. Chem.* **245**, 6489–6497 (1970)]. Designations for the atoms of the protoheme IX group are according to the Brookhaven National Laboratory Protein Data Bank (Bernstein *et al.*, 1977; see Figure 1.1).

Throughout this thesis, amino acid residue numbers in cytochrome *c* are designated according to the numbering system for vertebrate cytochromes *c*. In particular, with respect to higher eukaryotic cytochromes *c*, the yeast iso-1-protein has a 5-residue long extension at its amino-terminus, and the residues in this extension have sequence numbers −5 to −1.

ACKNOWLEDGMENTS

I would like to offer thanks to the numerous people who have made contributions to the work described in this thesis. Chris Sherwood initially established conditions suitable for growth of yeast iso-1-cytochrome *c* crystals. Connie Leung kindly provided many of the crystals used in structural determinations. Gary Pielak, Rob Cutler, Alf Gartner, Linda Pearce, Jeanette Johnson, and others have supplied protein samples for crystallization. Past and present members of the laboratories of Gary Brayer, Grant Mauk and Michael Smith have been the source of numerous stimulating and enlightening discussions. I am particularly indebted to Stephen Evans for writing many invaluable and clever programs used in structural analyses. I sincerely acknowledge my research supervisor, Gary Brayer, for the insights, guidance, and encouragement he has provided during my stay in his lab. I am also grateful to my other two supervisory committee members, Grant Mauk and Ross MacGillivray, for their careful reading of this thesis and for considerate advice and suggestions. Also, the Medical Research Council of Canada is acknowledged for provision of a Studentship. Finally, I must extend my deepest appreciation to my family for their patient support over these past years.

I. INTRODUCTION

A. The cytochromes *c*

Cytochrome *c* is an electron transfer protein that occurs in the energy-transducing electron transport chains of many organisms. Because of its ubiquity, small size, solubility, and ease of extraction and purification, and the great importance of electron transfer reactions such as those carried out by this protein, cytochrome *c* has been extensively studied (for review articles, see Margoliash and Schejter, 1966; Harbury and Marks, 1973; Dickerson and Timkovich, 1975; Salemme, 1977; Ferguson-Miller *et al.*, 1979; Timkovich, 1979; Meyer and Kamen, 1982; Mathews, 1985; Williams *et al.*, 1985b; Pettigrew and Moore, 1987; Pielak *et al.*, 1987a; Dixon, 1988).

The cytochromes *c* are proteins which contain a heme prosthetic group, derived from the protoporphyrin IX skeleton. The heme is attached covalently to the polypeptide chain by thioether bonds formed through the addition of a cysteinyl sulfhydryl group to each of the two heme vinyls. The two cysteine residues involved occur in the sequence Cys-X-Y-Cys-His, with the histidine residue contributing one of the axial ligands to the heme iron. The various classes of cytochrome *c* differ in the location along the polypeptide chain of the heme attachment site, the number of heme groups per polypeptide chain, the identity of the sixth axial ligand, and the occurrence of additional prosthetic groups (e.g. flavins). The best characterized class is mitochondrial cytochrome *c*. Cytochromes *c* of this type have a single heme bound near the N-terminal end of the polypeptide chain; have as the sixth axial ligand a methionine side chain, provided by a residue occurring near the C-terminus; and have a relatively positive reduction potential.

The mitochondrial cytochromes *c*, which include proteins of both eukaryotic and prokaryotic origin, have been further subgrouped according to the length of the polypeptide chain (Dickerson, 1980). Members of the small subclass have 79-99 amino acid residues, the medium subclass 100-112, and the large subclass 113-134. Eukaryotic cytochromes *c* belong to the medium subclass, and have highly conservative amino acid sequences. In contrast, among prokaryotic cytochromes *c* of the mitochondrial type, there is considerable diversity in both size (members of

all three subclasses occur) and amino acid sequence. Discussion herein will be almost completely restricted to mitochondrial eukaryotic cytochrome *c*.

1. General biochemical properties of eukaryotic cytochrome *c*

In mitochondria, cytochrome *c* is a peripheral membrane protein, and occurs in the intermembrane space. It serves to carry an electron from cytochrome *c* reductase (coenzyme Q–cytochrome *c* oxidoreductase, complex III) to cytochrome *c* oxidase (complex IV), two membrane–bound complexes of the respiratory chain. That cytochrome *c* is the penultimate carrier in this electron transport chain is consistent with its relatively high reduction potential (~260 mV).

Cytochrome *c* is characterized by its covalently bound protoheme IX prosthetic group (Figure I.1). The heme group is attached to the polypeptide chain through thioether linkages between its two vinyl groups and two cysteinyl side chains (Cys14 and Cys17) from the protein molecule. [Three exceptions are observed; in the cytochromes *c* of *Euglena gracilis*, *Crithidia*

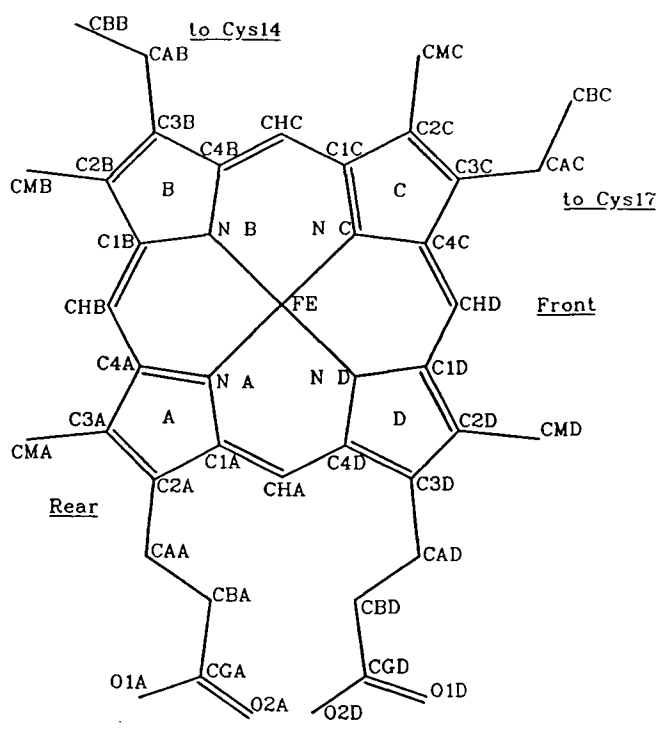

Figure I.1. Schematic representation of the protoheme IX group of cytochrome *c*. The heme atoms and pyrrole rings are labelled according the convention (Bernstein *et al.*, 1977) used throughout this thesis.

oncopelti and *Crithidia fasciculata*, Cys14 is replaced by alanine and only a single thioether linkage occurs (Pettigrew *et al.*, 1975; Hill and Pettigrew, 1975).] The UV-visible spectra of metalloporphyrins are characterized by the presence of three bands, whose positions are influenced by the nature of the side chain substituents on the heme chromophore. For cytochrome *c*, the α, β and γ (Soret) bands occur at 550-558 nm, 521-527 nm and 415-423 nm, respectively.

The heme iron atom of cytochrome *c* is relatively inert; at physiological temperature and pH, cytochrome *c* is not readily auto-oxidizable and is resistant to binding of exogenous ligands (such as cyanide and carbon monoxide). In addition to forming coordinate bonds with the four pyrrole nitrogens, the central iron atom of the heme is liganded by histidine (His18) and methionine (Met80) side chains from the protein. Cytochrome *c* carries electrons via a reversible cycling between the +2 and +3 oxidation states of the active-site heme iron. It should be noted that because coordination of the iron ion to the heme requires the dissociation of two pyrrole hydrogens, the porphyrin portion of the heme carries a -2 charge. Thus cycling between the +2 and +3 oxidation states of the heme iron actually involves a change between 0 and +1 in the net charge of the heme. Coordination of the sulfur atom of Met80 to the heme iron is associated, when the iron is the oxidized state, with an absorbance at 695 nm arising from a charge-transfer transition (Adar, 1979). The 695 nm band is frequently used as an indicator of structural integrity.

Eukaryotic cytochromes *c* are typically 103-113 amino acid residues in length. One of the most notable features of the amino acid sequence is the high content of basic residues. Cytochrome *c* is thus a very basic protein, with a pI of ~10.

2. General structural properties of cytochrome *c*

a. Structural description

The tertiary structures of several mitochondrial-type cytochromes *c* have previously been determined. These include members from several eukaryotes and also from a diverse group of prokaryotes. Cytochromes *c* in the oxidized form have been studied from horse (2.8 Å; Dickerson *et al.*, 1971) and rice (1.5 Å; Ochi *et al.*, 1983). The structures of both reduced and oxidized

cytochrome *c* of bonito (2.3 Å and 2.8 Å, respectively; Tanaka *et al.*, 1975; Matsuura *et al.*, 1979), and tuna (1.5 Å and 1.8 Å, respectively; Takano and Dickerson, 1981a, b) are known. The structures of two members of the large class of prokaryotic mitochondrial cytochrome *c* have been determined: oxidized cytochrome c_2 from the purple non-sulfur bacterium *Rhodospirillum rubrum* (1.7 Å; Salemme *et al.*, 1973; Bhatia, 1981), and oxidized cytochrome c_{550} from the facultative nitrate-reducing bacterium *Paracoccus denitrificans* (2.45 Å; Timkovich and Dickerson, 1976). In addition, the structures of several prokaryotic cytochromes *c* of the small class are known: both oxidized and reduced cytochrome c_{551} of another denitrifying bacterium, *Pseudomonas aeruginosa* (1.6 Å; Matsuura *et al.*, 1982); oxidized cytochrome c_{555} from the green sulfur bacterium *Chlorobium thiosulfatophilum* (2.7 Å; Korszun and Salemme, 1977); and cytochrome c_{554} from the cyanobacterium *Anacystis nidulans* (3.0 Å; Ludwig *et al.*, 1982).

Comparison of the structures of these cytochromes *c* shows that they share the same basic cytochrome *c* fold. It has thus been proposed that all mitochondrial-type cytochromes *c* have descended from a common evolutionary ancestor (Dickerson *et al.*, 1976). The cytochrome *c* fold resembles a clamshell of protein enclosing the heme group within a hydrophobic core (Timkovich, 1979). The general fold of mitochondrial eukaryotic cytochrome *c* can be described as follows (see Figure I.2). The amino-terminal portion of the polypeptide chain (residues 1 to 12) forms an α-helix along the top of the molecule. Cysteines 14 and 17 form the thioether linkages to the heme group, and histidine 18 ligands the heme iron. Residues 19 through 46 form a series of loops which envelop the right face of the heme group. Residues 40 through 49 constitute the floor of the cytochrome *c* molecule. A short α-helix at the lower left formed by residues 50 to 55 is followed by an extended segment of polypeptide chain composed of residues 55 to 60, which makes up the lower part of the rear wall. Two consecutive α-helices formed by residues 61 to 69 and 70 to 74 constitute the left side of the molecule. Residues 76 through 87 form a segment of extended polypeptide chain which sweeps across the upper left face of the heme group, and which provides the methionine 80 ligand. Finally, residues from 88 to the C-terminus form an α-helix which runs diagonally across the back of the molecule. The regions of the cytochrome *c*

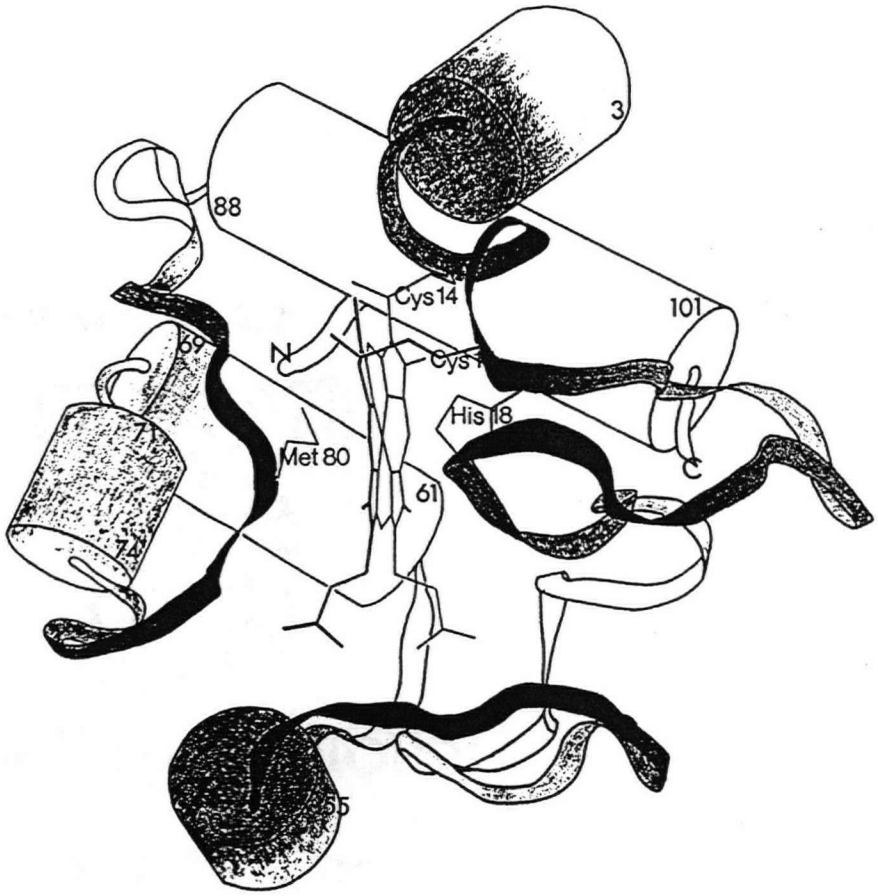

Figure I.2. Schematic representation of the structure of eukaryotic cytochrome *c*. Regions of the polypeptide chain having an α-helical conformation have been represented as solid cylinders, and regions forming loop structures as ribbons. Also shown are the heme group; the side chains of Cys14 and Cys17, which form thioether bonds to the heme; and the side chains of His18 and Met80, which form coordinate bonds to the heme iron. The N- and C- termini have been labelled, and the terminal residues of the helices have been numbered. The front of the molecule has been shaded darkly, while the rear of the molecule is unshaded. Note that the cytochrome *c* molecule is shown here in a standard orientation for viewing (Dickerson and Timkovich, 1975). Future uses of the designations front, left, right, bottom, rear, etc. will refer to this orientation. (The exposed edge of the heme group, at the front of the molecule, directly faces the viewer. The left and right sides of the molecule contain the Met80 and His18 ligands, respectively; the bottom of the molecule lies just below the heme propionic acid groups; and the rear of the molecule contains the C-terminal α-helix.)

molecule lacking extensive regular secondary structure have been categorized by Leszczynski and Rose (1986) as Ω-loops (continuous segments of polypeptide chain whose initiation and termination points lie adjacent in space). These loops span residues 18 to 32, 34 to 43, 40 to 54, and 70 to 84.

The polypeptide chains of prokaryotic cytochromes *c*, in comparison to that of eukaryotic cytochrome *c*, typically possess a number of deletions and/or insertions which occur at the external periphery of the molecule, and often at a point of chain reversal at surface loops (Timkovich, 1979). Upon comparing the structures of tuna cytochrome *c*, *Rhodospirillum rubrum* cytochrome c_2, *Paracoccus denitrificans* cytochrome c_{550}, and *Pseudomonas aeruginosa* cytochrome c_{551}, Chothia and Lesk (1985) have determined that six regions of these cytochrome *c* molecules have a common folding pattern. These six segments of polypeptide chain represent the core structure of the cytochrome *c* molecule, and all are found to constitute part of the heme pocket. The three segments encompassing residues 5–12, 60–71, and 90–101 are α-helical and form the top, lower-left and upper-rear walls, respectively, of the heme pocket. The remaining segments line the heme crevice: residues 14–20 contribute the histidine ligand and the two cysteine side chains involved in the thioether linkages; residues 24 to 33 pack directly against the right face of the heme group; and residues 79–81 pack against the left face of the heme and contribute the methionine ligand.

Several conserved structural features of the mitochondrial cytochrome *c* molecule are notable. The N-terminal and C-terminal α-helices are in contact, with the axes of the two helices being roughly perpendicular. Only a small portion of the surface of the heme group, at the front face of the cytochrome *c* molecule, is exposed to the external solvent medium. The heme group is encased by a hydrophobic core; the side chains which pack against the tetrapyrrole portion of the heme group in the interior of the cytochrome *c* molecule are almost exclusively hydrophobic or aromatic in character. The propionic acid side groups are buried in the interior of the molecule, where they form polar interactions with internal protein groups. The molecular surface of the highly basic eukaryotic cytochromes *c* is characterized by a conserved distribution of lysyl side chains, which encircle the exposed edge of the heme group. Occurring on the opposite (i.e. rearward) face of the molecule is a clustering of a small number of acidic side chains. This separation of positive and negative surface charges has been proposed to establish a dipole moment which serves to optimally orient the cytochrome *c* molecule for interaction with its redox partners

(Koppenol and Margoliash, 1982; Margoliash and Bosshard, 1983).

b. Alkaline isomerization

Theorell and Akesson (1941) identified spectroscopically five forms of ferricytochrome *c* occurring over the pH range 0 to 13. Form III, which is stable between pH 4 and 8, is the biologically functional form. The alkaline isomerization, or the conversion of form III to form IV (the alkaline form), occurs at pH ~9, and involves the deprotonation of one (or more) group(s), followed by the displacement of the Met80 heme ligand by some other strong-field ligand (Davis *et al.*, 1974). The form IV of ferricytochrome *c* cannot be reduced by cytochrome *c* reductase. Because it occurs under physiological conditions, and also inactivates the cytochrome *c* molecule, the alkaline isomerization has been suggested to have a role in regulating the activity of the electron transport chain (Gadsby *et al.*, 1987).

The identity of the sixth ligand to the heme iron in alkaline ferricytochrome *c* has not been established (Gadsby *et al.*, 1987). Possible candidates are the ϵ-amino group of a nearby lysine residue (Lys72, 73 or 79), the hydroxyl group of Tyr67, or a free hydroxide ion. The ligand exchange which occurs during the alkaline isomerization process has several other consequences. There is a slight conformational rearrangement of the cytochrome *c* molecule, and also a loosening of the left side of the heme crevice (Davis *et al.*, 1974; Wooten *et al.*, 1981). The reduction potential of the alkaline cytochrome *c* is decreased to ~120 mV (Davis *et al.*, 1974; Rodkey and Ball, 1950), which likely explains the impaired reducibility of the protein. In addition, the loss of Met80 ligation means that the alkaline isomerization can be conveniently monitored by measurement of the decrease in absorbance at 695 nm. The degree of susceptibility of various forms of cytochrome *c* to alkaline isomerization, or indeed to other denaturing agents (such as heat, urea, guanidine hydrochloride, alcohols, etc.), is frequently used as a measure of the strength of the intrinsic binding forces maintaining the native structure of the heme crevice (Osheroff *et al.*, 1980).

B. Aspects related to the electron-transfer role of cytochrome *c*

1. Electron transfer reactivity of cytochrome *c*

Cytochrome *c* carries out electron transfer reactions with a limited number of proteinaceous redox partners: within the electron transport chain, cytochrome *c* reductase and cytochrome *c* oxidase; and within the mitochondrial intermembrane space, cytochrome b_5 and sulfite oxidase (animals), and flavocytochrome b_2 and cytochrome *c* peroxidase (yeast). Cytochrome *c* forms fairly tight complexes (having dissociation constants in the range 1-5 X 10^{-7} M) with its redox partners. The rates of the electron transfer reactions can be assessed in terms of (apparent) second-order rate constants (measured experimentally under reaction conditions yielding pseudo first-order kinetics). The value of this rate constant is dependent on both the association constant for intermolecular complexation, and the first order rate constant for electron transfer within this precursor complex. Reported second order rate constants for electron transfer reactions between cytochrome *c* and its physiological redox partners are: 1.0 X 10^7 $M^{-1}s^{-1}$ for cytochrome *c* oxidase, 1.7 X 10^7 $M^{-1}s^{-1}$ for cytochrome *c* reductase, ~4.0 X 10^7 $M^{-1}s^{-1}$ for cytochrome b_5, and ~2.0 X 10^{10} $M^{-1}s^{-1}$ for cytochrome b_2 (Pettigrew amd Moore, 1987). The kinetic measurements show that these cytochrome *c*-mediated electron transfer reactions occur at rates close to the diffusion limit (Margoliash and Bosshard, 1983: Pettigrew and Moore, 1987).

In contrast to the high degree of selectivity it displays in reactions with its protein redox partners, cytochrome *c* transfers electrons fairly indiscriminately with numerous inorganic redox reagents. The second-order rate constants for the reduction of cytochrome *c* by various iron and ruthenium coordination complexes, as well as by dithionite, are all in the range 2-6 X 10^4 $M^{-1}s^{-1}$ (Ferguson-Miller *et al.*, 1979). That all of the inorganic redox reagents react at similar rates with cytochrome *c* suggests that a common reaction mechanism is used. The kinetics of electron transfer are most consistent with an outer-sphere reaction, likely via the exposed edge of the heme group (Timkovich, 1979).

It is notable that in comparison to its reactions with small molecule reagents, cytochrome *c* reacts with enzymic partners with a much higher degree of specificity, and at rates greater by two

to five orders of magnitude. These observations likely indicate that the polypeptide chain of cytochrome *c* provides specific interaction sites for both the recognition of the appropriate macromolecular redox partner, and the optimization of the electron transfer process itself (Dickerson and Timkovich, 1975).

2. Factors which influence heme reduction potential

The reduction potential of the heme iron of cytochrome *c* is of great importance, as it affects the ability of the protein to react with both the preceding and succeeding carrier in the electron transport chain, and in addition may determine the rate of the electron exchanges carried out by this protein. The majority of eukaryotic cytochromes *c* have a reduction potential (E_m) in the range 240–280 mV. This relatively positive reduction potential means that the neutral state of the heme (containing ferrous Fe^{2+}) is stabilized relative to the positively charged state (containing ferric Fe^{3+}). Three independent structural factors which additively influence the reduction potential of a heme group are the chemical nature of (i) the peripheral side chains of the heme pyrrole rings, (ii) the axial ligands to the heme iron, and (iii) the local environment of the heme moiety (reviewed in Mathews, 1985; Marchon *et al.*, 1982). This section discusses these factors with respect to their determining the reduction potential of cytochrome *c*. A fourth factor of possible importance is a conformational change, accompanying a change in oxidation state of the heme iron, in the surrounding polypeptide matrix. By strongly favoring one of the reduced or oxidized conformational forms, the polypeptide chain can potentially influence the relative stabilities of the alternative oxidation states of the heme iron.

The effect of the heme side chain substituents is inductive; electron-withdrawing groups destabilize a positive charge at the heme iron, and thus cause a more positive potential. The reduction potential of the heme iron of Fe-protoporphyrin is ~40 mV higher than that of Fe-mesoporphyrin, in which the vinyl groups are replaced by ethyl groups (Cowgill and Clark, 1952). Thus, saturation of the vinyl groups of protoheme in cytochrome *c* occurring upon formation of the thioether linkages would be expected to lower the reduction potential. However, the *Euglena* and *Crithidia* cytochromes *c*, in which only one of the vinyl groups forms a thioether

bond to the polypeptide chain, have reduction potentials not significantly different from those of more typical cytochromes *c* (Moore and Williams, 1977). In addition, the inductive effect of the cysteinyl sulfurs themselves on the heme reduction potential has not been thoroughly investigated.

The influence of the heme ligands on reduction potential is related to the electron-donating power of the ligating group. More strongly donating groups stabilize a more positive heme iron, and thus give rise to a lower reduction potential (Falk, 1964). Comparison of the reduction potentials of pairs of heme complexes which have imidazole at the fifth coordination site, but which differ in whether a histidine or methionine analogue forms the sixth ligand, has shown that ligation of the more weakly-donating methionine effects a ~160 mV increase in reduction potential (Mathews, 1985). As discussed earlier, the alkaline form of ferricytochrome *c* has a reduction potential ~140 mV lower than that of native cytochrome *c*, although the drop in reduction potential is likely not due entirely to the loss of Met80 ligation. In addition, replacement of Met80 with a His residue lowers the reduction potential of horse cytochrome *c* by ~220 mV (Raphael and Gray, 1989). Therefore, that cytochrome *c* uses as axial ligands a methionine in addition to a histidine side chain undoubtedly contributes to the its fairly high reduction potential.

Three aspects of the local environment of the heme have been shown to influence the reduction potential of the heme iron: the polarity of the heme microenvironment, the degree of solvent accessibility to the heme, and electrostatic interactions between the heme iron and neighboring charged groups. Kassner (1972) has compared the reduction potential of the bipyridine coordination complex of mesoheme methyl ester dissolved in benzene, with that of the corresponding complex of mesoheme (free acid) in aqueous solution. The heme complex in the non-polar solvent has a ~300 mV more positive reduction potential, which has been attributed to the greater electrostatic free energy required to separate, within a medium of low dielectric constant, a monocationic ferric iron centre from its countercharge(s) (Kassner, 1972, 1973; Churg and Warshel, 1986). This proposal undoubtedly accounts at least partially for the high reduction potential of cytochrome *c*, as the heme group of this protein is tightly enveloped within a crevice

lined primarily with non-polar amino acid side chains. Stellwagen (1978) has noted that within a series of heme proteins of known structure, the fraction of the total surface area of the heme excluded from external solvent correlates positively with a high reduction potential. He has therefore suggested that the exposure of the heme to the external medium is a major determinant of the stability of the reduced state of the heme iron, as it reflects the accessibility of the heme to exogenous oxidants. Stellwagen's results can also be interpreted in terms of Kassner's proposed model (Kassner, 1973), since an increase in solvent accessibility to the heme would likely effect a corresponding increase in the polarity of the heme environment (Schlauder and Kassner, 1979). Schlauder and Kassner (1979) have used solvent perturbation absorption difference spectroscopy to measure the extent of solvent accessibility to the heme group of both horse cytochrome *c* and *R. rubrum* cytochrome c_2. Their results suggest that the 60 mV lower reduction potential of the former protein is due in part to the greater extent of heme exposure in this protein. The third environmental factor which influences the reduction potential of cytochrome *c* is electrostatic interactions between the heme iron and neighboring ionizable groups (George *et al.*, 1966). The preponderence of basic side chains on the surface of the cytochrome *c* molecule, and in particular around the exposed edge of the heme group, may act to destabilize a positive charge on the heme iron. A number of experiments have indicated that changes in the charges on protein side chain groups can affect the reduction potential of the heme. Derivitization of single lysine residues with trifluoromethylphenyl isocyanate causes a drop of ~15 mV in the reduction potential of horse cytochrome *c* (H.T. Smith *et al.*, 1977), which Rees (1980) attributes to the effect of neutralization of the positive charge on the ϵ-amino groups. Moore *et al.* (1984) have suggested that in yeast cytochrome *c* deprotonation of His39, which occurs with pK_a of 6.9 and 7.3 in the oxidized and reduced forms of the protein respectively, causes the reduction potential to drop by ~25 mV. Moore (1983) has also proposed that the rear propionic acid group's ionization state, which may in turn be controlled by a salt bridge interaction with the side chain of Arg38, directly influences the reduction potential of the heme iron. The relative importance of the non-polarity of the heme environment, inaccessibility of the heme to solvent, and electrostatic interactions between the heme

iron and neighboring charged groups has been a subject of considerable debate (Stellwagen, 1978; Mathews, 1985; Moore *et al.*, 1986;). However, it is likely that collectively these factors, in controlling the nature of the microenvironment of the heme group buried within the polypeptide matrix of the cytochrome *c* molecule, are the major determinants of the fairly high reduction potential of this protein. It should also be noted that within the cytochromes *c*, the nature of the heme side chains and ligands does not vary (with the exception of the protozoan cytochromes noted above). Thus differences in the reduction potential between variants of cytochrome *c* must be caused by differences in the local environment of the respective heme groups.

3. Proposed mechanisms of electron transfer

Electron transfer reactions involving coordination complexes occur through either an inner-sphere or outer-sphere mechanism. The heme iron of cytochrome *c* is sequestered within the protein matrix, and is fairly inert to ligand exchange. Thus, the majority of electron transfers carried out by cytochrome *c* are via an outer-sphere mechanism. Inner-sphere reactions occur only with some small, highly reactive inorganic redox agents (dithionite, Cr^{3+} ion) and likely require opening of the heme crevice and displacement of the Met80 ligand.

Many outer-sphere mechanisms for the passage of electrons to and from the heme prosthetic group of cytochrome *c* have been suggested (see Harbury and Marks, 1973). The proposed pathways include: (i) via overlap of π-orbitals of the heme group with those of the redox partner, (ii) via the axial ligands, (iii) via the thioether groups linking the polypeptide chain to the porphyrin ring, (iv) via one or more aromatic side chains whose π-orbitals overlap with those of the heme group, (v) via segments of the polypeptide chain, (vi) via bound solvent or small solute molecules, (vii) via quantum-mechanical tunnelling, and (viii) via various combinations of the above.

A number of detailed pathways for cytochrome *c*, mainly of the type (ii) and (iv), have been proposed based on the inspection of the tertiary structure of this protein [see reviews by Salemme (1977) and Dickerson and Timkovich (1975)]. These mechanisms involve conduction of an electron between the heme iron and the surface of the molecule either via a series of

intervening aromatic side chains, or directly through the exposed heme edge. As general mechanisms for the electron-transfer function of cytochrome *c*, these early proposals have been largely discounted by the discovery of naturally-occurring variants of cytochrome *c* bearing incompatible amino acid substitutions at key residues, and by chemical modification studies (see Section I.D.1 below).

4. Redox-state dependent conformational changes

An abundance of experimental evidence indicates that structural differences exist between the oxidized and reduced forms of cytochrome *c* (see reviews by Margoliash and Schecter, 1966; Salemme, 1977). The conformational differences affect a number of surface properties: the two redox forms have differing chromatographic behavior on ion exchange resins, antigenic reactivity, binding affinities for ions, and rates of hydrogen-deuterium exchange. Structural differences between the two redox states are also manifest in differing chemical reactivity (Pande and Myer, 1980) and pK_a's (Robinson *et al.*, 1983) of certain side chain groups, and in differing binding affinities for both enzymic redox partners and phospholipid membranes. A large body of data indicates that the reduced form has a more compact and more stable structure. Ferrocytochrome *c* has a greater resistance to unfolding induced both by chemical denaturants (acid, base, methanol, urea, guanidine hydrochloride) and by heat, to ligand displacement by exogenous substituents, and to proteolytic digestion; and has a lower viscosity in solution (Fisher *et al.*, 1973), a lower compressibility (Eden *et al.*, 1982), and a smaller radius of gyration (by ~1 Å) in low ionic strength solution (as indicated by small-angle X-ray scattering) (Trewhella *et al.*, 1988). The increased stability of reduced cytochrome *c* is likely due in part to the 1000-fold increase in strength of the Met80-heme iron coordinate bond in this form of the protein, although differences between the two redox forms in the interaction between the delocalized π-electrons of the heme molecular orbitals and surrounding side chains within the heme pocket have also been suggested to have a role (Timkovich, 1979).

The techniques of CD and ORD spectroscopy, NMR spectroscopy and X-ray crystallography have been used to evaluate directly the redox-state dependent structural differences.

These differences have been studied in the hope that they will shed light on the mechanism of the electron transfer process. CD and ORD spectroscopy indicate that the oxidation state of the heme iron does not significantly affect the polypeptide conformation of cytochrome *c*, although the environments of one or more aromatic side chains may be perturbed (Myer, 1968). The regions of the polypeptide chain indicated by NMR spectroscopy to differ structurally between the two redox forms of horse cytochrome *c* include residues 39 to 43 and 50 to 60, and also groups in the vicinity of the side chains of His18, Trp59 and Tyr67 (Feng *et al.*, 1990). The redox-state sensitive regions of the cytochrome *c* molecule indicated by NMR spectroscopy and by X-ray crystallography (see below) are in good agreement.

Takano and Dickerson (1980, 1981a,b) have compared the crystal structures of oxidized and reduced tuna cytochrome *c*. Their studies show that the most striking differences between the two redox forms occur in the vicinity of a water molecule buried at the lower left of the heme crevice. Upon oxidation of cytochrome *c* this water molecule, which forms hydrogen bonds to Asn52 ND2, Tyr67 OH and Thr78 OG1, shifts ~1 Å in a direction toward the heme iron. This movement is accompanied by a shift of the tyrosyl ring of Tyr67 toward the rear of the molecule, and an upward rotation of the side chain of Asn52. This set of structural changes is postulated to be triggered by either the attraction of the water molecule to the increased positive charge on the ferric heme iron, or the weakened ability of the Met80 SD atom to donate an electron pair in the hydrogen bond to Tyr67 OH. In addition, there is a small shift of the heme group within its crevice, in a direction toward the surface of the molecule. This shift may also effect movements of adjacent protein groups, including Cys17, His18, Trp59, Phe82 and Ala83. Smaller movements are observed in the regions of polypeptide chain spanning residues 38 to 43 and 47 to 57. Overall, the shifts occurring upon oxidation of cytochrome *c*, in particular the closer approach of the buried water molecule to the heme iron and the more solvent-accessible positioning of the heme group, are consistent with a more polar heme environment required to stabilize the positive charge at the heme centre. There is also the possibility that through interactions with its reductase or oxidase, cytochrome *c* is induced to undergo shifts corresponding to the structural differences

between its two redox forms, thus facilitating either the acceptance or donation of an electron.

The magnitude of these oxidation-state dependent conformational changes, as indicated by the above methods, appears to be too small to account for the observed differences in the solution properties between the two redox forms of cytochrome *c*. The interpretation of many workers (Salemme, 1977; Timkovich, 1979; Moore and Williams, 1980a; Eden *et al.*, 1982; Feng *et al.*, 1989) is that differences in the chemical and physical properties of the two redox states arise from differences in dynamic behavior (the frequency and/or amplitudes of vibrational fluctuations in the structure of the molecule). The larger crystallographic temperature factors of ferricytochrome *c* do indicate that this structure has greater overall flexibility.

5. Modelling of intermolecular complexes between cytochrome *c* and its physiological redox partners

Structural models have been proposed for the bimolecular complexes of cytochrome *c* with cytochrome b_5 and with cytochrome *c* peroxidase (reviewed by Pettigrew and Moore, 1987; Mathews, 1985; Poulos and Finzel, 1984). These models have been constructed by manual adjustment of the relative orientations of the crystallographically determined structures of the individual proteins. The rationale behind this modelling is to optimize both the occurrence of favorable ionic interactions between positively charged groups on cytochrome *c* and negatively charged groups on the redox partner, and the steric fit between the complementary molecular surfaces.

In the optimal complex modelled for cytochrome *c* and cytochrome b_5 (Salemme, 1976), lysines 13, 27, 72 and 79 of the former protein form salt bridges with Glu48, Glu44, Asp60 and an exposed heme propionate, respectively, of the latter protein. The iso-1-cytochrome *c*-cytochrome *c* peroxidase complex (Lum *et al.*, 1987; and also Poulos and Kraut, 1980) contains the following hydrogen bond and salt bridge interactions: Lys5-Asp33, Thr12-Asn87, Arg13-Asp37, Gln16-Gln86, Lys(Tml)72-Asp217, Lys86-Asp37, Lys87-Asp34 and Gly83-His181. In addition, there is a hydrophobic interaction between Ala81 and Leu182, as well as the antiparallel juxtaposition of the N-terminal α-helix of cytochrome *c* with a helical segment of cytochrome *c*

peroxidase. In both these complexes, there is a high degree of complementarity in the surface topographies of the partner protein molecules, which suggests that solvent is excluded from the intermolecular interface. Notably, the heme group of cytochrome *c* is parallel to the hemes of cytochrome b_5 and cytochrome *c* peroxidase in the respective complexes. In the cytochrome *c*–cytochrome b_5 complex, the closest approach of the heme groups of the two proteins is ~8 Å, and the iron–iron interatomic distance is ~16 Å; the corresponding values for the cytochrome *c*–cytochrome *c* peroxidase complex are ~16 Å and ~25 Å, respectively.

Several inferences can be drawn regarding intermolecular electron transfer within the modelled complexes. First, within the interfacial region, the proposed exclusion of solvent and the neutralization of complementary charges would tend to lower the dielectric constant of the medium separating the two redox centres, thereby facilitating the electron transfer process (Salemme, 1977; Poulos and Kraut, 1980). Second, the parallel orientations of the heme groups of the two proteins may serve to enhance interaction between the delocalized π–electron systems of these two groups. This interaction may be of particular mechanistic importance for the cytochrome *c*–cytochrome b_5 complex, in which the distance between heme edges is fairly short (Salemme, 1977). Third, in the cytochrome *c*–cytochrome *c* peroxidase complex, an extended series of planar, conjugated and aromatic groups occur in a parallel arrangement. These groups, which include the hemes of both molecules, Phe82 of cytochrome *c*, and the spatially adjacent His181 side chain of cytochrome *c* peroxidase, is postulated to form an intermolecular electron conduit composed of overlapping π–orbitals (Poulos and Kraut, 1980).

In general, the main features of the proposed complexes are corroborated by experimental evidence from many sources, including chemical modification of lysine side chains of cytochrome *c* (Margoliash and Bosshard, 1983) and of acidic groups on cytochrome b_5 (Mauk *et al.*, 1986) and cytochrome *c* peroxidase (Waldemeyer *et al.*, 1982); differential protection of protein groups upon complexation (Bechtold and Bosshard, 1985); covalent crosslinking of the partner proteins (Waldemeyer and Bosshard, 1985); identification by NMR spectroscopy of the protein groups occurring at the intermolecular interface (Eley and Moore, 1983; Satterlee *et al.*, 1987; Moench *et*

al., 1987); and comparison of the kinetics of reduction by free flavin semiquinones of the free and complexed protein components (Hazzard *et al.*, 1987). However, representation of the interacting protein molecules as forming a static, rigid complex is likely not completely accurate. A number of recent studies suggest that an important requirement for productive interaction between cytochrome *c* and its redox partner is the ability of the two proteins to undergo dynamic movement relative to one another. For example, examination of the electrostatic potential surfaces of cytochrome b_5 and cytochrome *c* suggests that these two proteins can interact in a number of different relative orientations to form electrostatically stabilized complexes (Mauk *et al.*, 1986). Molecular dynamics simulations indicate that within this complex cytochrome b_5 migrates along the surface of cytochrome *c*, and also that the complex transiently contracts, bringing the two heme irons closer together (Wendoloski *et al.*, 1987). Brownian dynamics simulations have yielded similar results for the cytochrome *c*– cytochrome *c* peroxidase complex (Northrup *et al.*, 1988). In crystallographic studies of a 1:1 complex between cytochrome *c* and cytochrome *c* peroxidase, well-defined electron density for the former is not visible, suggesting that cytochrome *c* is positionally disordered with respect to cytochrome *c* peroxidase (Poulos *et al.*, 1987b). In addition, the finding that the maximal rate of electron transfer between ferrocytochrome *c* and ferrylcytochrome *c* peroxidase occurs not at low ionic strength, but an ionic strength ($\mu = 260$ mM) at which electrostatic interactions would be considerably masked, also suggests that flexibility in the complexation interactions is necessary (Hazzard *et al.*, 1988b). Collectively, these studies indicate that conformational flexibility in the intermolecular interactions between the electron transfer partners is required for productive complexation, likely by allowing the two proteins to adopt relative orientations that promote electron transfer, or by decreasing the distance separating the redox centres. The complexes proposed from modelling studies may represent the precursor in which electron transfer geometry is optimal.

C. Amino acid conservation in the cytochromes *c*

Because cytochrome *c* is a protein ubiquitous to all eukaryotes, and can be readily isolated, primary sequences have been determined for a large number (92) of eukaryotic species. The

degree of sequence divergence between the various species of cytochrome *c* has been used to trace the phylogeny of the eukaryotes. A comparison of the available primary structures indicates that cytochrome *c* is a highly conservative protein. The cytochromes *c* from two species whose evolutionary divergence occurred as long ago as one billion years have amino acid sequences that typically differ at no more than 50 of the ~104 positions. In the sequences of eukaryotic cytochromes *c*, there are 21 residues which are absolutely invariant, and an additional 11 which are

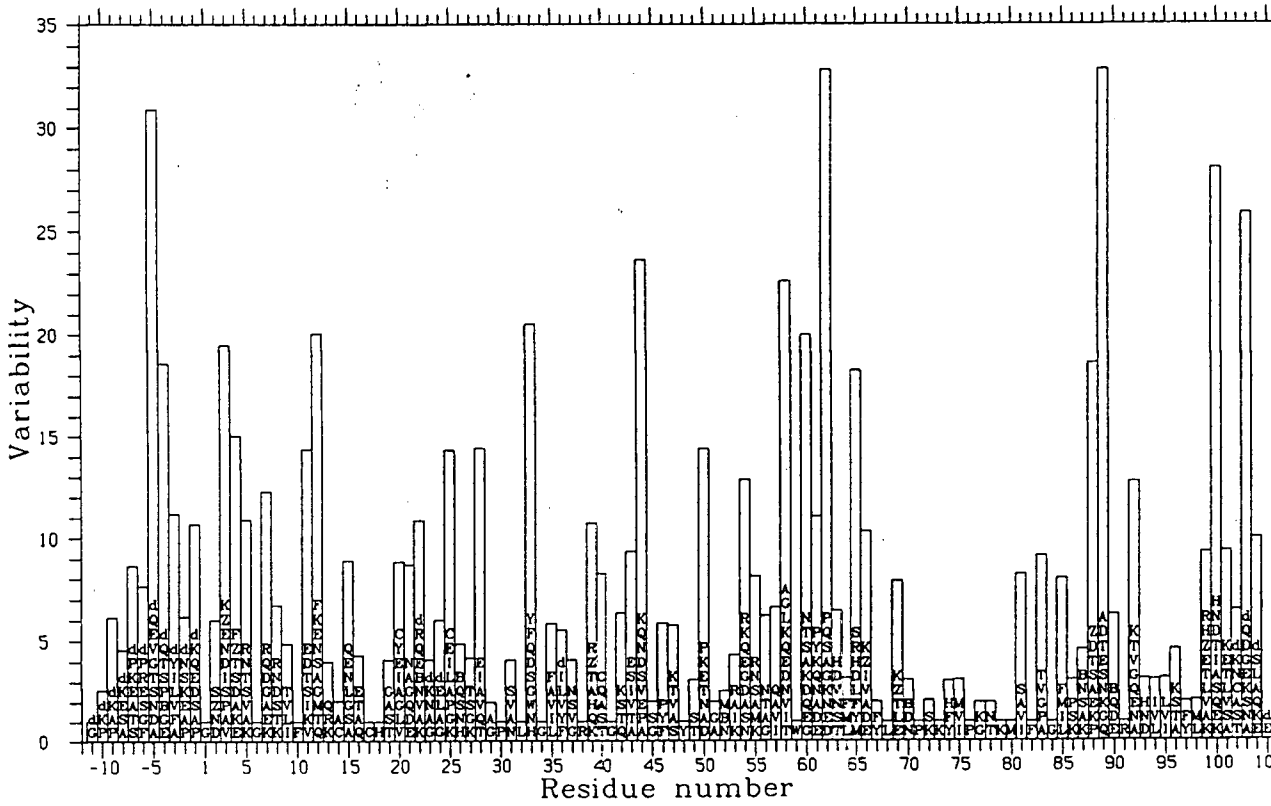

Figure I.3. Degree of amino acid variability along the polypeptide chain of eukaryotic cytochromes *c*. At each position, the variability (Wu and Kabat, 1970) is calculated as the number of different amino acids observed to occur at this position, divided by the frequency of occurrence of the most common amino acid (whose identity is given at the very bottom of the histogram). Other amino acids which occur at each position have also been listed. Note that a lower case 'd' indicates that the designated residue is deleted in some species. The sequences of 92 eukaryotic cytochromes *c* (Hampsey *et al.*, 1986) have been included in this analysis.

always one of two amino acids (see Figure I.3). Only 14 positions, having a variability index (Wu and Kabat, 1970) greater than 15, are highly variable.

The conservative residues in cytochrome *c* very likely have an important role in the function of the molecule. On the other hand, residues which show considerable variability in amino acid identity are not as likely to be critical to the proper functioning of cytochrome *c*. Consideration of the reactions which the cytochrome *c* molecule undergoes during its biosynthetic processing and its functioning in electron transfer suggests several roles that a conserved residue may play. These roles, which may be highly interdependent, include:

(i) serving as a site for post-translational modification of the polypeptide chain, and in particular for covalent coupling of the heme prosthetic group,

(ii) recognition of the receptor responsible for translocation of cytochrome *c* across the outer mitochondrial membrane,

(iii) directing the proper folding of the polypeptide chain, or stabilizing the native, folded state of the cytochrome *c* molecule,

(iv) directly participating in the electron transfer reaction,

(v) creating the appropriate heme environment, and in particular controlling the reduction potential of the heme iron,

(vi) binding of a small molecule effector (e.g. ions, ATP), and

(vii) forming complexation interactions with redox partners of cytochrome *c*.

A role for a conserved residue may be inferred from inspection of the structural environment of that residue within the cytochrome *c* molecule. The results of several such analyses of the invariant residues of cytochrome *c* are summarized in Table I.1. It must be emphasized, however, that functional roles for residues proposed on the basis of amino acid conservatism and structural considerations are tentative unless supported by additional, independent experimental evidence.

D. Structure-function studies in cytochrome *c*

The objective of the study of the relationship between a protein's structure and function is to elucidate the molecular basis of the protein's activity. A powerful approach in the investigation

Table I.1. Roles inferred from structural considerations for the invariant residues in eukaryotic cytochromes *c*

Residue	Structural environment	Inferred role(s)[a]
Gly1	Occurs at rear surface of molecule. Forms no strong intramolecular interactions.	Unknown. Site of acetylation for some mammalian cyt *c*'s
Gly6	Close contact with 97 CA, at interface between N- and C-terminal α-helices	Small volume allows close packing of helices
Phe10	Packed in upper rear of heme pocket	Contributes to hydrophobic heme pocket
Cys17	Front exposed thioether bond to heme group	Covalent heme attachment
His18	Ligand to heme iron	Heme ligand
Pro30	Forms bend in polypeptide chain adjacent to heme group. C=O group hydrogen bonds His18 imidazole.	Generates turn conformation. Stabilizes conformation of His18 imidazole. Contributes to hydrophobic heme pocket
Leu32	Packed in right rear of heme pocket.	Contributes to hydrophobic heme pocket
Gly34	Part of surface β-turn	Protein folding: allows type II tight-turn
Arg38	Electrostatic interaction with rear heme propionic acid group	Stabilizes buried heme propionic acid group. Controls redox potential of heme iron.[b]
Gly41	Part of surface β-turn	Protein folding: allows type II tight-turn
Tyr48	Hydrogen bonds rear heme propionic acid	Stabilizes buried heme propionic acid group
Trp59	Hydrogen bonds rear heme propionic acid	Stabilizes buried heme propionic acid group
Leu68	Packed at left rear of heme pocket	Contributes to hydrophobic heme pocket
Pro71	Junction between 60's and 70's α-helices	Protein folding: directs folding of helices
Lys73	Directed into solvent at left, front side of molecule	Intermolecular interactions
Pro76	Part of surface β-turn	Protein folding: generates turn conformation
Lys79	Forms hydrogen bond to 47 C=O on opposite side of heme crevice	Stabilizes protein conformation
Met80	Ligand to heme iron	Heme ligand
Phe82	At surface of molecule adjacent to heme group	Electron transfer. Maintenance of heme environment
Gly84	Packed against side chains of Phe82 and Leu85	Protein folding: allows otherwise inaccessible conformation of polypeptide chain
Arg91	At left surface of molecule. Forms hydrogen bond(s) to carbonyl groups from 80's loop	Stabilize conformation of 80's loop.

[a]Unless otherwise noted, the inferred roles are according to Dickerson and Timkovich, 1975; Timkovich, 1979; Takano and Dickerson, 1981; and Hampsey *et al.*, 1988, 1986.
[b]Moore *et al.*, 1984.

of structure-function relationships has traditionally been the analysis of variants. Thus the role carried out by a particular amino acid residue in a protein can potentially be established from a characterization of the structural and functional properties of variant proteins in which that residue

has been specifically modified. In particular, a role for a residue proposed on the basis of the inspection of protein structure, or amino acid sequence conservation, can be directly tested.

For a protein system to be amenable to this method of analysis, there are two prerequisites: an accurately known three-dimensional structure for the protein, and means to generate variants of the protein. As discussed above, a fairly extensive data base of cytochrome *c* structures has been available for some time. In the past two decades, a number of methods have been used to generate modified forms of cytochrome *c*. The information that these methods have yielded on the functional roles of specific residues of cytochrome *c* provide the foundation for further studies of structure-function relationships in this protein. Thus, previous work will be thoroughly reviewed in the following sections. It must be emphasized, however, that in none of the studies described has the structure of the variant cytochrome *c* been determined. Thus an assignment of a role to a residue based on the observed altered functional properties of the variant protein in which that residue has been modified must be regarded with some caution, as the modification may have unexpectedly perturbed the conformation of the protein molecule.

1. Chemical modification

a. Heme prosthetic group

Chemical modification involving the heme group has been of three types: complete removal of the heme moiety to yield apocytochrome *c*, removal of the heme iron to yield porphyrin-cytochrome *c* (Fisher *et al.*, 1973), and replacement of the iron by another metal centre. The heme thioether bonds can be cleaved by treatment with silver or mercury salts. Spectroscopic and physicochemical measurements indicate that the resultant polypeptide chain of apocytochrome *c* has a random, unfolded conformation. Removal of the heme iron can be effected by protonating the heme pyrrole nitrogens using anhydrous hydrogen fluoride. The porphyrin-cytochrome *c* appears to retain a native conformation. These results suggest that (i) packing of hydrophobic groups against the faces of the heme group, and possibly the interaction of polar groups with the heme propionic acid groups, are important in determining the overall folding of the polypeptide chain (Dickerson and Timkovich, 1975), and (ii) the coordinate bonds formed to the His18 and Met80

side chains are not critical determinants of the protein fold. Metallocytochromes *c*, in which the heme iron is replaced by another metal ion (Cu, Zn, Mn, Ni, Co, or Sn), can be prepared from porphyrin-cytochrome *c*, but have not been extensively characterized (Erecinska and Vanderkooi, 1978; Meade *et al.*, 1989).

b. Histidine and methionine

Early experiments in which methionine and histidine side chains were modified were directed at identifying the heme ligands (Ferguson-Miller *et al.*, 1979). Methionine can be carboxymethylated by treatment with a haloacetic acid. Derivatization of Met12, 33 or 65 had no substantial effects; however, derivatization of Met80 greatly perturbed the cytochrome *c* protein, thus suggesting that Met80 is a heme ligand. Similarly, modification of histidine residues using diazonium-1H-tetrazole had identified His18 as a heme ligand. The altered cytochrome *c* in which ligation of Met80 to the heme is disrupted by either side chain carboxymethylation or oxidation (to the sulfoxide) is auto-oxidizable, can be bound by exogenous heme ligands, and fails to react with both cytochrome *c* oxidase and reductase (Myer *et al.*, 1980). More recently, cytochrome *c* carboxymethylated at Met65 and Met80 has been shown by CD spectroscopy to have no alterations in overall conformation, but is more susceptible to denaturation by urea and has a greatly decreased reduction potential (-215 mV) (Santucci *et al.*, 1987). Collectively, these results indicate that coordination of Met80 to the heme iron is required for maintenance of both the structural integrity of the heme pocket, and a high reduction potential. Note that the inhibited activity of the carboxymethylated cytochrome *c* with oxidase and reductase does not establish that the side chain of Met80 plays a direct role in electron transfer, as the inhibition may be the result of either a disrupted polypeptide chain conformation, or the perturbed reduction potential.

In recent experiments, the side chains of histidine residues have been modified by coordination of a redox centre, such as ruthenium pentaammine. These studies are aimed at determining the influence of the intervening protein medium on electron transfer between the ruthenium and heme iron redox centres. They appear to indicate that electron transfer that does not occur via the exposed heme edge is relatively slow, and that the protein envelope imparts a

preferred directionality to electron transfer between the Ru and Fe sites (Nocera *et al.*, 1984; Bechtold *et al.*, 1986).

c. Tryptophan and tyrosine

Cytochrome *c* possesses an invariant tryptophan and several conservative tyrosine residues. Because these aromatic side chains appear to form conduits from the heme group to the surface of the molecule, chemical modification has been used to probe their participation in electron transfer. Proteins specifically modified at Trp59, Tyr67 and Tyr74 have been generated (reviewed by Ferguson-Miller, 1979; Myer *et al.*, 1980). The indole ring of Trp59 has been modified through formylation of the amide function by formic acid, and through oxidation by N-bromosuccinimide. The phenolic ring of Tyr67 can be nitrated by tetranitromethane, or iodinated using I_2/KI_3. The modification of these two internal residues was found not only to impair activity with oxidase and reductase, and reducibility by ascorbate, but also to cause the loss of Met80 ligation to the heme, a significant lowering in reduction potential, and an opening of the heme crevice as indicated by susceptibility of the heme iron to autooxidation and binding by CO. Therefore, these derivatives do not confirm the direct participation of Trp59 and Tyr67 in the electron transfer reaction, but do implicate these residues in maintaining the native conformation of the cytochrome *c* molecule. In particular, owing to its participation in a hydrogen bond to the Met80 side chain, the side chain of Tyr67 has been proposed to function in stabilizing the Met80 coordinate bond to the heme iron. The side chain of the surface residue Tyr74 can be preferentially iodinated by the enzyme lactoperoxidase. Normal reactivity of this derivative with both oxidase and reductase has shown that Tyr74 is not required for the electron transfer activity of cytochrome *c*.

d. Lysine and arginine

The interaction of cytochrome *c* with its redox partners is known to be electrostatic, as it is inhibited by high ionic strength and by the presence of polylysine. Intermolecular complexation is thought to be mediated by complementary ionic interactions between lysine side chains on cytochrome *c*, and negatively charged acidic groups on the other protein or on the phospholipid

membrane. Early experiments (reviewed by Dickerson and Timkovich, 1975) in which lysine residues were non-selectively acetylated, succinylated, trifluoroacetylated, or trinitrobenzylated confirmed the importance of lysines in the interaction of cytochrome *c* with its redox partners: the elimination of only two to six positively-charged ϵ-amino groups was sufficient to completely inhibit activity with cytochrome *c* oxidase. In contrast, modifications which preserve the positive charge on the lysine side chains, such as guanidination by O-methylisourea or acetimidylation by methyl acetimidate of ϵ-amino groups (Wallace and Harris, 1984), have little effect.

More recent studies have employed derivatives of cytochrome *c* singly modified at various lysine residues by trifluoroacetylation, trifluoromethylphenylcarbamolyation, or 4-carboxy-2,6-dinitrophenylation (Margoliash and Bosshard, 1983). These studies have shown that the lysines of greatest importance in the binding of horse cytochrome *c* to all redox partners tested (including cytochrome *c* oxidase, cytochrome *c* reductase, purified cytochrome c_1, yeast cytochrome *c* peroxidase, sulfite oxidase, cytochrome b_5 and yeast cytochrome b_2) occur at positions 8, 13, 27, 72, 86 and 87. A derivative of yeast (*Candida krusei*) cytochrome *c* in which the side chain of Arg13 is converted to a bisphenylglyoxal guanidinium compound also has reduced activity with reductase and oxidase (Margoliash *et al.*, 1973). These critical basic residues encircle the exposed edge of the heme group, and outline a region on the surface of the cytochrome *c* molecule which likely serves as the common enzymic interaction domain. The centre of the proposed interaction domain comprises several hydrophobic groups from residues 81 and 82, and also the positive end of the calculated dipole of the molecule.

The derivatives of cytochrome *c* modified at single lysine residues have also been used to map the preferred sites of reaction between cytochrome *c* and a number of small-molecule redox agents, including $[Fe(CN)_6]^{3-}$, $[Co(phen)_3]^{3+}$ (phen=1,10-phenanthroline), $[Fe(EDTA)]^{2-}$, and $[Co(sep)]^{2+}$ (sep=sepulchrate, 1,3,6,8,10,13,19-octa-azabicyclo[6.6.6]icosane) (Butler *et al.*, 1983; Armstrong *et al.*, 1986). These studies also identify the solvent-accessible edge of the heme as the site of electron exchange between cytochrome *c* and non-protein coordination complexes. The experiments in which lysine residues of cytochrome *c* have been chemically modified thus clearly

demonstrate that the exposed edge of the heme group serves as the most favorable pathway for electron movement to and from the heme iron.

2. Naturally occurring variants from different species of organisms

The cytochromes *c* from different species of organisms constitute a set of naturally-occurring variants of this protein. Historically, an amino acid substitution observed in a newly-sequenced species of cytochrome *c* has been useful in disproving a proposed mechanism of electron transfer involving amino acid residues previously thought to be invariant (Dickerson and Timkovich, 1975). More recently, amino acid differences that occur between homologous cytochromes *c* have been utilized in the assignment of resonances in NMR spectra to specific amino acid side groups (Williams *et al.*, 1985a). The various cytochromes *c* available have commonly been used in assessing the degree of cross-reactivity between the cytochrome *c* of one species, and the cytochrome *c* reductase or oxidase of another (Timkovich, 1979). Protein partners from species as diverse as horse and the bacteria *R. rubrum* and *P. denitrificans* have been found to react with each other. This cross-reactivity indicates that the cytochromes *c* from these species possess similar domains for interaction with redox partners, and comparison of their structures implicates the conserved collar of positively charged side chains around the exposed heme edge (Salemme, 1977). Attempts have been made in attributing differences in binding affinity to the presence or absence of specific lysine residues (Ferguson-Miller *et al.*, 1976; Errede and Kamen, 1978). Amino acid compositions have also been correlated to the overall structural stability and resistance to denaturation of eukaryotic cytochromes *c* (Nall and Landers, 1981; Zuniga and Nall, 1983). However, these latter analyses are inconclusive due to the large number of amino acid differences between the proteins considered. In general, the use of naturally occurring variants in the study of structure-function relationships in cytochrome *c* has been hampered because there are few pairs of proteins which differ by only a single amino acid substitution.

3. *In vivo*-generated mutants

Because yeast is a unicellular microorganism, and is thus easy to manipulate both biochemically and genetically, it has historically served as the eukaryotic subject of experiments in classical molecular genetics. Sherman and his coworkers have generated numerous single-site amino acid replacements in yeast iso-1-cytochrome *c* (reviewed in Hampsey *et al.*, 1988, 1986). The earliest variants were isolated by direct screening of mutagenized yeast for strains deficient in heme proteins. Assays for heme proteins used were *in vivo* spectroscopy (see below), or benzidine staining as a means for detecting heme-catalyzed peroxidase activity. Later, selection methods were used which were more expedient, and which also allowed strains that contained functionally defective cytochromes *c* to be isolated. The selection involved culturing of the yeast in the presence of chlorolactate, which inhibits the growth of cells, such as those containing functional cytochrome *c*, capable of metabolizing lactate. Functional variants of iso-1-cytochrome *c* were then obtained by selection for growth on lactate of revertants of the various mutants. The majority of the variants obtained have not been extensively characterized. Typically, the yeast strains containing the variant cytochromes *c* were analyzed in two ways: (i) by *in vivo* spectroscopy of the Cα absorption band, which gives an indication of both proper covalent attachment of the heme, and the level of cytochrome *c* protein, and (ii) by growth on lactate medium, which gives an assessment of functionality of the variant protein.

In total, 177 different single-site amino acid replacements at 60 distinct positions along the polypeptide chain of iso-1-cytochrome *c* have been generated (Hampsey *et al.*, 1988). (It should be noted, however, that 69 of these are of limited interest for structure-function studies, as they occur in the highly variant N-terminal nine residues.) In addition, numerous truncations and deletions at the N-terminus, an internal deletion of residues 40 through 53, and multiple-site mutants have also been obtained. Of the single amino acid replacements observed, 30 occurring at 16 sites have been found to severely disable iso-1-cytochrome *c* activity. Hampsey *et al.* (1988, 1986) have suggested roles for the residues occurring at these sites in the wild-type protein. The only replacements found to cause a complete absence of iso-1-cytochrome *c* protein occurred at

Cys14, Cys17 and His18. Thus these residues, which form covalent bonds to the heme group, are inferred to have an essential role in attachment of heme to the apoprotein, or in import of cytochrome *c* into the mitochondrion. The remaining amino acid replacements yielding non-functional iso-1-cytochromes *c* resulted in subnormal, but detectable, intracellular levels of the protein. These replacements occur at residues which are inferred to have a role in either forming the hydrophobic heme pocket (Leu32, Trp59, Tyr67, Leu94 and Leu98), or directing the folding of the polypeptide chain (Gly6, Gly29, Pro30 and Pro71).

Interestingly, the collection of variant iso-1-cytochromes *c* includes several functional proteins that bear amino acid substitutions at evolutionarily invariant residues. The nature of the amino acids which can functionally replace Pro71 and Trp59 provides clues as to the role of these two residues in the wild-type protein. Pro71 occurs at the junction between two contiguous *a*-helices in the iso-1-cytochrome *c* molecule. Ernst *et al.* (1985) have shown that the Leu71 variant is non-functional, and that the absence of revertants containing Tyr, Glu, Gln, Lys, Phe and Arg at this site suggests that these position-71 variants are also non-functional. In contrast, revertant iso-1-cytochromes *c* containing Val71, Thr71, Ser71 or Ile71 are functional (but have moderately reduced activities). Based on a theoretical analysis, Ernst *et al.* have determined that whereas the permissible residues have both an appropriate steric size and a propensity to adopt the backbone conformation observed for Pro71 in wild-type iso-1-cytochrome *c*, the non-permissible residues do not meet these criteria. Therefore, they conclude that the role of Pro71 is to direct the folding of the polypeptide chain. Trp59 occupies the lower rear of the hydrophobic heme pocket, and forms a hydrogen bond to the rear propionic acid group of the heme. Schweingruber *et al.* (1979) have found that variants with Gly, Ser or Cys at position 59 are non-functional, and from the absence of corresponding functional revertants, infer that variants with Gln, Ala, Thr, Pro and Val at this position are as well. However, revertants with Phe, Tyr and Leu59 are completely to partially functional, although less stable structurally. These results suggest that the primary role of Trp59 is to provide a large hydrophobic group for lining the heme pocket, and that the hydrogen bond between the heme propionic acid and the Trp59 side chain is not absolutely

required. This interpretation is supported by the complete to partial functionality of second-site revertants in which Gly, Ser or Cys are retained at position 59, but are accompanied by a phenylalanine at the spatially adjacent position 40. The other substitutions observed at evolutionarily invariant residues include Tyr48 to Lys, Lys73 to Glu, Pro76 to Leu, and Arg91 to Gly or Leu. The degrees of stability and functionality of the variant proteins have not been reported, nor has the spectrum of amino acid residues permissible at these positions been determined. Thus these mutations do not provide definitive information on the roles of these invariant residues. However, that certain amino acid substitutions are tolerated at evolutionarily invariant residues indicates that functional requirements at these residues may be less strict than commonly assumed, and that evolutionary invariance does not necessarily imply functional invariance (Hampsey *et al.*, 1986, 1988).

Notably, with the exception of two replacements by proline residues, all amino acid substitutions that substantially impair function occur at residues that are evolutionarily invariant, or highly conserved. In contrast, at non-conservative positions in the amino acid sequence, a large number of replacements having minimal effects on the function of iso-1-cytochrome *c* can be accommodated. At a given position along the polypeptide chain, tolerance of single-site amino acid substitutions in iso-1-cytochrome *c* agrees well with variability in amino acid identity between different species of cytochrome *c*. Thus, together the two methods of analysis firmly establish that the invariant residues are critical to the proper functioning of cytochrome *c*, whereas other regions of the polypeptide chain have less important roles.

Several single-site amino acid substitutions have been useful in the analysis of intermolecular complexation reactions and structural stability in cytochrome *c*. These are described in the remainder of this section.

Hazzard *et al.* (1988a) have studied substitutions at Arg13, Gln16 and Lys27. Their results support the proposed involvement of these three residues in forming complexation interactions with cytochrome *c* peroxidase (Poulos and Kraut, 1980). In comparison to the wild-type protein, there is an increase for the Ile13 and Gln27 variants, and a decrease for the Lys16 variant, in the

accessibility of the cytochrome *c* heme to reduction by flavin within the bimolecular complex between cytochrome *c* and cytochrome *c* peroxidase. The reduction of the ferryl radical form of cytochrome *c* peroxidase by iso-1-cytochrome *c* was also affected, with the Ile13 substitution increasing the reaction rate, and the Lys16 and Gln27 substitutions decreasing the rate. A Ser16 substitution appears to have little effect on either complexation or electron transfer, suggesting that the hydrogen bonding interactions formed normally by Gln16 can be duplicated by a serine at position 16.

More extensive studies of amino acid substitutions at the highly conserved Lys27 have examined the importance of this residue to iso-1-cytochrome *c* stability and reactivity with redox partners. Variant iso-1-cytochromes *c* bearing Leu27, Gln27, Trp27 or Tyr27 substitutions were found to support the growth of yeast in lactate medium at 85% of normal for the former two, and at 65% of normal for the latter two. The Leu27 and Gln27 variants appear to be slightly, and the Tyr27 and Trp27 variants substantially more susceptible to denaturation both by guanidine hydrochloride and by heat (Hickey *et al.*, 1988). These results suggest that Lys27 has a role in conferring stability to the iso-1-cytochrome *c* protein. The Leu27 and Gln27 variants and the wild-type protein all react at very similar rates in single-turnover assays with cytochrome b_5 and cytochrome *c* peroxidase, and in steady state assays with cytochrome b_2, cytochrome *c* peroxidase and cytochrome *c* oxidase (Das *et al.*, 1988). The Gln27 variant was actually shown to bind cytochrome *c* peroxidase with 10 times the normal affinity. The finding that replacements of Lys27 with uncharged side chains has little effect on reactivity of iso-1-cytochrome *c* with its physiological redox partners is contrary to expectations based upon evidence from chemical modification and modelling studies. The interpretation proposed by Das *et al.* is that although interactions formed by Lys27 do not significantly contribute to the binding energy within the complex between cytochrome *c* with its redox partner, Lys27 does have a role in accelerating electrostatically intermolecular complexation.

Further studies on the series of functional variants at position 71 (Ramdas *et al.*, 1986; Ramdas and Nall, 1986) suggest a role for Pro71 in conferring stability to the cytochrome *c*

protein. The visible absorption spectra of the Thr71 and Ile71 variants have a reduced intensity of the 695nm band and small shifts in the Soret band, indicating that these variants have a weaker Met80–heme iron coordinate bond. Characterization of behavior of the iso–1–cytochromes *c* upon treatment with guanidine hydrochloride shows that the Val71, Thr71 and Ile71 variants have an increased susceptibility to denaturation and less cooperative unfolding transitions. The three variants also show enhanced rates for the fast phase of unfolding, a result consistent with the interpretation of Ernst *et al.* (1985) (see above) that Pro71 has a role in influencing polypeptide chain folding.

4. Semi–synthesis

The technique of semi–synthesis (Offord, 1981, 1987), in which a naturally occurring protein is partially fragmented, the polypeptide fragment containing the residue of interest is specifically modified, and then the complement of fragments is reconstituted, has been used to generate numerous variant cytochromes *c*. Horse cytochrome *c* has been most frequently used as the source of starting materials for cytochrome *c* semi–synthesis. The most commonly employed fragments of this protein are: the cyanogen bromide cleavage products (1–65)H, (66–104), (66–80), and (81–104); and the peptides resulting from limited tryptic digestion, (1–38)H and (39–104), as well as (1–25)H and (28–38), and (1–53)H and (54–104) (where H denotes the fragment containing the covalently bound heme group). Prior to its use in reconstitution, a fragment may be modified by chemical derivitization, or by stepwise truncation and resynthesis at its termini to produce small–scale amino acid deletions, substitutions, or insertions. Alternatively, a fragment can be replaced by one from a homologous protein, or can be completely synthesized to create any sequence desired. The reassembly of the set of polypeptide fragments may or may not involve reformation of the original peptide bonds. Certain fragments of the polypeptide chain of cytochrome *c* will spontaneously form a non–covalent complex which has conformational and functional properties very similar to those of the native protein molecule. This property is shared by several other small, robust proteins, and in the case of cytochrome *c*, is attributed to the dominant role of the heme group in nucleating the folding of the polypeptide chain (Harbury,

Table I.2. Summary of site-specific modifications of cytochrome *c* generated using semi-synthesis

Modifications made[a]	Reconstitution[b]	Characterization of properties[c]	Comments / Interpretations	Ref.[d]
Pro30 Substitution by Gly **Leu32** Substitution by Val, Phe, Ile or Nva **Gly34** Substitution by Ala or Ser **Leu35** Substitution by Ile, Lys or Thr	Non-covalent complex of (1-25)H· (28-38)*· (39-104)	All of the substitutions at residues 30, 32 (except Nva), and 34 decrease both the binding affinity of the fragment (28-38) to the (1-25)H·(39-104) complex, and the stability of Met80 coordination to the heme iron. The Gly30 variant shows increased autooxidizability, and inhibited activity with lactate dehydrogenase. In the Nva32 variant, the ligation of Met80 to the heme iron is weakened, and reactivity with cytochrome b_2 is decreased. The Ile35 substitution has only slight effects.	Both Leu32 and Leu35 line the hydrophobic heme crevice. Leu32 is shown to be particularly sensitive to replacements (by other non-polar side chains) and is proposed to form specific intramolecular interactions required to stabilize the folded state of the protein molecule (in particular the Met80-Fe coordinate bond). In this regard, Pro30 and Gly34 are also important, but Leu35 less so. In comparison to the other complexes, for the Gly30 and Nva32 variants, binding of the (28-38) fragment is less markedly increased in the reduced relative to the oxidized protein. Leu32 and Pro30 are thus implicated in providing the cytochrome *c* molecule rigidity in the reduced state, and may influence the reduction potential of the protein by stabilizing this conformational state.	1, 2
Arg38 Guanidinium group modified using pentane-2,4-dione; cyclohexane-1,2-dione; or camphorquinone-10-sulfonic acid.	Covalent coupling of (1-65)H and (66-104)	E_m~170mV. Loss of 695nm band. Very low biological activity.	Arg38 forms a number of hydrogen bonds which link the floor (residues 36 to 59) and the main body of the molecule, and also interacts with the buried rear propionic acid group of the heme. The addition of bulky substituents would disrupt these interactions. Thus Arg38 is inferred to have a structural role in stabilizing the molecular conformation.	3

Deletion, substitution by Gly. Restoration of Arg.	Non-covalent complex of (1-37)H· (39-104) or (1-37)H· (38-104)	Complexes possess 695nm band. For Arg38, E_m~220mV and 70% biological activity. For Gly38 substitution and the deletion, E_m<150mV and ~15% biological activity. The Arg38 complex has a greater resistance to further digestion by trypsin.	As above, the inferred role of Arg38 is to provide binding energy for the bottom loop of the molecule. Loosening of this loop results in a greater exposure of the heme, and thus a lowered reduction potential. This interpretation is supported by the properties of the complex (1-38)H· (39-104) (E_m=150mV and ~25% biological activity). In this complex, the break in the polypeptide chain is such that Arg38 is retained in the main body of the molecule, where it cannot act in attaching the floor to the main body of the molecule.	4. Also 5.
Substitution by Lys and Gln	Non-covalent complex (1-37)H· (38-104)	Reduction potential and biological activities, of Lys38: 166mV and 26%; of Gln38: 148mV and 19%; and of Arg38: 215mV and 55%.	Lys can provide a counter-charge for the heme propionate, but does not fully restore a normal reduction potential or electron transfer activity. The results therefore suggest that the extensive hydrogen-bonding capacity of the Arg side chain is specifically required at position 38.	6
Lys39				
Deletion or substitution by Orn	Non-covalent complex (1-38)H· (40-104), or (1-38)H· (39-104)	Biological activity of both variant complexes is not significantly altered from that of the (1-38)H· (39-104) complex having Lys at position 39.	Residue 39 is fairly strongly conserved, with Lys39 occurring in all animal species. The studies on these non-covalent complexes are taken to support a role of Lys39 in stabilizing the conformation of the bottom Ω-loop of the protein molecule. The Glu39 substitution, which occurs within in the stable complex having Arg38 within the Ω-loop (see above), has the greatest effect on the functional properties. The other substitutions, which occur within a complex already destabilized by the absence of Arg38 within this Ω-loop, have lesser effects.	5
Substitution by Glu	Non-covalent complex (1-37)H· (38-104)	Biological activity and reduction potential decreased to 34% and 190mV, respectively, from 55% and 215mV in the (1-37)H·(38-104) complex having Lys39.		6
Substitution by Thr or Tyr	Non-covalent complex of (1-39)H*· (40-104)	Biological activity decreased, from 25% in the [Lys39](1-39)H· (40-104) complex, to ~19% in both variant complexes.		6
Deletion, or substitution by Ala, Gly, Phe or trifluoro-acetylated-Lys	Covalent religation of (1-39)H and (41-104), or (1-39)H and (40-104)	All variants have reduced stability (pK for alkaline isomerization lowered from 9.7 to ~8.2), decreased reduction potentials (~190mV), and impaired biological activity (20-30%).	The results are again interpreted in terms of the inability of the substituted amino acids to form an interaction which stabilizes the conformation of the bottom Ω-loop. It is proposed that the structural destabilization results in the lowered reduction potentials, which in turn cause the decrease in biological activity.	7

Thr40				
Substitution by Lys, Val and Phe	Covalent religation of (1–40)H and (41–104)	All variants have a decreased reduction potential (~190mV) and impaired biological activity (~25%). Val40 and Phe40 variants have pK's for alkaline isomerization of 8.0 and 8.3, respectively.	In eukaryotic cytochromes *c*, residue 40 is most frequently threonine or serine. This pattern of conservation and the experimental results suggest that the side chain of residue 40 forms a hydrogen bond which is important in stabilizing the conformation of the bottom Ω–loop of the protein molecule.	7
Asn54 and Lys55				
Substitution by Ala54 and Lys55 deletion	Covalent religation of (1–54)H and (56–104)	Biological activities for these variants, and also for an Ala insertion following Lys55, are decreased to ~20%.	Residues 54 through 56 are highly variable in identity. The loss of function is interpreted in terms of structural perturbation of the bottom Ω–loop of the protein molecule.	7
Met65				
Substitution by Hse	Covalent coupling of (1–65)H and (66–104)	Normal electron transfer activity, reduction potential and spectral properties. NMR studies reveal no significant structural perturbations.	Homoserine at position 65 does not alter the structural or functional properties of cytochrome *c*. These results are consistent with the high variability of residue 65. (Serine occurs here in several species.)	8, 9
Glu66				
Substitution by Gln, Lys and Nva	Covalent coupling of (1–65)H and (66–104)	Nva66: no 695nm band; only 20% biological activity. Gln66 and Lys66: slightly elevated reduction potential; decrease of 0.7 in pK for alkaline isomerization; >90% biological activity. Lys66: decreased ability to bind ATP.	Glu66 is a fairly conservative acidic residue occurring in the amphipathic 60's helix. The Nva66 variant demonstrates that hydrophobicity cannot be tolerated at position 66. Inferred roles of Glu66: inducing α–helical conformation; providing a salt bridge to stabilize protein against denaturation; ATP binding.	10
Tyr67				
Substitution by Phe and FPhe	Covalent religation of (1–65)H, (66–79)* and (80–104)	Phe67 variant retains 56% of normal activity in succinate oxidase system; but FPhe67 is completely inactive.	Highly conserved Tyr67, together with Thr78, Asn52 and an internal water molecule, form a network of hydrogen bonds at the lower left of the heme crevice, with Tyr67 directly hydrogen bonding the Met80 ligand. Tyr67 is also part of the left aromatic channel. The results show that Tyr67 is not essential for electron transfer activity. An electronegative fluorine substituent may perturb the electronic environment of the heme.	11
Substitution by FPhe	Covalent religation of (1–65)H, (66–79)* and (80–104)	E_m = 199mV. Rate of electron transfer in ascorbate–cytochrome *c* oxidase system increased.	The electron transfer activity determined for the FPhe67 variant is in disagreement with results from ref. 11. The significance was not discussed.	12

Substitution by Phe	Covalent religation of (1–65)H and (66–104)*	E_m = 199mV. 50% decrease in biological activity. Shift in position of α- and β-bands in absorption spectrum. Increased resistance to denaturation by alkaline pH (pK 10.7 vs. 9.3) and temperature (T_m = 107°C vs. 64°C). Protein is oxidized rapidly by molecular oxygen.	That the Phe67 variant has an increased resistance to denaturation shows that the major role of Tyr67 is not in structural stabilization. The altered spectral properties suggest that Tyr67 has a direct influence on the π-electron system of the heme, which is inferred to affect the reduction potential. Tyr67 may also have a role in preventing the access of oxygen to the heme iron, and in establishing the reduction potential of the protein. The lowered electron transfer activity is attributed to the decreased reduction potential.	13
Lys72, Lys73 and Lys79				
Substituted individually (and also in groups) by acetylated-Lys	Covalent religation of (1–65)H, (66–79)* and (80–104)	All variants possess the 695nm band. Modifications of Lys72 and Lys73 cause a 2.5-fold increase in K_m for activity with cytochrome *c* oxidase. Modification of Lys79 has only a minor effect.	The results are consistent with chemical modification studies on the intact protein. Lys72 and Lys73 are confirmed to be part of the binding domain on cytochrome *c* for cytochrome *c* oxidase.	14
Tyr74				
Substitution by Leu	Covalent religation of (1–65)H, (66–79)* and (80–104)	Activity slightly lowered in both succinate oxidase (by ~40%) and ascorbate–cytochrome *c* oxidase systems. Greater thermolability, as the 695nm band is lost at a lower temperature.	The highly conserved Tyr74, Trp59 and Tyr67 form an aromatic cluster at the left side of the molecule. The results show that Tyr74 is not essential for electron transfer activity, but suggest that it may have a role in stabilizing the conformation of the left side of the molecule.	11, 14
Thr78				
Substitution by Val	Covalent religation of (1–65)H, (66–79)* and (80–104)	E_m = 170mV. Loss of 695nm band. Electron transfer activity in ascorbate–cytochrome *c* oxidase assay is decreased.	Thr78 is a member of the hydrogen bonding network at the lower left of the heme crevice. The significant perturbation of functional properties in the Val78 variant suggests that Thr78 has a role in structural stabilization.	12
Substitution by Asn and Aba	Covalent coupling of (1–65)H and (66–104)*	Asn78 has a slight, and Aba78 a large decrease in intensity of the 695nm band. E_m(Aba78) = 251mV; E_m(Asn78) = 266mV. Both variants, but Aba78 in particular, show a large decrease both in stability against alkaline and thermal denaturation, and in biological activity.	These results also implicate Thr78 in stabilization of the heme crevice. Note that Asn78 does occur in one eukaryotic cytochrome *c* (from *Chlamydomonas*).	13

Met80				
Substitution by His	Covalent religation of (1–65)H and (66–104)*	The absorption spectrum of this variant is fairly similar to that of the native protein. The reduction potential is lowered to 40mV.	Met80 has a role in establishing the fairly high reduction potential of the heme iron.	15
Ile81				
Substitution by Ala, Leu and Val	Covalent religation of (1–65)H, (66–79) and (80–104)*	The reduction potentials of these variants are not significantly perturbed. Although the activity of the variants in mediating electron transfer from ascorbate to cytochrome *c* oxidase is not greatly altered, some differences in K_m are apparent.	Replacement of Ile81 with other aliphatic side chains does not affect the reduction potential. Slight differences in activity with cytochrome *c* oxidase suggest that residue 81 has a role in interaction with oxidase.	16
Phe82				
Substitution by Leu	Covalent religation of (1–65)H, (66–79) and (80–104)*	E_m = 265mV. Activity in mediating electron transfer from ascorbate to cytochrome *c* oxidase is not greatly affected.	Phe82 has been proposed to participate directly in the electron transfer reaction. However, these results show that it is not absolutely required. The role of the side chain at this position may be to provide hydrophobicity.	12
Ala83				
Substitution by Pro	Covalent coupling of (1–65)H and (66–104)*	Slightly decreased susceptibility to denatuation by alkaline pH (pK 9.0 vs. 9.3) and temperature (T_m 53° vs. 65° C). Slightly increased reduction potential (266mV), but decreased biological activity (~80%).	Ala83 and Pro83 are highly conserved in animals and plants respectively. The Ala to Pro substitution has minor effects. The decrease in electron transfer activity suggests that residue 83 has a role in interaction with cytochrome *c* reductase.	13
Arg91				
Guanidinium group modified using pentane–2,4–dione; cyclohexane–1,2–dione; or camphorquinone–10–sulfonic acid.	Covalent coupling of (1–65)H and (66–104)	No effect on reduction potential, 695nm band, or biological activity.	Addition of bulky substituents would disrupt the hydrogen bonds between Arg91 and the 80's region of polypeptide chain. Thus, the results do not support the proposed role of Arg91 in stabilizing the conformation of the 80's loop.	3
Guanidinium group modified using pentane–2,4–dione	As above	Decreased ability to bind ATP.	Arg91 is implicated as a component of an ATP binding site on the surface of the cytochrome *c* molecule.	17

Tyr97				
Substitution by Leu	Covalent coupling of (1–65)H, (66–79) and (80–104)*	E_m = 319mV. Slightly reduced activity in mediating electron transfer from ascorbate to cytochrome *c* oxidase.	Highly conserved Tyr97 and Phe10 form an aromatic channel at the right side of the molecule. The results show that Tyr97 is not essential for electron transfer activity. The interaction of these aromatic side chains may be an important structural element, as the reformation of the 65–66 peptide bond is inhibited in this complex.	12
Residues 40 to 55				
Deletion	Covalent religation from (1–39)H and (56–104)	Variant protein possesses 695nm band; is highly unstable (pK for alkaline isomerization 7.4); has a decrease in reduction potential by 130mV; and has only 15% biological activity.	Residues 40 to 55 form the bottom Ω-loop of the cytochrome *c* molecule. It is proposed that the role of this loop is to shield the bottom edge of the heme from solvent and thereby control the reduction potential of the protein. The deletion of the sequence 40 to 55 generates a protein which is an analog of the small bacterial cytochromes *c*.	18
Residues 60 to 103 or 104				
Chimera of different species of cytochrome *c*	Covalent coupling of (1–65)H of horse, with (65–104) of cow, or (65–103) of yeast cytochrome *c*	All chimeric proteins have full biological activity. The measured reduction potentials are: horse, 255; cow, 257; horse–yeast, 263; cow–yeast, 268; and yeast, 270mV.	Between yeast, cow and horse cytochromes *c*, 10 amino acid differences occur in the (66–103) fragment. The variant residues are inferred to have a role in controlling the reduction potential of the protein.	19

---

[a]The normally-occurring amino acid residues, given in bold face, are those in horse cytochrome *c*. All modifications described have been carried out using this protein. Abbreviations used: Aba = L-*α*-aminobutyric acid, FPhe = L-fluorophenylalanine, Hse = L-homoserine, Nva = L-norvaline, Orn = L-ornithine.

[b]Fragments labelled with a '*' have been constructed de novo, by either stepwise synthesis, or condensation of partial fragments.

Note that some of the variants are non-covalent complexes of protein fragments.

Variants derived from the covalent coupling of fragments (1–65)H and (66–104) also contain a Met to Hse substitution at position 65.

[c]The properties of the intact, native protein include: E_m = 260mV; pK for alkaline transition ~9.3; and the presence of a 695nm absorption band.

Biological activity is typically assayed from the polarographic measurement of the ability of the variant protein to restore succinate oxidation by a preparation of cytochrome *c*-depleted mitochondria (from rat liver). Less frequently used as a substrate in these assays is ascorbate-TMPD.

[d]References: (1) Poerio *et al.*, 1986. (2) Juillerat and Tanuichi, 1986. (3) Wallace and Rose, 1983. (4) Proudfoot *et al.*, 1986. (5) Harris and Offord, 1977. (6) Proudfoot and Wallace, 1987. (7) Proudfoot *et al.*, 1989. (8) Corradin and Harbury, 1974. (9) Boswell *et al.*, 1981. (10) Wallace and Corthesy, 1986. (11) Koul *et al.*, 1979. (12) ten Kortenaar *et al.*, 1985. (13) Wallace *et al.*, 1989. (14) Boon *et al.*, 1979. (15) Raphael and Gray, 1989. (16) Boots and Tesser, 1987. (17) Corthesy and Wallace, 1986. (18) Wallace, 1987. (19) Wallace *et al.*, 1986.

1977; Wallace and Proudfoot, 1987). This property is also exploited in the final covalent coupling of semi-synthetic fragments through conformationally-directed religation: within certain non-covalent complexes, the close proximity of the α-carboxyl and α-amino groups of contiguous peptide fragments promotes the spontaneous formation of a peptide bond. Generally, for the peptide bond to form the carboxyl group must be activated, and this has most commonly involved the homoserine lactone present at the C-terminus of the cyanogen bromide cleavage peptide (1-65)H. Coupling to this fragment of the (66-104) fragment generates the intact cytochrome *c* molecule bearing a methionine to homoserine substitution at the position 65 religation site. [This substitution does not significantly affect the structural or functional properties of the product cytochrome *c* (Boswell *et al.*, 1981; Corradin and Harbury, 1974).] More recently, proteases used under conditions which favor bond synthesis have been used to attach activating groups [in the form of an esterifed (dichlorophenyl) amino acid, another C-terminal activated peptide fragment, or an activating group itself (Boc-hydrazide, methanol)] to the C-terminus of a peptide (Wallace *et al.*, 1989). In one report, the reverse proteolytic action of the enzyme clostripain, in the presence of 90% glycerol, was used to catalyze the direct recoupling of fragments (1-38)H and (38-104) (Juillerat and Homandberg, 1981). Because they require that reactive amino acid side groups be protected, standard procedures for activating and coupling peptide fragments are employed only during intermediary manipulations of isolated fragments [e.g. in the coupling of the fragments (66-79) and (80-104)].

Through the use of semi-synthetic methods, more than 50 single-site amino acid modifications have been made at 23 different positions along the polypeptide chain of cytochrome *c*. The results of these studies are summarized in Table I.2. It should be noted that for investigating the roles of specific amino acid residues, non-covalent complexes are less informative, as the occurrence of internal polypeptide chain termini introduces a pair of charged groups (the α-amino and α-carboxylate groups) and in addition may substantially alter the local conformation of the protein molecule. Although such complexes may be largely functional, they are usually described as having a looser conformation than the native molecule (Wallace and Proudfoot, 1987).

5. Site-directed mutagenesis

Difficulties with either the interpretation of the experimental results obtained, or the specific targeting of the site of modification have hampered the preceding methods for generating variant proteins in providing a thorough characterization of structure-function relationships in cytochrome *c*. With chemical modification, problems in attaining product homogeneity complicate the assignment of the cause of altered functional properties to a specific residue: side reactions can modify other protein groups, and the protein species derivatized at the desired site must be separated from species derivatized at other sites. In addition, the choice of target residue is restricted, as not all residues are susceptible to chemical modificiation (e.g. aliphatics and phenylalanine). Also, because derivitization of a side chain usually causes an increase in steric bulk, the method is not useful for modification of internal residues as it would result in gross conformational rearrangements. Naturally occurring variants from different species of organisms also cannot provide a definitive assignment of a role to a particular residue, because they almost always contain amino acid substitutions at multiple sites (Dickerson and Timkovich, 1975). In addition, by definition they do not supply variants at conservative positions in the amino acid sequence. Variants obtainable from *in vivo* mutagenesis are usually confined to those arising from single-nucleotide substitutions, and may also be restricted by the method of selection employed.

Direct chemical synthesis and site-directed mutagenesis are two methods which in principle permit a variant protein to be designed specifically. The direct synthesis approach may be handicapped if the protein contains certain post-translational modifications (such as covalent heme attachment), although this problem may be overcome to some extent using semi-synthesis. The utility of semi-synthesis is, however, limited by the requirement for a polypeptide chain cleavage site near the residue of interest. In the cytochrome *c* system, it is clear that site-directed mutagenesis emerges as the most general and effective method. The modifications which can potentially be generated are diverse, including amino acid substitutions, deletions, and insertions. The only shortcoming of the method is that changes are restricted to the 20 naturally occurring amino acids. Because of its power and convenience, site-directed mutagenesis is currently being

widely used to generate an ever-expanding collection of variant cytochromes *c*. The variants which have been fairly extensively characterized and thus provided useful information on structure-function relationships in cytochrome *c* are reviewed here. A summary of studies on the Phe82 and Arg38 variants of yeast iso-1-cytochrome *c*, which are major subjects of the work done in this thesis, is deferred to Sections I.E and I.F.

Asn52 and Tyr67, which together with Thr78 participate in hydrogen bonding a buried water molecule positioned below Met80 in the heme crevice, have both been replaced with non-hydrogen bonding side chains. The Phe67 variant of rat cytochrome *c* (Luntz *et al.*, 1989) has a slightly perturbed absorption spectrum, a 35 mV decrease in reduction potential (to 224 mV), and normal reactivity in transferring electrons from ascorbate-TMPD to cytochrome *c* oxidase. Most notable, however, is a significant increase in the stability of the structure of this variant, as indicated by the greater resistance to denaturation by both alkaline and acid pH and by heat or urea. The Ile52 substitution in yeast iso-1-cytochrome *c* (Das *et al.*, 1989) was first identified as a global second-site suppressor of mutations giving rise to unstable proteins. The Ile52 variant is more resistant to thermal denaturation than the Asn52 wild-type protein ($T_m = 64^\circ$ C, vs. 47° C). The increased stability of both the Phe67 and Ile52 variants is suggested to be due to several factors: the increased hydrophobicity introduced at internal sites of the protein molecule, the increased helix-forming propensity of the substituted amino acids (both Tyr67 and Asn52 occur within α-helical segments of polypeptide chain), and the inferred loss of the buried water molecule. These results indicate that neither the interaction between Tyr67 OH and the Met80 ligand sulfur, nor the internal hydrogen bonding network have major roles in stabilizing the protein structure. In addition, this network, whose water molecule is positionally sensitive to the redox state of the heme iron, appears not to be critical to the electron transfer function of the cytochrome *c* molecule. The roles of Tyr67 and Asn52 are inferred to be in stabilizing the buried water molecule, which may in turn have a role in conferring conformational flexibility to the adjacent polypeptide chain, a region of the cytochrome *c* molecule known to form interactions with redox partners. It should be noted that these studies predict that amino acid substitutions at the

third residue in the network, Thr78, would also give rise to a more stable protein (Luntz *et al.*, 1989), but semi-synthesis has shown this not to be the case (see Table I.2).

The trimethyllysine at residue 72 in yeast iso-1-cytochrome *c* has been replaced with an arginine (Holzschu *et al.*, 1987). Lys72 is strongly conserved in eukaryotic cytochromes *c*, and has been implicated by chemical modification studies (see Section I.D.1) to be involved in the binding of cytochrome *c* to redox partners. In yeasts and plants, Lys72 is specifically methylated by an S-adenosylmethionine protein-lysine N-methyl transferase (DiMaria *et al.*, 1979). The Arg72 variant iso-1-cytochrome *c* supports normal growth of yeast cells on lactate medium, but has a slightly lowered stability, and moderately altered reaction kinetics with cytochrome b_2. These results suggest that Tml72 may have a role in conferring structural stability to the iso-1-cytochrome *c* protein, or in recognition of redox partners. Replacement of Tml72 with an Asp has also been shown to affect both binding and electron transfer between iso-1-cytochrome *c* and cytochrome *c* peroxidase (Hazzard *et al.*, 1988a).

The three invariant proline residues have been the target of mutagenesis, with a focus on their role in stabilizing the native conformation of the cytochrome *c* molecule. The replacement of Pro30 with a threonine substantially decreases protein stability: the Thr30 variant of yeast iso-2-cytochrome *c* supports only poor growth of the yeast cells on non-fermentable media, and occurs at lowered levels *in vivo* (Wood *et al.*, 1988a). At Pro76, substitution by Gly lowered the pK_a for alkaline transition from ~8.4 to ~6.6, increased the susceptibility to denaturation by guanidine hydrochloride, and increased the solvent accessibility of tyrosine side chains (Wood *et al.*, 1988a,b; Nall *et al.*, 1989). For this Gly76 variant, the fast phase of unfolding induced by guanidine hydrochloride appears to occur at ~3 times the normal rate (Wood *et al.*, 1988b). Pro76 in cytochrome *c* is thus indicated to function in stabilizing the native versus the unfolded conformation of the polypeptide chain. In addition, the abnormal refolding of the denatured Gly76 protein, as monitored spectroscopically, implicates Pro76 in having a role, through cis-trans isomerization of its imide peptide bond, in the slow phase of polypeptide chain folding of cytochrome *c*. Amino acid replacements at Pro71 in yeast iso-1-cytochrome *c* obtained *in vivo*

(see above) have been supplemented by a Thr71 variant of iso-2-cytochrome *c* generated using site-directed mutagenesis (White *et al.*, 1987). The two studies agree in the importance of Pro71 in stabilizing the structure of cytochrome *c*. The Thr71 variant of iso-2-cytochrome *c* has a lowered pK_a for alkaline isomerization (to 6.6 from 8.5) and a decreased resistance to denaturation by guanidine hydrochloride.

Sorrell *et al.* (1989) used mixed-oligonucleotides to mutate the heme ligand His18 of yeast iso-2-cytochrome *c*. Of several amino acid replacements generated (including Tyr, Thr, Cys, Gln, Ser and Leu) only Arg resulted in a variant iso-2-cytochrome *c* which supported yeast growth (at a subnormal rate) on glycerol. The Arg18 variant protein possessed the 695nm absorbance band, was not auto-oxidizable, and did not bind carbon monoxide, thus indicating the heme iron had a complete coordination sphere. The variant protein had anomalous electrochemical behavior, although it appeared to have a normal reduction potential (~270mV).

Fetrow *et al.* (1989) have investigated the role of the four Ω-loops of yeast iso-1-cytochrome *c* (see Section I.A.2.a) as independent functional units. Yeast strains bearing mutations in which residues 22 to 28, or 74 to 78 of iso-1-cytochrome *c* are deleted are devoid of cytochrome *c* protein and do not grow on glycerol or lactate media. However, replacement of the loop consisting of residues 19 to 28 with the corresponding segments from other cytochromes *c* (tuna, *R. rubrum* c_2, *P. denitrificans* c_{550} or *P. aeruginosa* c_{551}) yielded partially to completely functional proteins. These results suggest that the presence of this loop is required for cytochrome *c* biosynthesis and activity, but that substantial sequence variability within the loop can be accommodated. Iso-1-cytochrome *c* with a deletion of residues 37 to 40 or 43 to 50 was present at about 50% the normal intracellular level and retained 10-20% activity, thus indicating that the residues within these two loops are not absolutely required for cytochrome *c* heme attachment, mitochondrial import, or enzymic interactions. It should be noted that most of the deletions and replacements generated are fairly severe, and the resultant proteins have not been extensively characterized, thus rendering the conclusions less than definitive.

E. Yeast iso-1-cytochrome *c*

In baker's yeast (*Saccharomyces cerevisiae*) two isozymes of cytochrome *c* occur, iso-1- and iso-2-cytochromes *c*. Iso-1-cytochrome *c*, the subject of the present studies, is the more abundant form, typically representing 95% of the total.

The amino acid sequence of yeast iso-1-cytochrome *c* was determined chemically by Narita *et al.* (1969) and Lederer *et al.* (1972). Although a component of the mitochondrion, yeast iso-1-cytochrome *c* is encoded by a nuclear gene. Cloning and sequencing of this gene (Smith *et al.*, 1979) confirmed the amino acid sequence determined from the protein. The biosynthesis of iso-1-cytochrome *c* involves the following steps: translation of apo-iso-1-cytochrome *c* on cytosolic ribosomes; removal of the amino-terminal methionine residue of the polypeptide chain through the action of an aminopeptidase (Tsunasawa *et al.*, 1985); trimethylation of the ϵ-nitrogen of Lys72 (DiMaria *et al.*, 1979); and covalent attachment of protoheme IX to two cysteinyl residues of apo-iso-1-cytochrome *c* through the action of a specific heme lyase (Dumont *et al.*, 1988). This final step has been shown to be concomitant with translocation of the holo-iso-1-cytochrome *c* molecule across the outer mitochondrial membrane. The fully processed protein contains 108 amino acid residues, and has a calculated molecular weight of 12720. In comparison to vertebrate cytochromes *c*, which are typically 103 or 104 amino acid residues in length, yeast iso-1-cytochrome *c* has a 5 residue-long extension at its amino terminus.

In addition to acting as an electron carrier in the electron transfer chain, yeast cytochrome *c* reacts with two other redox partners occurring within the intermembrane space of mitochondria, cytochrome b_2 (yeast lactate dehydrogenase) and cytochrome *c* peroxidase. Cytochrome b_2 contains both a heme and a flavin prosthetic group, and mediates electron transfer from L-lactate into the electron transport chain. Cytochrome *c* peroxidase catalyzes the reduction of hydrogen peroxide to water and acquires reducing equivalents specifically from ferrocytochrome *c*.

Yeast iso-1-cytochrome *c* clearly represents an ideal system for investigation of structure-function relationships in cytochrome *c*. Its amino acid sequence can be manipulated using both

site-directed and classical (*in vivo*) mutagenesis. Expression in yeast cells allows for a useful assay of the *in vivo* functionality of variant cytochromes *c*, and the protein can be readily prepared and purified in sufficient quantity for *in vitro* characterizations. The study of many aspects of the yeast cytochromes *c*, such as their high resolution NMR spectroscopy (Pielak *et al.*, 1988a) and their protein folding properties (Nall and Landers, 1983), is well developed. In addition, two well-studied soluble proteins, cytochrome *c* peroxidase and cytochrome b_2, provide convenient systems for the analysis of the interaction of cytochrome *c* with its physiological redox partners.

F. Site-directed mutagenesis of Phe82 in yeast iso-1-cytochrome *c*

Phenylalanine 82 was chosen as a target for mutagenesis for several reasons (Pielak *et al.*, 1985): its phylogenetic invariance suggested an important role; its side chain, positioned at the surface of the molecule and parallel to the heme group, has an interesting structural environment; and the phenyl group is not amenable to chemical modification. Various functions have been proposed for Phe82. Phe82 may have a structural role in establishing the appropriate heme environment. In addition, modelling studies indicate that the phenyl ring of Phe82 is positioned such that it could form part of the route for electron transfer between the heme groups of cytochrome *c* and cytochrome *c* peroxidase (see Section I.B.5). Molecular dynamics simulations of the modelled electron transfer complex between cytochrome *c* and cytochrome b_5 also suggest that the side chain of Phe82 may transiently adopt a conformation such that it can bridge the heme groups of the two molecules, thereby facilitating electron transfer between these groups (Wendoloski *et al.*, 1987).

Phe82 of yeast iso-1-cytochrome *c* has been replaced by tyrosine, serine, glycine, isoleucine, leucine, and alanine. The experimentally determined properties of these position 82 variants are summarized in Table I.3. In comparison to the wild-type protein, all of the variants have a lowered reduction potential, with the greatest decrease observed for the Gly82 and Ser82 variants (~50mV). These results suggest that the heme environments of the variant iso-1-cytochromes *c* are perturbed. The variants display a range of values for reactivity with $Fe(EDTA)^{2-}$, which provides a measure of the exposure of the cytochrome *c* heme group to the

Table I.3. Properties of position 82 variants of iso-1-cytochrome *c*

Variant	E_m (mV)[a]	pK_a for alkaline transition[b]	Reactivity with $Fe(EDTA)^{2-}$[a] (10^4 $M^{-1}s^{-1}$)	Relative steady state reactivity with CCP[c]	Reactivity with Zn-CCP[d]	
					k_t (s^{-1})	k_b (s^{-1})
Phe	290	8.5	7.2	100%	166	1.9×10^4
Ser	255	7.7	14.8	70%	151	2.3
Gly	247	7.7	13.7	20%	13	1.4
Tyr	280	8.4	6.2	30%	173	1.5×10^4
Ile	273	7.2	9.4	n.d.[e]	56	3.0
Leu	286	7.2	8.8	n.d.	93	2.0
Ala	260	n.d.	9.9	n.d.	n.d.	n.d.

References: [a]Rafferty *et al.*, 1990. [b]Pearce *et al.*, 1989. [c]Pielak *et al.*, 1985. [d]Liang *et al.*, 1987, 1988.
[e] n.d.: not determined

external medium. The apparent pK_a for the alkaline isomerization of all of the variants is lower than that of the wild-type protein (Pearce *et al.*, 1989), indicating that the heme crevice in the variant proteins has a lower stability. This apparent pK_a is determined by two factors, (i) the dissociation constant, K_H, of the ionizable group which is deprotonated during the alkaline isomerization process, and (ii) the equilibrium constant, K_C, for conversion of the native to the alkaline conformation (Davis *et al.*, 1974). Study of the kinetics of the alkaline isomerization reaction indicates that for the Ile82 and Leu82 variants, the drop in the apparent pK_a is due entirely to a drop in the pK_H of the titrating group (from ~11 to ~9.5), whereas for the Gly82 and Ser82 variants, both pK_H and K_C are lowered (from ~11 to ~9.1, and from ~300 to ~20). Little change is observed in the intrinsic (UV) CD spectra of the Gly82, Ser82 and Tyr82 variants, indicating that the overall secondary structure occurring in these proteins is not affected (Pielak *et al.*, 1986). However, all of the variants have altered CD spectra in the Soret region, with the most notable difference being the decreased intensity of the negative Cotton effect at 417 nm in the spectra of the oxidized forms of the variant proteins (Pielak *et al.*, 1986; Rafferty *et al.*, 1990).

These spectral differences are attributed either to the loss of the direct interaction between the π-orbitals of the Phe82 phenyl ring and those of the heme group, or to the altered conformation of the polypeptide chain in the vicinity of the heme group. In addition, the Soret CD spectra of the variants and of the wild-type protein are very similar at temperatures above 40° C, which suggests that amino acid substitutions at position 82 may decrease the stability of the native conformation of the polypeptide chain surrounding the heme. Collectively, the results from studies of the reduction potential, alkaline isomerization, $Fe(EDTA)^{2-}$ reactivity and circular dichroism of the variant proteins suggest that they all have a lowered stability and a moderate perturbation in heme environment and possibly in the conformation of the adjacent polypeptide chain.

The electron transfer activity of the position 82 variant iso-1-cytochromes *c* has been assessed by a number of means. All of the variant proteins, as the sole source of cellular cytochrome *c*, support viable growth of yeast cells under non-fermenting conditions, thus demonstrating that they are capable of carrying out electron transfer with both cytochrome *c* reductase and oxidase of the mitochondrial electron transfer chain. In addition, all of the variants show at least partial activity in *in vitro* steady-state assays with cytochrome *c* peroxidase. Using semi-synthesis, ten Kortenaar *et al.* (1985) (see Table I.2) have also shown that a Leu82 variant of horse cytochrome *c* has nearly normal activity in mediating electron transfer from ascorbate-TMPD to cytochrome *c* oxidase. These results indicate that Phe82 is not essential to the mechanism of electron transfer of cytochrome *c*, but do not rule out a role of Phe82 in optimizing the efficiency of the electron transfer reaction. In addition, it is conceivable that in the absence of Phe82, electron transfer occurs via an alternative mechanism.

The identity of residue 82 does affect the steady state reactivity of the cytochrome *c* variants with H_2O_2-oxidized cytochrome *c* peroxidase (Table I.3). However, the results of these steady state assays are not straightforwardly interpretable, as they are influenced by both the rate of association/dissociation of cytochrome *c* and cytochrome *c* peroxidase, and the rate of the electron transfer reaction within the intermolecular complex.

Liang *et al.* (1987, 1988) have examined the intermolecular electron transfer rates within a preformed 1:1 complex between iso-1-cytochrome *c* and cytochrome *c* peroxidase in which the heme iron has been replaced by a zinc ion. The reaction scheme is described below, with ZnP representing zinc-substituted cytochrome *c* peroxidase, and Fe(II,III) cyt *c* representing the two oxidation states of cytochrome *c*. Flash photolysis of the initial complex **A** generates the zinc-protoporphyrin triplet state (^{3}ZnP), which can either decay back to the ground state or reduce ferricytochrome *c* with rate constant k_t.

$$\underset{\mathbf{A}}{[\text{ZnP, Fe(III)cyt } c]} \xrightarrow{h\nu} \underset{\mathbf{A^*}}{[^3\text{ZnP, Fe(III)cyt } c]} \xrightarrow{k_t} \underset{\mathbf{B}}{[\text{ZnP}^{+\cdot}\text{, Fe(II)cyt } c]}$$

The resulting intermediate **B** returns to the ground state by thermal electron transfer, with rate constant k_b, from Fe(II)cyt *c* to the cation radical $ZnP^{+\cdot}$.

$$\underset{\mathbf{B}}{[\text{ZnP}^{+\cdot}\text{, Fe(II)cyt } c]} \xrightarrow{k_b} \underset{\mathbf{A}}{[\text{ZnP, Fe(III)cyt } c]}$$

The thermal reaction is analogous to the physiological oxidation of ferrocytochrome *c* by H_2O_2-oxidized cytochrome *c* peroxidase. The rate constant k_b can alternatively be measured within a complex in which cytochrome *c* is initially in the reduced state. In this case, the triplet-state zinc protoporphyrin is oxidized by an exogenous oxidant ($[Co(NH_3)_5Cl]^{2+}$) to the cation free radical, which is subsequently reduced by cytochrome *c*.

$$\underset{\mathbf{A}}{[\text{ZnP, Fe(II)cyt } c]} \xrightarrow{h\nu} [^3\text{ZnP, Fe(II)cyt } c] \xrightarrow{[Co(NH_3)_5Cl]^{2+}} \underset{\mathbf{B}}{[\text{ZnP}^{+\cdot}\text{, Fe(II)cyt } c]} \xrightarrow{k_b} [\text{ZnP, Fe(III)cyt } c]$$

$$[Co(NH_3)_5Cl]^{+} \longrightarrow [Co(H_2O)_6]^{2+}$$

This scheme precludes the possibility that structural rearrangements associated with the reduction of cytochrome *c* in the conversion of **A*** to **B** are determining the rate of the subsequent decay of the intermediate **B**.

The rate constants k_t for the photoinitiated electron transfer are similar for the wild-type protein and the Ser82 and Tyr82 variants, moderately lowered for the Leu82 and Ile82 variants, and substantially lowered for the Gly82 variant (see Table I.3). This rate is affected by the distance between the redox sites. The data thus indicate that the complex formed between

zinc–cytochrome *c* peroxidase and cytochrome *c* is similar for all position 82 variants, except the Gly82 variant, which is inferred to form a perturbed complex. The rate of the thermal electron transfer reaction is strongly dependent on the identity of the residue at position 82: the variants with non–aromatic side chains (Ser, Leu and Ile) react ~10^4 times slower than the wild–type and Tyr82 variants.

It is notable that in reactions where ferrocytochrome *c* is oxidized (i.e. those with cytochrome *c* peroxidase–H_2O_2 and with Zn^{++}–cytochrome *c* peroxidase) a decrease in reduction potential of the cytochrome *c* heme iron would be expected, in the absence of other effects, to increase the reaction rate. The experimental data (Table I.3) show that the various position 82 variants have a wide range of values for these reaction rates, and that the rates do not correlate with cytochrome *c* reduction potentials, as is most evident for the Gly82 and Ser82 variants. The results therefore emphasize that the nature of the side chain present at position 82 does directly affect the electron transfer activity of cytochrome *c*.

An aromatic side chain at position 82 is inferred to facilitate electron transfer by forming hole superexchange interactions between cytochrome *c* and cytochrome *c* peroxidase (Liang *et al.*, 1987, 1988). Modelling of the intermolecular complex suggests the pathway would include the heme and Phe82 phenyl ring of cytochrome *c*, and the imidazole ring of His181 and heme of cytochrome *c* peroxidase (Section I.B.5). CD spectroscopic studies indicating that the π–electron systems of Phe82 and of the heme in iso–1–cytochrome *c* are coupled (Pielak *et al.*, 1986) lend some support to this proposal. However, more recent studies on the kinetics of electron transfer within the Zn–cytochrome *c* peroxidase – iso–1–cytochrome *c* complexes (Everest *et al.*, 1990) downplay the importance of hole–superexchange pathways. These studies instead suggest that the rate of interconversion between conformational substates of cytochrome *c* having differing reactivity with Zn–cytochrome *c* peroxidase has a role in determining the rate of interprotein electron transfer.

G. Site-directed mutagenesis of Arg38 in yeast iso-1-cytochrome *c*

Arg38 is invariant in eukaryotic cytochromes *c* and also occurs in several bacterial cytochromes c_2. The guanidinium group of this residue interacts with the rear buried propionic acid group of the heme. Two structural roles have been proposed for Arg38. Arg38 has been suggested to function in stabilizing the protein by providing a counter-charge for the internal carboxylate group (Takano and Dickerson, 1981a). Wallace also suggests Arg38 stabilizes the fold of the protein, although he emphasizes that the network of hydrogen bonds formed by the guanidinium group serves to bind the bottom loop of the molecule to the main body (see Table I.2). Moore has proposed two roles related to the electrostatic interactions between Arg38 and heme group. First, Arg38 may lower the pK_a of the adjacent propionic acid group to a value below ~4.5, and thereby contribute to the pH-independence of the heme reduction potential within the physiological pH range. This proposal is based on the observed variation with pH of the reduction potential of cytochromes c_2 which possess amino acid replacements at Arg38 (Moore *et al.*, 1984). Second, the side chain of Arg38, through direct electrostatic interaction with the heme iron, may act as a trigger for remote conformational changes upon oxidation/reduction of the cytochrome *c* molecule (Moore, 1983).

Cutler *et al.* (1989) have replaced Arg38 in yeast iso-1-cytochrome *c* with lysine, histidine, glutamine, asparagine, leucine and alanine. These amino acid substitutions have very little effect on either the electron transfer or structural properties of the iso-1-cytochrome *c* protein: all of the variants support the growth of yeast on lactate, exhibit normal *in vitro* reactivity with cytochrome *c* peroxidase, and have NMR and absorption spectra very similar to those of the wild-type protein. The focus of the work by Cutler *et al.* was on the proposed role of Arg38 in controlling the reduction potential of iso-1-cytochrome *c*. The variation of reduction potential with pH for all of the variants was found to be identical to that for the wild-type protein. This result suggests that Arg38 is not responsible for maintaining a low pK_a for the heme propionate group. However, the identity of the amino acid at position 38 does have a marked effect on the reduction potentials of the variant proteins, which were observed to be as follows: Arg38

(wild-type): 272mV; Lys38: 249mV; His38: 245mV; Gln38: 242mV; Asn38: 238mV; Leu38: 231mV; and Ala38: 225mV. It is notable that the observed shifts in reduction potential for the Lys38 and Gln38 variants follow the same pattern observed for the same substitutions generated semi-synthetically in non-covalent complexes of horse cytochrome *c* (Proudfoot and Wallace, 1987; see Table I.2).

H. Site-directed mutagenesis of Cys102 in yeast iso-1-cytochrome *c*

Amongst eukaryotic cytochromes *c*, yeast iso-1- and bullfrog cytochromes *c* are unique in having a cysteine at position 102. In the latter protein, Cys102 forms an intramolecular disulfide bond with Cys20 (Brems *et al.*, 1982). In yeast iso-1-cytochrome *c*, Cys102 is of particular interest because its free sulfhydryl group is highly reactive toward both chemical modification (Zuniga and Nall, 1983; Motonaga *et al.*, 1965) and oxidation (Cutler *et al.*, 1987), and in addition is involved in the covalent dimerization of iso-1-cytochrome *c* via intermolecular disulfide bond formation (Bryant *et al.*, 1985).

The reactivity of Cys102 complicates electrochemical analyses of iso-1-cytochrome *c*, as the protein is readily autoreduced and exhibits non-Nernstian behavior. Dimerization in the absence of exogenous thiol reducing agents also hinders the study of the oxidized form of the protein. Two approaches to alleviate these problems have been used. The free sulfhydryl group has been blocked by modification with iodoacetamide (Nall and Landers, 1983) or methyl thiosulfonate (Ramdas *et al.*, 1986). Alternatively, site-directed mutagenesis has been used to replace Cys102 in yeast iso-1-cytochrome *c* with either a threonine (Cutler *et al.*, 1987) or a serine (Mayo *et al.*, 1986) residue. In the work of Cutler *et al.* threonine was chosen as a replacement because it is the most frequently occurring residue at this position in eukaryotic cytochromes *c*. The Thr102 variant iso-1-cytochrome *c* has been shown by Cutler *et al.* to be functionally indistinguishable from the genuine wild-type protein. It has normal behavior *in vivo*, as it is transported into the mitochondria, is methylated, and supports growth of yeast cells on lactate. In addition, it has a normal absorption spectrum, including a 695 nm band, well-behaved electrochemistry, and a typical reduction potential. Iso-1-cytochrome *c* bearing Thr102 thus serves as the "wild-type" background

for a large part of the structural and functional studies described in this thesis.

I. Objectives of the present research

The underlying objective of the studies described in this thesis is to further the understanding of the structure and function of the cytochrome *c* molecule. The first goal of the current project was to obtain an accurate description of the three-dimensional structure of the wild-type yeast iso-1-cytochrome *c* protein. The structure of yeast iso-1-cytochrome *c* was compared to those of tuna and rice cytochromes *c*. As these cytochromes *c* possess about 60% amino acid sequence divergence, these comparisons provide information on the effect of differences in amino acid sequence on various structural aspects of the cytochrome *c* molecule. Equally importantly, the wild-type iso-1-cytochrome *c* structure will provide a basis for both the interpretation of the properties of singly-mutated iso-1 variant proteins, and the design of new variants. A major focus of this work was to determine the structural and functional role of the evolutionarily conserved residue phenylalanine 82. This involved determining the effects of amino acid substitutions at position 82 on the structure of iso-1-cytochrome *c*, and interpreting differences in functional properties of the variant proteins in terms of the altered structural properties. The role of the invariant residue arginine 38 was similarly, although less extensively studied.

Appendix: Basic primer on structure determination using X-ray crystallography

This appendix describes the basic theory and some practical details of the use of the X-ray diffraction method in determining the three-dimensional structure of a protein molecule. Its intent is to provide the reader who is unfamiliar with the method sufficient background to comprehend the crystallographic aspects of the research described in this thesis. For a more thorough discussion, the reader is referred to Blundell and Johnson (1976) or Sherwood (1976).

1. Crystalline structure

A crystal is a solid in which the component molecules are arranged such that there is long-range, well-defined order in three dimensions. Within the crystal, relatively precise

intermolecular interactions are formed, with the molecules packing in a symmetrical manner. The crystal can be represented as a fundamental unit cell which repeats in a three-dimensional lattice. Each identical unit cell contains one or more asymmetric units which are related to one another by well-defined symmetry elements. Generally, an asymmetric unit is composed of one protein molecule. Thus the objective of the X-ray crystallographic experiment is to determine the atomic structure of the asymmetric unit.

2. Diffraction theory

The interaction of X-rays with irradiated matter is mainly electronic. This occurs through a process called Thomson scattering, in which the electric field of the incident X-ray causes an electron to oscillate, and thereby emit its own electromagnetic radiation in all directions. When X-rays interact with an object (i.e. a collection of atoms), each electron in the object acts as an independent scatterer. However, superposition of waves emitted by different electrons occurs. The intensity and phase of the total wave scattered by an object in a particular direction is entirely dependent on the distribution of electrons within the object. This concept is the foundation of structure determination by X-ray diffraction methods, and is expressed in the cardinal equation of X-ray scattering:

$$\vec{F}(\vec{S}) = \int_{\substack{\text{all}\\ \text{object}}} \rho(\vec{r}) \exp i(\vec{S}\cdot\vec{r})\, d\vec{r}, \qquad (1)$$

which states that $F(\vec{S})$, the total wave scattered in the direction $\vec{S}$, consists of the sum of individual waves $\exp i(\vec{S}\cdot\vec{r})$ scattered in the direction $\vec{S}$ by infinitesimal volumes, $d\vec{r}$, of electron density, $\rho(\vec{r})$ located at the position $\vec{r}$. The integration is over all the infinitesimal volumes in the object. Mathematically, the set of total waves, called the diffraction pattern, represents the Fourier transform of the distribution of electron density in the object.

3. Basic X-ray crystallographic theory

For a regularly arranged group of objects, such as a unit cell arranged in a crystal lattice, all of the individual unit cells scatter X-rays in an identical manner. But there is destructive interference of all scattered X-rays except for those travelling in a few certain directions with

respect to the translational symmetry of the crystal lattice. A convenient representation of this interference is that for a crystal's real lattice, there is a reciprocal lattice which defines the permissible directions of the diffracted X-rays. The reciprocal lattice contains a sampling of the diffraction pattern of the unit cell. Each position in this lattice represents a permitted diffracted wave, called a reflection, and is specified by a triplet of integers h,k,l.

For diffraction by a crystal, two simplifications can be made to the Fourier transform of the electron density (equation 1). The continuously variable direction of the diffracted X-ray, $\vec{S}$, can be replaced by the discrete direction specified by hkl. Second, because appreciable electron density is localized near atomic nuclei, appropriate substitution of the atomic electron density function ρ_j and the atomic position (x_j, y_j, z_j) can be made. The integral is then replaced by a summation over the n individual atoms contained in the unit cell:

$$\vec{F}(hkl) = \sum_{j=1}^{n} \rho_j \exp 2\pi i(hx_j + ky_j + lz_j). \qquad (2)$$

This is the structure factor equation, which again shows that individual reflections, referred to as structure factors, will have an amplitude (intensity) and phase determined by the atomic positioning within the unit cell. Equally importantly, this equation allows one to calculate exactly how a known distribution of atoms within the unit cell would diffract X-rays.

The atomic electron distribution function in equation 2, ρ, has two components. The first, the atomic scattering factor, is determined by the number and spatial arrangement of electrons around a positionally fixed atomic nucleus, and thus is dependent only on the identity of the atom (i.e. the atomic number). The second, the atomic temperature factor B, corrects for thermal motion of the atomic nucleus (and of its associated electrons). Atomic motion permits the electrons to occupy a larger volume of space, thus decreasing the effective electron density and the scattering power of the atom. Each atom has an associated B-factor, which provides a measure of the mean-square displacement of the atom from its average position.

A property of the Fourier transform is that it can be inverted. Inversion of the Fourier transform of the electron density (eq. 1) yields the electron density equation,

$$\rho(\vec{r}) = \int_{all\ \vec{s}} \vec{F}(\vec{S}) \exp -i(\vec{S}\cdot\vec{r})\, d\vec{S},$$

or equivalently,

$$\rho(x,y,z) = \sum_{h=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} |F(hkl)| \exp i\alpha(hkl) \exp -2\pi i(hx + ky + lz). \quad (3)$$

Thus, the distribution of electron density in the unit cell can be deduced from the Fourier transform of the diffraction pattern. In practice, equation 3 is used to calculate the amount of electron density present at selected points throughout the unit cell. The resultant electron density map allows for a reconstruction of the atomic structure within the unit cell. However, the diffracted waves contain two pieces of information: the amplitude, which can be readily measured experimentally; and the phase, which is unobservable. [The electron density equation has been written in a form which emphasizes that the structure factor has these two components, $|F(hkl)|$ and $\alpha(hkl)$.] Assigning phase information to the reflection data constitutes the major obstacle in the crystal structure determination.

4. Patterson synthesis and molecular replacement methods

A Patterson map of the unit cell can be calculated using only the experimentally observable intensities of the diffracted X-rays:

$$P(u,v,w) = \sum_{h=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} |F(hkl)|^2 \exp -2\pi i(hu + kv + lw). \quad (4)$$

In contrast to the electron density map, which has peaks occurring at the positions of atoms in the unit cell, the Patterson map has peaks at positions corresponding to interatomic vectors within the unit cell. Thus, a peak at the point (u,v,w) in the Patterson map indicates that there exist in the real unit cell two atoms related positionally such that $u=x_1-x_2$, $v=y_1-y_2$ and $w=z_1-z_2$. Because of its much greater complexity (it contains n^2 peaks instead of only n, where n is the number of atoms in the real unit cell), the Patterson map cannot be interpreted to yield directly the atomic positions in a protein molecule. However, the experimentally determined set of interatomic vectors which is obtained from this type of map is particularly useful in a number of search methods, collectively referred to as molecular replacement.

Molecular replacement (Rossman, 1973) is one method for determining the phases of the observed diffraction data, and requires that a preliminary model be proposed for the three-dimensional structure of the unknown protein. Generally, this preliminary model is based on the

known structure of a highly related protein, frequently the homologous protein from another species of organism. The modelled structure is not immediately useful in assigning phase information to the diffraction intensities measured for the unknown protein, which represent the Fourier transform of the contents of the unknown protein's crystallographic unit cell. Therefore, what is required is that the search molecule be orientationally and translationally positioned with respect to the symmetry-related molecules in this unit cell. The search for this positioning is carried out in two stages. In the first stage, the rotation function is used to correctly orient the model. In this procedure, a set of intramolecular interatomic vectors is calculated for the search model. This set of vectors is then systematically rotated, in steps about three mutually orthogonal axes, and a comparison is made with a Patterson map calculated with the observed diffraction data. The comparison is usually restricted to short vectors (within a molecule's radius from the Patterson origin), in order that only intramolecular interatomic vectors in the observed vector set be included. The superposition of the modelled and observed sets of interatomic vectors will be maximal when the orientation of the search model coincides with that of one of the molecules in the unit cell of the unknown protein. Once the correct orientation of the search model has been thus defined, the translation function is used in the second stage to position the oriented model with respect to the symmetry elements present in the crystal unit cell. In this procedure, a symmetry axis is chosen, and the model is systematically translated to all the grid points on a plane perpendicular to the chosen axis. For each prospective positioning of the search molecule, the symmetry related molecule is generated, and then a set of intermolecular interatomic vectors is calculated and compared against the Patterson map. The correct location for the search model with respect to the symmetry axis can hopefully be identified as a point which gives rise to a high agreement between the observed and calculated interatomic vectors sets. If the positioning of the search model with respect to a different symmetry axis can be determined, from another two-dimensional search about this second axis, then the positioning of the search molecule within the unit cell is defined. The entire contents of the unit cell can then be generated by applying the crystallographic symmetry elements to the approximately oriented and positioned search molecule.

Finally, phases can be calculated for each of the experimentally observed structure factor amplitudes, and an initial electron density map computed.

5. Refinement

That the true values of the phases of the diffraction intensities cannot be directly measured has several consequences. Most importantly, the phases as determined by calculation or by indirect experimental means serve only as estimates for the true phases, and will undoubtedly contain substantial error. Removal of this error, while at the same time maximizing the accuracy of the atomic model, is the objective of structural refinement, the process by which the agreement between the proposed structure and the observed diffraction data is optimized. Refinement is necessarily an iterative process, in which indicated corrections are made to the proposed model and new structure factors (amplitudes and phases) having (hopefully) greater accuracy are calculated and then used to initiate another cycle. Because accurate structural information can be obtained only after extensive refinement, protein crystallographers expend a great deal of effort in refinement of their structures.

Assessment of the agreement between the calculated and observed diffraction patterns is based on the structure factor amplitudes. The most commonly used index is the conventional R-factor, defined as

$$R = \sum_{hkl} ||F_o(hkl)| - F_c(hkl)|| \; / \; \sum_{hkl} |F_o(hkl)|.$$

A low R-factor implies that there is good agreement between the calculated phases and the true phases. Clearly, the R-factor will have a minimum value of 0 when the agreement between F_O and F_C is perfect. Typically, well-refined protein structures have values below ~0.20.

Iterative examination of Fourier maps constitutes one strategy in structural refinement. As an example, inspection of the initial map calculated with phases estimated using molecular replacement methods may indicate that corrections to the original model are required. The revised model can then be used to obtain improved estimates for the calculated phases, and another electron density map computed. This procedure is carried out repetitively, with the aim that the calculated phases will converge to the real phases. It should be noted that because the electron

density map is synthesized from observed intensities combined with phases calculated using the current proposed model, the features of these maps will be biased toward the original structure. In practice, manual adjustments to the model are interspersed with automated refinement procedures, which are used to speed up the refinement process and to increase the accuracy of the determined atomic positions. With these methods, the positional and temperature factor parameters of the atoms of the proposed structure are adjusted by least-squares analysis to minimize the overall discrepancy between calculated and observed structure factor amplitudes.

Another effective tool in the refinement process is the difference Fourier map, which identifies features of electron density accounted for in one set of structure factors, but not in another. For structural refinement, it is useful to compare the observed and calculated structure factors according to the following equation:

$$\Delta\rho(x,y,z) = \sum_{h=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} [F_o(hkl) - F_c(hkl)] \exp i\alpha_c(hkl) \exp -2\pi i(hx + ky + lz). \quad (5)$$

Positive peaks will occur in this map at the positions of previously unincluded atoms (e.g. solvent atoms). Wrongly positioned atoms in the model will appear within negative peaks, possibly adjacent to positive peaks indicating the correct positioning. Frequently, coefficients of the type $2F_o - F_c$ are used to calculate maps that combine features of both the direct Fourier (F_o) and difference Fourier ($F_o - F_c$) maps.

When two homologous proteins have the identical (isomorphous) crystal form (e.g. yeast iso-1-cytochrome *c* and its variants), the phases of the two independent sets of diffraction intensities will in general be very similar. In this case, a difference map of the type

$$\Delta\rho(x,y,z) = \sum_{h=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} [F(hkl)(native) - F(hkl)(variant)] \exp i\alpha_c(hkl,native) \cdot \exp -2\pi i(hx + ky + lz). \quad (6)$$

can be used to identify conformational differences between the two proteins.

Another important aspect of the structural refinement process is resolution. Resolution is the minimum distance between two groups which appear in the electron density map as having distinct or separate electron density peaks. The resolution of a structure is critically dependent on the completeness of the diffraction intensity data set used in the structural determination. Equation 3 shows the image of the electron density is synthesized from a Fourier series containing an

infinite number of terms. In practice, the series is truncated, partly because higher order terms are too weak to be measured (i.e. are beyond the diffraction limit of the crystal). A decrease in the quality of the electron density map accompanies a decrease in the number of reflections that contribute to the synthesis of the map. This problem is compounded by the fact that it is the high order terms which contain information on fine details. Furthermore, in least-squares structural refinement, the atomic parameters (x,y,z and B) are adjusted against the measured intensities. Inclusion of a greater number of reflections means that these parameters are determined by a greater number of observations. These considerations demonstrate that increasing the resolution allows for an increased accuracy in the determination of atomic positions. In general, the accuracy of a structural determination is very difficult to assess, but it must be emphasized that the level of uncertainity will be greater for structures determined at low resolution.

II. EXPERIMENTAL PROCEDURES USED IN STRUCTURAL ANALYSES

A. Source of protein for crystallization experiments

Wild-type yeast (*Saccharomyces cerevisiae*) iso-1-cytochrome *c* (Lot no. 33F-7260) was purchased from Sigma Chemical Co. (St. Louis). The protein as supplied is greater than 95% isozyme-1, as confirmed using denaturing SDS-polyacrylamide gel electrophoresis, and was used without further purification. For use, the dry protein was dissolved to a suitable concentration in buffer, and any insoluble material was removed by centrifugation.

Iso-1-cytochrome *c* proteins possessing amino acid substitutions were obtained from yeast strains (GM-3C-2) that lacked functional iso-1- and iso-2-cytochrome *c* genes, and contained plasmids (YEp13) bearing iso-1 genes encoding the mutated protein. Isolation of iso-1-cytochrome *c* from yeast cultures (typically 40 l) was performed essentially as described by Cutler *et al.* (1987). Briefly, cytochrome *c* was released from yeast cells by ethyl acetate extraction, and then purified by successive CM-cellulose and CM-Sepharose ion exchange, and Sephadex G-50 gel filtration chromatographies. The protein was concentrated and exchanged into the appropriate buffer by ultrafiltration, and stored in liquid nitrogen or in a -70°C freezer.

B. Growth of crystals

The initial screening that determined conditions suitable for the formation of wild-type iso-1-cytochrome *c* crystals has been described by Sherwood and Brayer (1985). A pH in the range 5.8 to 7.2, a final ammonium sulfate concentration of at least 88% saturation, and at least 20 mM DTT were required to effect crystal growth. In order that the parameters be established appropriate for the growth of large, single crystals of iso-1-cytochrome *c*, with minimal concomitant formation of microcrystalline and precipitated material, a large number of growth conditions were finely sampled, using the vapor diffusion technique in tissue culture plates (McPherson, 1982). The optimal conditions using this method were determined to be the equilibration of 5 μl of protein solution (25 mg/ml) containing 78% saturated (3.17 M) ammonium sulfate, 40 mM dithiothreitol, and 0.1 M sodium phosphate (adjusted to pH 6.2 using ammonium

hydroxide or acetic acid), against 1.0 ml of 92% saturated (3.74 M) ammonium sulfate in the same buffer.

Although large crystals of iso-1-cytochrome *c* could be obtained using this hanging drop procedure, trials set up under these conditions often did not yield crystals, due apparently to the absence of nucleation sites. For this reason, free interface diffusion (Salemme, 1972), in which the formation of crystal nuclei is favored by the protein initially experiencing a supersaturating concentration of precipitant, was attempted and found to be the most reliable method for producing the large crystals required for collection of high resolution data. Crystallization conditions were selected based on the optimized conditions determined using the vapor diffusion method. The final ammonium sulfate concentration (90%) was chosen to be slightly lower than the 92% used in the vapor diffusion experiments, in which a slightly higher precipitant concentration is required to promote crystal nucleation. This lower final concentration would be expected to favor slower crystal growth and less loss of protein through precipitate formation. A typical trial, carried out in a melting point capillary (1.5 mm inner diameter), involved the layering of 5 μl of a concentrated protein solution (~200 mg/ml) in a buffer containing 0.1 M sodium phosphate (pH 6.2) and 40 mM dithiothreitol, on top of 45 μl of 100% saturated (4.06 M) ammonium sulfate in the same buffer. A fair amount of turbidity formed initially at the interface between the protein and saturated ammonium sulfate solutions. The resultant precipitate would then sink into the column of ammonium sulfate. But as the two solutions mixed, and the concentration of precipitant in the ammonium sulfate phase decreased, much of the precipitate was observed to redissolve. Crystal growth initiated at nuclei that had formed within the initial precipitate, and generally occurred at the original interface or along the trail made by this precipitate as it fell into the column of ammonium sulfate. Crystals grew to a maximal size of 0.9 x 0.6 x 0.5 mm in about two weeks. At the end of this period, the solution had become completely clear (the initial protein solution was bright red), with all the protein incorporated into the crystals present. In general, very little precipitate remained.

The iso-1-cytochrome *c* proteins containing amino acid substitutions proved to be much less cooperative in forming large, single crystals. Both the hanging drop and free interface diffusion methods, employing similar conditions to those used to grow wild-type iso-1-cytochrome *c* crystals, could effect crystallization of the variant proteins. However, during the early stages of most crystallization trials, much of the protein fell out of solution in the form of brown, denatured protein precipitate or fine, needle-shaped aggregates, instead of becoming incorporated into crystal nuclei. This problem could be alleviated to some degree using the hair-seeding technique developed by Leung *et al.* (1989). In this method, a well-formed crystal of wild-type iso-1-cytochrome *c* is crushed in a buffered solution of ammonium sulfate. The tip of a human hair is immersed in this seeding solution, and then passed through a droplet of protein solution containing the variant iso-1-cytochrome *c*. Precipitate formation is reduced, as this process must compete with the incorporation of protein into the seeded crystal nuclei. The difficulty with this procedure is in introducing only a small number of initiation sites for crystal growth. As was frequently the case, the presence of numerous nucleation sites gave rise to a shower of small, but well-formed, crystals. Thus, in general, only small crystals of the variant iso-1-cytochrome *c* proteins could be obtained. This limited the resolution at which the mutant structures could be determined.

The exact growth conditions and the sizes of the crystals employed in the studies discussed herein are summarized in Table II.1. It is notable that in contrast to Cys102-containing iso-1-cytochrome *c* proteins, the Gly82 variant in which Cys102 was replaced by Thr102 was well-behaved in the absence of highly reducing conditions. The optimal growth of crystals of this variant was obtained in solutions that contained no dithiothreitol.

C. Characterization of the crystals

1. Oxidation state of crystalline iso-1-cytochrome *c*

Crystals grown in the presence of a high concentration of reducing agent (40 mM dithiothreitol) would be expected to contain cytochrome *c* protein in which the heme iron is in the reduced (Fe^{2+}) state. This was verified, as a crystal of wild-type iso-1-cytochrome *c* dissolved in

Table II.1. Growth conditions, sizes and unit cell dimensions of crystals of wild-type and variant iso-1-cytochromes *c*

Protein crystal	Crystallization method and conditions[a]	Crystal size (mm)	Unit cell dimensions (Å) a (=b)	c
Wild-type (Cys102)			36.46	136.86[b]
B	HD } Final A.S. 92%, 45mM DTT, pH 6.3	0.6x0.4x0.4		
C	HD }	0.8x0.5x0.5		
D	HD. Final A.S. 92%, 45mM DTT, 0.2 M NaCl, pH 6.0	0.7x0.5x0.4		
E	FID }	0.8x0.6x0.5		
F	FID } Final A.S. 90%, 40mM DTT, pH 6.2	0.8x0.6x0.5		
G	FID }	0.9x0.7x0.5		
Ser82(Cys102)	FID. Final A.S. 90%, 40mM DTT, pH 6.2	0.3x0.3x0.2	36.45	136.48
Gly82(Cys102)	FID. Final A.S. 90%, 45mM DTT, pH 6.2	0.3x0.3x0.1	36.45	136.57
Gly82(Thr102)			36.48	137.20[b]
B	HD,S } Final A.S. 92%, 0mM DTT, pH 6.0	0.5x0.3x0.2		
C	HD,S }	0.5x0.4x0.3		
Tyr82(Thr102)	HD,S. Final A.S. 90%, 20mM DTT, pH 6.0	0.5x0.4x0.1	36.43	136.67
Ile82(Thr102)	HD,S. Final A.S. 90%, 20mM DTT, 0.15 M NaCl, pH 5.4	0.5x0.4x0.2	36.48	136.79
Ala38(Thr102)	HD,S. Final A.S. 90%, 20mM DTT, pH 6.0	0.3x0.2x0.1	36.44	136.84

[a]HD: hanging drop; FID: free interface diffusion; S: seeded.
All crystallization trials used sodium phosphate buffer (0.1 M) and were conducted at room temperature. Final A.S. refers to the final ammonium sulfate concentration, in per cent saturation (100% ≡ 4.06 M).
[b]Averaged over all crystals used of this protein.

phosphate buffer yielded a visible absorption spectrum typical of that for reduced cytochrome *c*. Crystals of the Gly82(Thr102) variant of iso-1-cytochrome *c* were grown in the absence of dithiothreitol. Spectroscopic measurements showed that this protein was crystallized in its oxidized form.

2. Treatment of crystals prior to X-ray analysis

A curious observation was that after 1-2 months, wild-type iso-1-cytochrome *c* crystals left undisturbed in their original hanging drops became internally disordered and lost their ability

to diffract X-rays, despite exhibiting no altered morphological characteristics. Because it was felt that decomposition of the reducing agent may be a factor in the crystal disordering, prior to their analysis using X-ray diffraction, crystals were treated in the following manner. The crystal was removed from the vessel in which it was originally grown and immersed in ~100 μl of freshly prepared mother liquor contained in a depression plate. The mother liquor was identical in composition to that used to grow the crystal, except for an ammonium sulfate concentration greater by ~0.16 M (4% saturation). The slightly higher ammonium sulfate concentration was used to stabilize the crystal during subsequent manipulation during mounting.

In the mounting procedure, a crystal was drawn into a thin-walled glass capillary having a diameter (0.7 or 1.0 mm) just large enough to accommodate the crystal. In order that crystal slippage and X-ray scattering by excess fluid surrounding the crystal be minimized, the crystal was then dried as thoroughly as possible using a finely-drawn out glass capillary and strips of filter paper. Finally, a short column of mother liquor was introduced at each end of the mounting capillary, which was then sealed with wax.

3. Space group and external morphology of the crystals

X-ray precession photographs of wild-type iso-1-cytochrome *c* crystals are shown in Figure II.1. Based on the 4-fold symmetry of the hk0 layer of the diffraction pattern, and on the systematic absence of axial reflections with h,k $\neq$ 2n, and l $\neq$ 4n, the space group of the crystals was assigned to be $P4_32_12$, or its enantiomorph $P4_12_12$. Unit cell parameters were initially determined from precession photographs. They were later more accurately determined, from diffractometer settings of 25 strong reflections with $20.7^\circ < 2\theta < 32.3^\circ$, to be $a=b=36.46(5)$ Å, and c=136.86(12) Å (standard deviations in parentheses). The unit cell dimensions observed and the crystallographic symmetry present indicated that only one iso-1-cytochrome *c* molecule occurs in the asymmetric unit (Matthews, 1968). In all cases, whether or not seeding with wild-type iso-1-cytochrome *c* crystals was performed, crystals of the variant iso-1-cytochromes *c* were isomorphous with those of the wild-type protein (see Table II.1.).

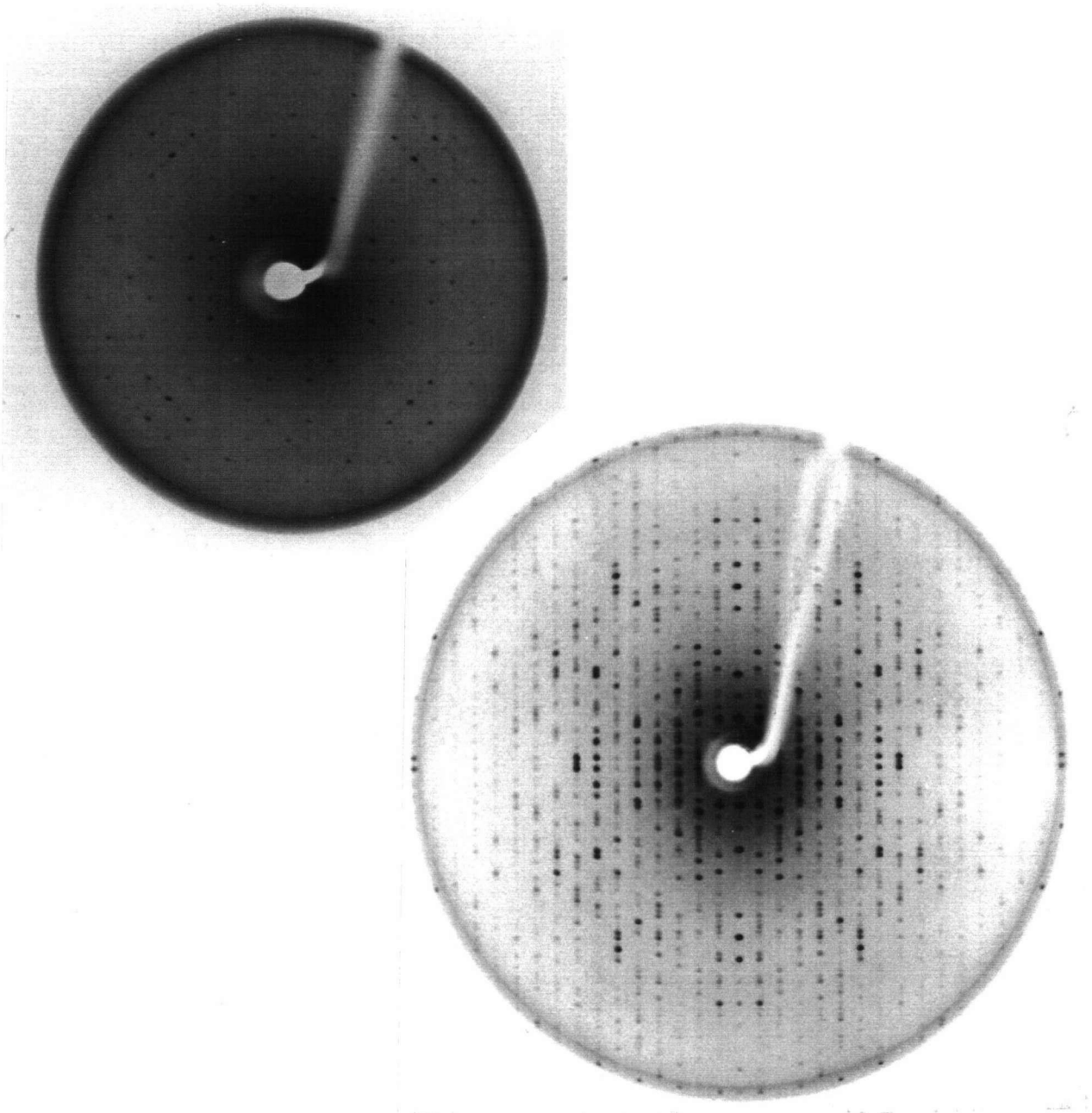

Figure II.1. Precession photographs from crystals of wild-type yeast iso-1-cytochrome *c*.
(*Top*) The hk0 zone of the reciprocal lattice. Evident is the four-fold symmetry about an axis perpendicular to the plane of the photograph and passing through the origin of the layer. The **a*** and **b*** axes run horizontally and vertically. Axial reflections having h (or k) odd are systematically absent. The outermost spots on this photograph correspond to a resolution of 2.98 Å. The photograph was taken with a precession angle (μ) of 15° and an exposure time of 20 hours.
(*Bottom*) The h0l zone of the reciprocal lattice. The **a*** axis runs horizontally, and the **c*** axis vertically. The systematic absence of reflections with $l \neq 4n$ is evident along the **c*** axis. The outermost spots on this photograph correspond to a resolution of 2.25 Å. The photograph was taken with $\mu = 20°$ and an exposure time of 12 hours.
For both photographs, the crystal-to-film distance was 75mm, and CuKα radiation ($\lambda = 1.5418$ Å) was used, with power to the X-ray tube of 26mA, 40kV.

Crystals of reduced iso-1-cytochrome *c* have a deep orange-red color and a tetragonal bipyramidal shape (Figure II.2). As determined from precession photography, the orientation of the reciprocal lattice with respect to the external morphology of the crystals is as follows: the *a* and *b* crystallographic unit cell axes make a 45° angle with the equatorial edges of the tetragonal bipyramid, while the *c* axis is perpendicular to the equatorial plane. Thus, the *c* axis runs parallel to the external 4-fold symmetry axis of the crystal.

D. Collection of X-ray diffraction data

All reflection data were collected on an Enraf-Nonius CAD4F11 diffractometer having a crystal-to-detector distance of 368 mm, and a helium-purged path for the diffracted beam. The incident radiation was nickel-filtered CuKα, produced from a fine focus (0.75 x 15.0 mm) X-ray

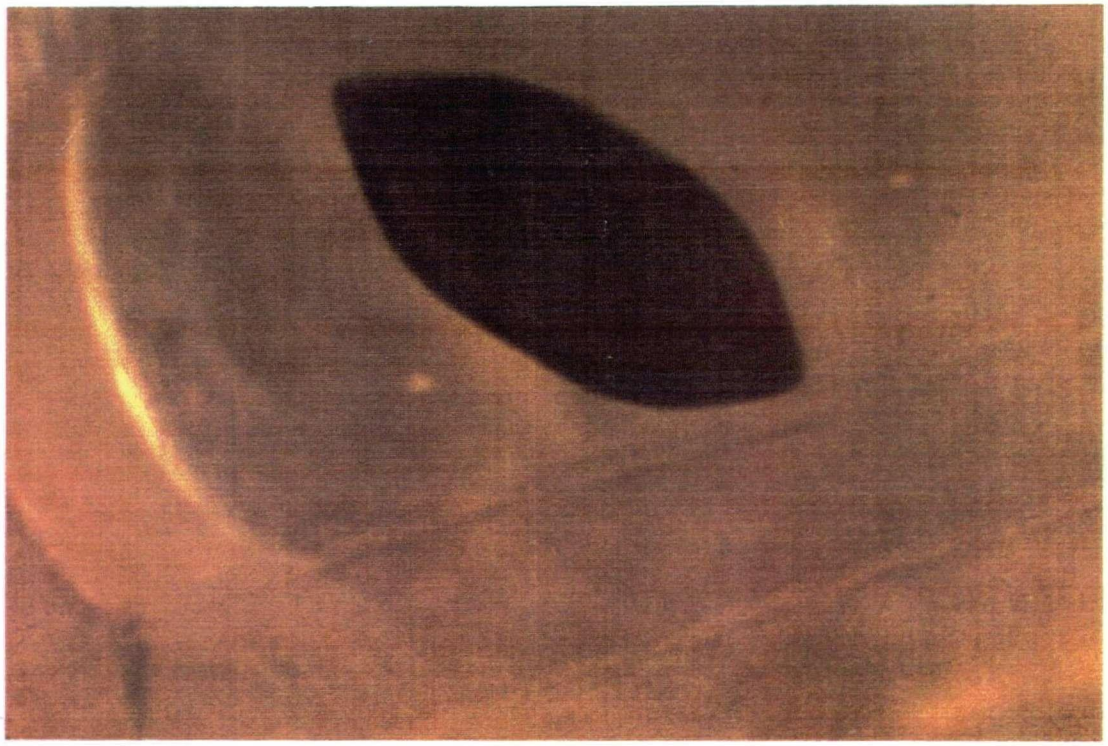

Figure II.2. Photomicrograph of a crystal of wild-type yeast iso-1-cytochrome *c*. The solution surrounding the deep red crystal is the mother liquor from the hanging drop within which the crystal was grown. The *a,b* axis of the crystal runs between the pointed ends of the crystal, whereas the *c* axis runs perpendicular to this axis and within the plane of the photograph. The maximum dimension of the crystal shown is ~0.9 mm.

tube operated at 26 mA, 40 kV. All data collections were carried out at an ambient temperature of 15°C.

The strategy for data collection, chosen to maximize the amount of data that could be collected from any one crystal, employed fairly narrow scan widths, coupled with relatively slow scan speeds. Typically, continuous 0.5–0.6° omega scans at a speed of 0.55°/min were used. Normally for a given crystal, data were collected in spherical shells, beginning with the highest resolution shell. Background counts for each reflection were measured using scans 1/6 of the total scan width on either side of the reflection peak. Crystal decay and slippage were monitored by measuring at approximately 2.5 hour intervals the intensities of three strong standard reflections. Yeast iso-1-cytochrome *c* crystals were fairly resistant to radiation damage; the typical decay experienced by a crystal during data collection (generally ~11000 reflections, 3 weeks) was about 35%, based on intensity control reflections.

E. Data processing

The general procedures used to process diffraction intensity measurements into structure factor amplitudes are described here; specific details pertaining to individual data sets are given in later sections. Correction for background, absorption, crystal decay, Lorentz and polarization effects were applied to each set of diffraction intensities. For most data sets, background corrections simply involved, for each individual reflection, the subtraction of the measured background counts from the counts for the peak. Some sets were processed using an alternative handling of backgrounds, in which for each individual reflection, the background correction applied was based on an average of measured backgrounds for neighboring reflections in reciprocal space. The gains yielded by this method are that the merging statistics for multiply measured reflections are significantly improved and the number of negative net intensities is reduced, and were most evident for data sets containing many weak intensities. Treatment of absorption effects employed the empirical method of North *et al.* (1967), in which a phi-dependent correction factor for each reflection is obtained from an averaged absorption curve generated from measurements made every 5° on two strong phi-independent reflections. Corrections for crystal decay were applied based on

an interpolated curve generated from the decay profiles of the standard reflections.

Each set of structure factors was placed on an absolute scale, on the basis of one of two methods. For strongly diffracting crystals from which data had been collected over a large range of theta, Thiessen and Levy's implementation of Wilson's statistical method (Thiessen and Levy, 1973; Wilson, 1942) was used to derive an estimate for the scale factor. For data sets collected from more weakly diffracting crystals, the Wilson method was found to yield unreliable results. For these data sets the scale factor was obtained by performing a linear rescale, based on F_O, against the common reflections in a set of merged, scaled structure factors measured from three crystals of wild-type iso-1-cytochrome *c*.

F. Structure solution of iso-1-cytochrome *c* using molecular replacement methods

1. Construction of a search model for yeast iso-1-cytochrome *c*

For use as a search model in the molecular replacement method, a three-dimensional atomic model of yeast iso-1-cytochrome *c* was constructed based on the known tertiary structures of tuna and rice cytochromes *c*. The initial step in model construction was the incorporation from tuna ferrocytochrome *c* of atoms of the polypeptide backbone, of the heme group, and of amino acid side chains of residues (63 in total) conserved in the primary structures of tuna and yeast iso-1-cytochromes *c*. Then, for amino acid residues (40 in total) with non-identical side chains, the appropriate yeast side chains were appended via a least-squares fitting procedure (R.J. Read, unpublished) based on the positions of the N, CA, and C atoms. The conformation of each new yeast iso-1-cytochrome *c* side chain was modelled, where possible, by maintaining the torsion angles of the original tuna side chain. The homologous residues of rice cytochrome *c* were used to model the first five N-terminal residues of yeast iso-1-cytochrome *c*, which are absent in the tuna protein. These residues were approximately positioned in the yeast iso-1-cytochrome *c* model structure through the least-squares superposition of the alpha-carbon backbones of tuna and rice cytochromes *c*. In order that incorrect bond distances and unfavorable interatomic contacts be removed, the atomic model for yeast iso-1-cytochrome *c* derived to this point was subjected to eight cycles of structural idealization (Hendrickson, 1985).

2. Rotation function search

The set of molecular replacement programs used in these studies was an early version of the MERLOT program package developed by P. Fitzgerald (1988) which was adapted for use on the University of British Columbia Amdahl 5850 computer. The preliminary model constructed for yeast iso-1-cytochrome *c* was oriented, with respect to the crystal axis directions of the unit cell, using Crowther's fast rotation function (Crowther, 1973). Structure factors for the model were calculated with the yeast iso-1-cytochrome *c* coordinates centred at the origin of a triclinic unit cell with orthogonal cell axes 70 Å in length. The overall temperature factor for the model was assumed to be 15 Å^2. An $F \geq 3\sigma_F$ criterion was applied in the selection of reflections used from the measured diffraction data set. The Patterson search of the asymmetric volume of rotation space ($0 \leq \alpha \leq 90°$, $0 \leq \beta \leq 90°$, $0 \leq \gamma \leq 360°$) was carried out using 1.25° steps in α, and 5° steps in β and γ. The search radius about the Patterson origin was limited to 20 Å, and origin removal was applied. Several different resolution ranges were used in the selection of the structure factors to be included in the calculation of the rotation function. Finer scans, employing 1° steps in α, β, and γ, about possible orientations as located by the coarse rotation function, were subsequently conducted using a program written by E.E. Lattman.

3. Searches for the translational positioning of the model for yeast iso-1-cytochrome *c*

a. Translation function search

The translation function of Crowther and Blow (1967) was utilized in initial attempts to determine the translational positioning of the oriented model of yeast iso-1-cytochrome *c* within the crystallographic unit cell. Searches were calculated on a 0.75 X 0.75 Å grid for each of five independent Harker sections, using 946 reflections with $F \geq 3\sigma_F$ in the resolution range 10 to 4.0 Å. In addition, when the search was for the set of intermolecular vectors between two molecules having relative positions dependent on the space group enantiomorph, both of the possible Harker sections were calculated. For example, if molecule 2 is related to molecule 1 by a counterclockwise rotation of 90° about the positive *z* axis, then the set of intermolecular vectors from molecule 1 to molecule 2 appears most strongly on section $w=1/4$ for space group $P4_12_12$,

and on section $w=3/4$ for space group $P4_32_12$. In this case, both $w=1/4$ and $w=3/4$ Harker sections were calculated.

b. Packing analysis

The symmetry elements present in the crystallographic unit cell severely restrict the positioning of molecules within the cell. For instance, a molecule cannot be situated such that its centre is very near a two-fold rotation axis in the crystal if interpenetration by the molecule related to it by this rotation axis is to be avoided. A crystal packing analysis, based on that described by Bott and Sarma (1976), was used to define the regions in the crystallographic unit cell capable of accommodating the oriented, modelled molecule of yeast iso-1-cytochrome *c*. In this procedure, a skeleton consisting of the α-carbons of the 81 residues located at the surface of the molecule, as determined using the method of Connolly (1983), was oriented from the results of the rotation function analysis. This atomic skeleton was systematically translated in 1 Å steps through the region of the unit cell ($0 \leq x \leq 1/2$, $0 \leq y \leq 1$, $0 \leq z \leq 1/2$), and at each position the immediately adjacent, symmetry-related skeletal molecules were generated and the number of close (5 Å or less) intermolecular contacts was calculated. In order that computing time be reduced, contact calculations were not initiated at a given position if two symmetry-related molecules were found to interpenetrate (i.e. if the centres of these two molecules were closer than 25 Å, the estimated miniumum diameter for yeast iso-1-cytochrome *c*), and calculations at a given position were terminated once 15 close contacts had accumulated. The packing analysis was carried out in both enantiomorphic space groups.

c. Correlation coefficient searches

Those regions of the unit cell defined as permissible by the packing analysis were then searched for the translational positioning of the oriented model for yeast iso-1-cytochrome *c* which would yield the greatest agreement between calculated and observed structure factors. This agreement was measured using the following correlation coefficient:

$$\frac{\Sigma\ [(F_o-\bar{F}_o)\cdot(F_c-\bar{F}_c)]}{[\Sigma(F_o-\bar{F}_o)^2\cdot\Sigma(F_c-\bar{F}_c)^2]^{1/2}},$$

where F_o is an observed structure factor magnitude, F_c is the magnitude of the corresponding structure factor calculated for the positioned model for yeast iso-1-cytochrome *c* (all atoms), and $\bar{F}_o$ and $\bar{F}_c$ are the respective average values over the reflections used in the calculation. The value of the correlation coefficient obtained is independent of the relative scales of the observed and calculated structure factor magnitudes, thereby eliminating potential scaling errors. The searches were carried out in both enantiomorphic space groups, and usually on a 1 Å grid. At each point, the correlation coefficient was calculated for the 416 reflections with $F \geq 3\sigma_F$ in the resolution range 10 to 5.0 Å. The computer program used to perform this analysis was written by Dr. W. Hutcheon.

4. Rigid body refinement of orientational and translational positioning

The orientational and translational positioning of the preliminary model for yeast iso-1-cytochrome *c*, as determined by the rotation function, packing analysis and correlation coefficient searches, was refined using a rigid body procedure (Ward *et al.*, 1976). The three rotational and three translational parameters were adjusted so as to minimize the conventional crystallographic R-factor. For each cycle of refinement, each parameter in turn was offset by a small increment (initially, offsets of +/- 0.5° for rotations, and +/- 0.5 Å for translations were used), while the other parameters held their original values. The R-factor, for the 581 reflections with $F > 3\sigma_F$ in the resolution range 10 to 4.5 Å, was calculated for the positioned model after each adjustment. At the start of the next cycle, each parameter was assigned its value that gave the lowest R-factor during the previous cycle. The size of the increments for each of the parameters was decreased as the analysis proceeded, and the refinement was concluded when no further reduction in R-factor could be obtained.

G. Structure refinement procedures

The stereochemically restrained least-squares method of Hendrickson and Konnert (1981) was used to refine the structures of yeast iso-1-cytochrome *c*. This method adjusts the atomic parameters of the model to improve both the agreement between observed and calculated structure factor magnitudes, and the accordance of bonded and non-bonded interatomic distances present with those occurring in small-molecule structures. Stereochemical restraints were applied to: covalent bond lengths, bond angles, planarity (of amide, carboxyl and guanidinium groups, and aromatic rings), chirality at asymmetric carbon atoms, non-bonded interatomic contacts, hydrogen bond distances, conformational torsion angles, and similarity in temperature factors of covalently bonded atoms.

For all refinements, data of lower resolution than 6 Å were excluded, as these are most severely affected by disordered solvent (Schoenborn, 1988). In addition, a criterion of either $F \geq 2\sigma_F$ or $F \geq 3\sigma_F$ was used in the selection of reflections to be included.

Most of the refinement cycles were computed on an IRIS 3030 workstation, although some of the early refinement cycles for the wild-type iso-1-cytochrome *c* structure were performed on the Amdahl 5850. The same general strategy was used in the refinements of all structures. The emphasis placed by the refinement program on both stereochemical ideality and agreement between observed and calculated structure factors can be varied. Typically, fairly tight weightings on ideal geometry (see Table II.2) were maintained throughout the course of the refinements. In some cases after the atomic model had been adjusted manually, restraint weightings on bond distances were loosened slightly, and then subsequently tightened after two or three cycles. The weighting on structure factors was maintained at 1/2 to 1/3 the overall average discrepancy between F_o and F_c. All atoms in the atomic models, including those belonging to solvent molecules, were refined with full occupancies. Progress in the refinement, in terms of the drop in R-factor and the magnitude of positional shifts yielded, declines markedly after about five consecutive cycles, and thus usually refinement was interrupted after not more than eight cycles so that the model could be examined in electron density maps.

Table II.2. Typical weighting on stereochemical ideality used in restrained least-squares refinement of iso-1-cytochrome *c* structures

Stereochemical restraint type	Refinement weighting used
1-2 bond distance	0.020 Å
1-3 angle distance	0.030 Å
1-4 planar distance	0.050 Å
1-2 bond involving hydrogen atom[a]	0.012 Å
1-3 bond involving hydrogen atom	0.027 Å
Planar	0.020 Å
Chiral centre	0.150 Å^3
Non-bonded contact[b]	
single torsion	0.250 Å
multiple torsion	0.250 Å
possible hydrogen bond	0.250 Å
Staggered ($\pm 60°$, $180°$) torsion angle	19.0°
Planar torsion angle (0 or $180°$)	3.0°
Temperature factor	
1-2 bond (main chain)	1.20 Å^2
1-3 angle (main chain)	1.50 Å^2
1-2 bond (side chain)	1.70 Å^2
1-3 angle (side chain)	2.00 Å^2
1-2 bond involving hydrogen atom	3.00 Å^2
1-3 bond involving hydrogen atom	5.00 Å^2

[a]The restraints involving hydrogen atoms apply only to the refinement of the wild-type iso-1-cytochrome *c* structure.

[b]This class of restraint incorporates a reduction of 0.1 Å from the radius of each atom involved in a contact.

H. Electron density maps

1. Map inspection

When necessary, manual adjustments to the atomic models were made through inspection of difference electron density maps. Electron density maps were computed using the programs of the ROCKS system (Reeke, 1984), or the fast Fourier program FSFOR written by W. Furey and adapted to the IRIS 3030 by A. Berghuis. The coefficients used were of the type ($2F_o-F_c$, α_c), (F_o-F_c, α_c), and (F_o-F_c', α_c') where F_c' and α_c' represent structure factors and phases calculated

without the inclusion of the portion (up to 9%) of the atomic model to be examined. Map interpretation was assisted by the use of the program FRODO (Jones, 1978), which displays a superposition of the atomic skeleton on envelopes indicating the spatial distribution of electron density. The version of FRODO used was adapted to the IRIS 3030 by S.J. Oatley, and was further modified by S.A. Evans.

2. Modelling of solvent structure

Water molecules were introduced into significant peaks of otherwise unoccupied electron density observed in $2F_o-F_c$ or F_o-F_c maps, and only at positions where the new water molecule could form hydrogen bonds with protein atoms or other (previously identified) water molecules. The authenticity and positioning of all water molecules were monitored throughout the course of refinement, using $2F_o-F_c$ and F_o-F_c' maps. In the latter case, water molecules were deleted from the atomic model, and were retained in the refinement only if electron density peaks subsequently appeared at the sites they originally occupied.

3. Preliminary difference maps in the analysis of variant yeast iso-1-cytochromes *c*

The first step in the structural analysis of a variant iso-1-cytochrome *c* was the inspection of difference electron density maps. A map with the coefficients F_o(wild-type) minus F_o(variant) and phases calculated for the wild-type structure was used initially to locate significant structural differences between the wild-type and variant iso-1-cytochromes *c*. For the calculation of these maps, it was found that unless the input reflections were carefully selected so as to include only those having significant intensity differences, the resultant maps contained appreciable noise and were difficult to interpret. When comparing diffraction intensities measured from crystals of wild-type iso-1-cytochrome *c* with those measured from crystals of the variant proteins a problem occurs for reflections which are observed in only one of the two data sets: it is uncertain whether the observed intensity difference arises from structural differences between the two iso-1-cytochromes *c*, or is due to errors in intensity measurement. A primary consideration in this respect is that the crystals used of the variant iso-1-cytochromes *c* diffracted more weakly,

resulting in a greater number of weak intensities and also in greater uncertainty in the measured intensities in the variant data sets. The strategy chosen in an effort to accommodate these problems was to include in the calculation of the difference map only those reflections which met one of the following conditions: the reflection was observed ($F \geq 3\sigma_F$ for the variant, and $F \geq 2\sigma_F$ for the wild-type data set) in both data sets; or the reflection was unobserved in one data set, but fairly strong ($F \geq 3\sigma_F$ in the variant, and $F \geq 6\sigma_F$ for the wild-type data set) in the other. Usually, regions of the iso-1-cytochrome *c* molecule identified by the F_o(wild-type)-F_o(variant) map as being perturbed structurally were subsequently examined in a $2F_o$(variant)-F_c' map. In this latter map, the region of interest is excluded from the calculation of structure factor magnitudes and phases, thus minimizing the bias toward the conformation of this region present in the wild-type structure.

I. Inclusion of hydrogen atoms

Hydrogen atoms were refined in the atomic model of wild-type iso-1-cytochrome *c*. They were included at a point in the refinement when it was judged that for many hydrogen atoms, particularly those on main chain amide nitrogens and well-ordered aliphatic and aromatic carbons, electron density was visible on $2F_o-F_c$ maps as small bulges protruding from the density belonging to their heavier parent atoms. It was thus felt that inclusion of hydrogens in the refinement model would improve both the bonding geometry and thermal parameter values of the heavy atoms, and the agreement between calculated and observed structure factors.

Hydrogen atoms were initially introduced into the iso-1-cytochrome *c* structure in ideal X-ray positions using the program HAFFIX (Honzatko *et al.*, 1985), and were assigned a temperature factor of 1.3 times that of the parent heavy atom. Hydrogens having ambiguous positions, such as those on hydroxyl groups, were positioned manually so as to form reasonable hydrogen bonds. Water molecules were left as neutral oxygen atoms. During structure refinement, separate restraint weighting was maintained for bonding geometry and thermal parameters involving hydrogen atoms (Wlodawer and Hendrickson, 1983) (see Table II.2). Tight restraints were applied to bonding geometry, so that hydrogen atoms would be tightly coupled to their parent atoms. The

thermal parameters of hydrogen atoms were not as tightly restrained, but in most cases showed no tendency to assume values greatly different from those of their parent atoms. The torsion angles of rotor methyl carbons and amine nitrogens were not restrained.

J. Comparison of cytochrome *c* structures

Tuna [both reduced (1.5 Å) and oxidized (1.8 Å)] (Takano and Dickerson, 1981a,b) and rice (1.5 Å) cytochromes *c* (Ochi *et al.*, 1983) are determined to resolutions comparable to that of yeast iso-1-cytochrome *c* (wild-type). Coordinates for these other cytochromes *c*, entries 4CYT, 3CYT and 1CCR respectively, were obtained from the Protein Data Bank (Bernstein *et al.*, 1977). The direct structural comparison of the various species of cytochrome *c* required that these other cytochromes *c* be superposed on yeast iso-1-cytochrome *c*. This was carried out by using an iterative least-squares procedure to determine for each cytochrome *c*, the rotation and translation to be applied in order to minimize the positional difference between corresponding main chain atoms (residues 1 to 103 only) of yeast iso-1- and this other cytochrome *c*. The rotation matrices and translation vectors used are listed in Table II.3. Because the variant iso-1-cytochromes *c* were crystallized isomorphously with wild-type iso-1-cytochrome *c*, direct structural comparisons of the variant and wild-type proteins did not require this preliminary superposition.

Since the structure of yeast iso-1-cytochrome *c* is in the reduced state, most of the comparisons pertaining to the tuna molecule discussed herein refer to tuna ferrocytochrome *c*. Specific instances where comparisons are made with tuna ferricytochrome *c* are indicated.

Table II.3. Rotation matrices and translation vectors applied for structural superposition of tuna and rice cytochromes *c* onto yeast iso-1-cytochrome *c*

	Rotation matrix			Translation vector
a. Tuna (reduced)				
	0.27945	0.38538	0.87943	-7.79118
	-0.70389	-0.54071	0.46062	12.86195
	0.65303	-0.74774	0.12016	-7.61227
b. Rice (oxidized)				
	-0.67668	-0.00102	-0.73628	22.80440
	0.58991	0.59763	-0.54299	22.47183
	0.44057	-0.80177	-0.40380	35.40532
c. Tuna (oxidized, outer molecule)				
	0.55471	-0.18403	0.81144	-20.39990
	0.68101	0.66073	-0.31569	15.07751
	-0.47805	0.72771	0.49184	10.90701
d. Tuna (oxidized, inner molecule)				
	-0.05221	-0.59576	-0.80146	18.20815
	-0.87437	-0.36044	0.32489	55.23132
	-0.48244	0.71774	-0.50210	10.77491

The application of the rotation matrix and vector which superposes the atomic coordinates (x_o, y_o, z_o) of tuna or rice cytochrome *c* onto those of yeast iso-1-cytochrome *c* is as follows:

$$\begin{bmatrix} \text{Rotation} \\ \\ \text{matrix} \end{bmatrix} \begin{bmatrix} x_o \\ y_o \\ z_o \end{bmatrix} + \begin{bmatrix} \text{V} \\ \text{e} \\ \text{c} \end{bmatrix} = \begin{bmatrix} x_o^y \\ y_o^y \\ z_o^y \end{bmatrix}.$$

III. THE STRUCTURE OF WILD-TYPE YEAST ISO-1-CYTOCHROME *c*

A. Details of the structure solution

1. X-ray diffraction data

Diffraction data to 1.23 Å resolution for wild-type iso-1-cytochrome *c* were collected from six crystals (see Table II.1), as detailed in Table III.1. Background averaging was used in the processing of all diffraction data beyond 2.0 Å resolution. The data sets B, C, and F were placed on an absolute scale using the method of Wilson (1942) (Thiessen and Levy, 1973). The

Table III.1. Crystals of wild-type iso-1-cytochrome *c* used in data collection to 1.23 Å resolution

Data set	Resolution range (Å)	Number of reflections[a]		Scale factor[b]	Merging R[c] (on I)(%)	Scaling R[d] (on F) (%)
		measured	unique			
B	20 - 2.5	11568	6860	14.91	2.7	-
C	20 - 1.7	11530	10940	8.25	2.2	-
D	1.75 - 1.50	6101	5831	9.11	1.1	4.6
E	1.75 - 1.40	11114	10486	7.85	1.5	5.2
F	20 - 2.00	12079	6812	7.89	4.1	-
G	1.50 - 1.23	14532	13868	8.32	2.8	5.6

[a]The unique region of reciprocal space is 1/16 of the sphere. The volume normally collected was h: $0 \longrightarrow \infty$, k: $0 \longrightarrow$ h, l: $0 \longrightarrow \infty$. For crystal B, both hkl and khl were measured, and for crystal F both hkl and -h,-k,-l.

[b]The scale factors given are the multiplicative factors which place the structure factors in each individual data set on an absolute scale.

[c]The merging R is defined as $\dfrac{\sum_{hkl} \sum_{i=1}^{n} | I_i(hkl) - I(hkl) |}{\sum_{hkl} \sum_{i=1}^{n} I_i(hkl)}$.

The calculation of the merging R includes all reflections measured more than once (i.e. duplicates, symmetry mates and Friedel mates) and for data sets with a small amount of replicated data (sets C, D, E and G) is dominated by the measurement of intensity control reflections.

[d]The scaling R is calculated including only those reflections in the two-dimensional zone (hk0 or h0l) used in the linear rescale against the same set of merged reflections from Crystals B, C and F. The definition of the scaling R is the same as that for the merging R, except F is used in place of I.

remaining data sets were placed on the same scale by performing a linear rescale against a set of merged data from crystals B, C and F, using as a basis a two-dimensional zone of structure factors from 20 to 2.0 Å (h0l, 955 reflections, for data sets E and G; or hk0, 135 reflections, for set D) that was measured from all crystals. Data sets B and C were used in the molecular replacement analyses.

For use in high resolution refinement, a complete data set to 1.23 Å was obtained by merging the six independent sets of structure factors, as outlined in Table III.2. The majority of the data from 20 to 1.4 Å has been measured from at least two separate crystals. The fairly high merging R-factors for data of greater resolution than 1.7 Å are typical of high resolution analyses (Karplus and Schulz, 1987) and are indicative of the overall weaker intensity of high resolution

Table III.2. Summary of merging of data sets for wild-type iso-1-cytochrome *c*

Resolution range (Å)	Crystals used	Independent measurements	Unique reflections	Merging R[a] (on F) (%)
20 - 2.50	B, C, F	10633	3602	4.7
2.50 - 2.00	C, F	6423	3215	9.3
2.00 - 1.75[b]	C,D,E,G	3845	3204	22.8
1.75 - 1.70	C, D, E	2629	890	21.5
1.70 - 1.50	D, E	9502	4751	15.1[c]
1.50 - 1.40	E, G	6766	3383	20.7[c]
1.40 - 1.23	G	8741	8741	-
6.00 - 1.23	All	48539	27786	8.1

[a]The definition of this merging R is the same as that given in the footnotes to Table III.1. In each specified resolution range, the merging R value calculated includes reflections which have been measured from more than one crystal.

[b]In this resolution range, only hk0 (C and D) or h0l (C, E and G) reflections had measurements from more than one crystal.

[c]The calculated merging R value for these shells includes only those replicately measured reflections for which $F_{ave} \geq 3\sigma_F$.

reflections.

2. Structure solution using molecular replacement

a. Rotation function

Several trials of the rotation function, using various resolution limits in the selection of diffraction data to be included in the calculation, were performed in attempts to determine the orientation of the model of yeast iso-1-cytochrome *c*. In retrospect, it is evident that as higher resolution data were included, the rotation function solution obtained more closely approached the optimal orientation (see Table III.3), although in all trials, the strongest peak corresponded to the correct solution. For the Crowther fast rotation function, the discrimination of the correct solution from both background and false peaks also improved with the inclusion of higher resolution data. It was based on this consideration that the orientation chosen for use in further molecular replacement searches was the result of trial D ($\alpha = 28.2°$, $\beta = 83.9°$, $\gamma = 228.5°$).

Table III.3. Summary of results of the rotation function searches for yeast iso-1-cytochrome *c*

	Rotation function	Resolution limits (Å)	Model	α	β	γ	Peak height[a] (σ)	Next highest peak[a] (σ)
A	Coarse	10.0 - 5.0		32	83.6	217.3	4.4	3.8
	Fine	10.0 - 4.5		31	85	227		
B	Coarse	25.0 - 4.0	all	30	85	230	5.0	4.2
	Fine	25.0 - 4.0	atoms	30.6	86.5	226.5		
C	Coarse	10.0 - 4.0		29.6	83.2	225.2	6.0	4.5
	Fine	10.0 - 3.5		29.6	84	227.8		
D	Coarse	10.0 - 2.9	752	28.8	82.9	228	7.4	5.6
	Fine	10.0 - 3.1	atoms	28.2	83.9	228.5		
	Optimal orientation and translational positioning[b]			27.1 $x=0.115$	84.2 $y=0.545$	229.0 $z=0.0584$		

[a]Peak heights are given as multiples of the r.m.s. background level.

[b]The optimal orientation and translational positioning are taken to be those determined using rigid body refinement (see Section III.A.2.c).

That the performance of the rotation function improved with the inclusion of higher resolution diffraction data is in accord with the high degree of similarity between the structure of the search model used and that of the actual yeast iso-1-cytochrome *c* molecule. In contrast, inclusion of high resolution data has been found to affect adversely the accuracy of the rotation function in cases where the search model bears lesser structural homology to the actual molecule (Read and James, 1988). An additional factor in trial D yielding the best results may be that an edited search model was used in this run. The editing procedure removed portions of the search model having positions which may have been poorly predicted in the model-building process (atoms beyond CB of side chains of residues with high sequence variability within eukaryotic cytochromes *c*, and of lysine and glutamate side chains which projected into the solvent medium) and left 752 of the original 890 atoms.

b. Translational Searches

The use of the translation function of Crowther and Blow was unsuccessful in determining the translational positioning of the oriented model of yeast iso-1-cytochrome *c*. The most serious problems encountered were the low signal-to-noise ratio and the large number of peaks observed. Typically, the strongest peaks were only 2 to 3 times the r.m.s. background level and were only marginally higher than the next largest peak. A subsequent re-examination of the translation function maps, after the correct translational positioning had been determined, revealed that one search, involving the molecules at (x, y, z) and at ($-x$, $-y$, $1/2+z$), had actually yielded the correct solution as the strongest peak. However, no pair of translation function searches on individual Harker sections yielded a consistent value for any of the components of the translational positioning vector. On the other maps, the correct solution was found to correspond to the 5th, 6th or 7th largest peaks. A possible reason for the failure of the translation function is the relatively inaccurate orientational parameters used ($\alpha = 30.6°$, $\beta = 86.5°$, $\gamma = 226.5°$).

A packing analysis was used to define the regions of the unit cell in which the oriented model of yeast iso-1-cytochrome *c* could be positioned without conflicting with neighboring molecules generated by the crystallographic symmetry elements. This analysis located 7 distinct

regions (for each of the possible enantiomorphic space groups) at which fewer than 6 intermolecular contacts would occur. [For example, for the space group $P4_32_12$, these regions are: (0.0823± 0.0549, 0.521± 0.027, 0.0511± 0.0219); (0.480± 0.014, 0.974± 0.014, 0.0438± 0.0146); (0.0, 0.5, 0.161± 0.0146); along the line joining the points x=0.0, y=0.5 and x=0.5, y=0.0, with z=0.263± 0.0146; (0.480± 0.014, 0.041± 0.014, 0.321± 0.0146); (0.027± 0.014, 0.562± 0.014, 0.438± 0.0292); and (0.5, 0.027± 0.014, 0.431± 0.0219).] It is of note that this packing analysis was very restrictive in defining the molecular positioning consistent with the occupation of the observed crystallographic unit cell by eight iso-1-cytochrome *c* molecules: of 24605 grid points examined, only 306 (1.2%) were identified as being suitable positions for the centre of the yeast iso-1-cytochrome *c* model. Whereas this analysis restricted subsequent translational searches to 1.2% of the total search volume of the unit cell, a value of 13% was yielded by a simpler packing function that minimizes the interpenetration of solid sphere representations of the molecule (Hendrickson and Ward, 1976).

The translational positioning identified as being allowable based on packing can be understood upon considering the crystallographic symmetry elements present and the relative dimensions of the unit cell and of the cytochrome *c* molecule. That the *a* and *b* crystallographic cell axis lengths (36.46 Å) are only slightly longer than the approximate diameter of the cytochrome *c* molecule (25-30 Å) requires that a molecule centre be located as distant as possible from the 2_1 screw axes parallel to *a* and *b*, which lie in the $z=1/8$ and $z=3/8$ planes respectively. Also, a molecule centre cannot be positioned near the diagonal two-fold axes, which lie on the $z=0$, 1/4, and 1/2 planes. Thus the requirement for remoteness from both these 2_1 and diagonal 2-fold axes restricts the molecule centres to a positon most ideally very near the 4-fold screw axis at a *z* value of 1/16, 3/16, 5/16 or 7/16, or possibly adjacent to the diagonal 2_1 screw axes.

The next step involved translational searches for the maximum correlation coefficient between calculated and observed structure factors. The initial search was two-dimensional, using only 54 hk0 reflections (in the resolution range 10-5.0 Å) to determine the *x* and *y* components

of the translational positioning vector. This analysis, which is independent of the choice of the two enantiomorphic space groups, revealed that a correlation coefficient of 0.785 is obtained with the centre of the oriented model of yeast iso-1-cytochrome *c* positioned at $x=0.11$, and $y=0.55$. The results of the packing analyses indicated that for this positioning of x and y, the only z positioning allowed was 0.0438-0.0585 in the space group $P4_32_12$. (No solution could be found in space group $P4_12_12$.) A correlation coefficient search around this location yielded a value of 0.46 (10-5.0 Å, 416 hkl reflections) when the oriented model of yeast iso-1-cytochrome *c* was positioned at (0.110, 0.535, 0.0585), a point at which the packing analysis had indicated a total of 3 intermolecular contacts. Subsequently, all other regions identified as being allowed by packing, in both space groups, were searched. No correlation coefficient greater than 0.30 was found, which is indicative of the high degree of discrimination for the correct solution by the approach described. Also, on the basis of these results, the correct enantiomorphic space group was taken to be $P4_32_12$.

c. Rigid body refinement

The orientational and translational positioning of the model of yeast iso-1-cytochrome *c* determined using the rotation function, packing and correlation coefficient searches, ($\alpha=28.2^\circ$ $\beta=83.9^\circ$, $\gamma=228.5^\circ$; $x=0.110$, $y=0.535$, $z=0.0585$), was optimized through rigid body refinement. Systematic adjustment of each of the six parameters about their original values yielded a lowering in the R-factor from 0.54 to 0.51 (for 581 reflections, 10-4.5 Å), and the optimal positioning listed in Table III.3.

3. Structure refinement

Diffraction data of increasingly higher resolution were included as the refinement progressed. The course of the refinement can be conveniently discussed as two stages: to 2.8 Å resolution, fixing of the gross positioning of the molecule in the unit cell and correction of major errors in the conformation of the starting model; and at high resolution, attainment of maximum accuracy in atomic positions of the model.

The refinement to 2.8 Å included 82 least-squares cycles and two rounds of manual intervention as outlined in Figure III.1. Diffraction data ($F \geq 3\sigma_F$) from crystal C was used. Initially, an overall isotropic temperature factor of 14.8 Å^2 (as estimated by Wilson's statistical analysis) was assumed. Restrained refinement of individual atomic temperature factors was initiated after cycle 27. Major adjustments to the polypeptide backbone were performed manually on residues −5 to 2, 21 to 22, 52 to 63, and 85 to 88. Only 19 water molecules were added to the atomic model in this stage, after cycle 76. The overall r.m.s. shift in atomic positions between the

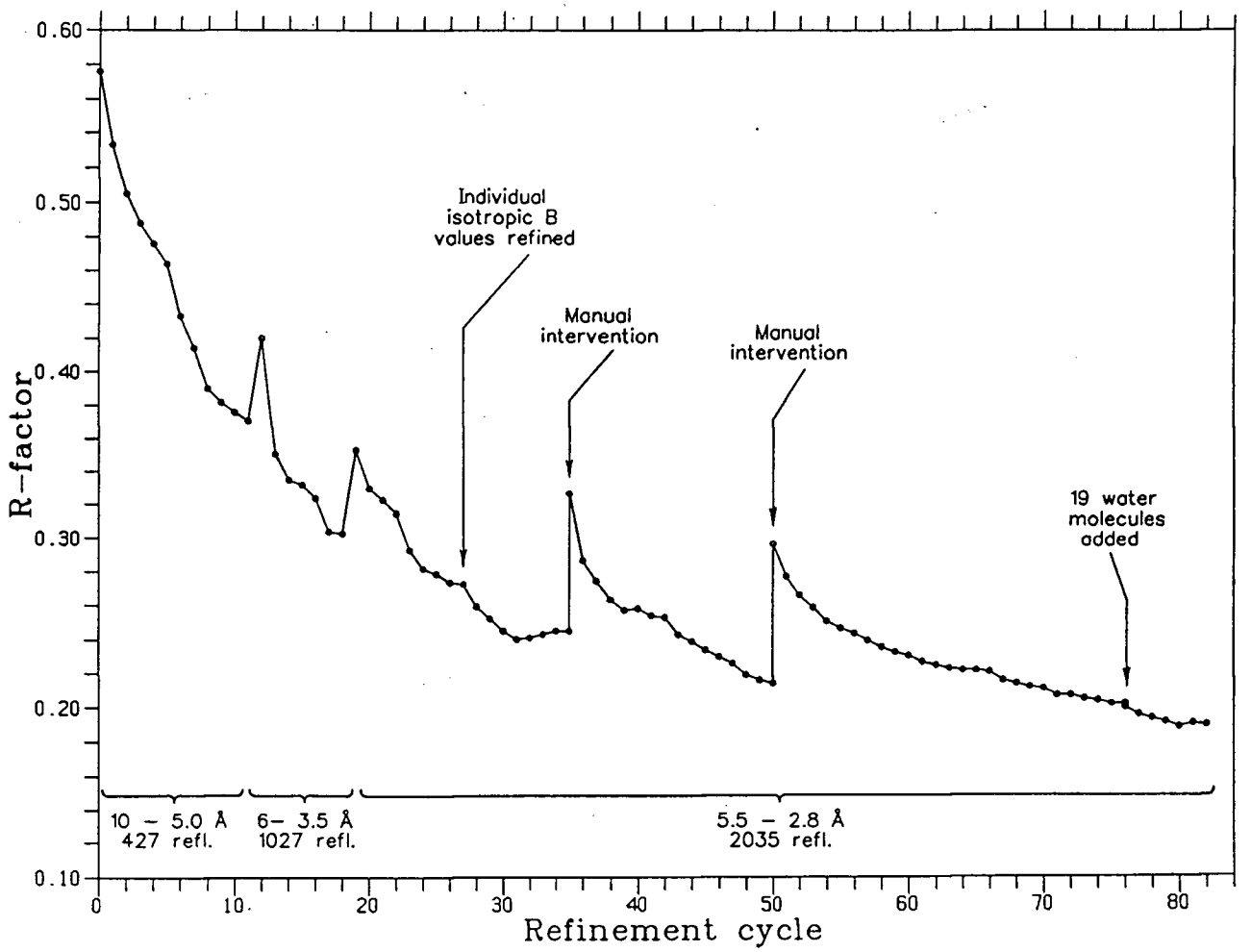

Figure III.1. Course of structural refinement of wild-type yeast iso-1-cytochrome *c* to 2.8 Å resolution. The plot shows the crystallographic R-factor as a function of the refinement cycle number. The resolution ranges of the diffraction data used are indicated below the plot. Points at which major revisions to the refinement model occurred are noted.

2.8 Å structure and the starting model was 1.0 Å.

High resolution structure refinement (Figure III.2) was initiated with a 2.8 Å resolution model refined to an R-factor of 0.192. The sigma cutoffs applied were $F \geq 2\sigma_F$ for data from 6.0–1.4 Å, and $F \geq 3\sigma_F$ for data from 1.4–1.23 Å. A more stringent rejection criterion was used for reflections below 1.4 Å resolution because of the greater uncertainty in these measurements and the lack of multiple observations (see Table III.2). The percentage of the total possible

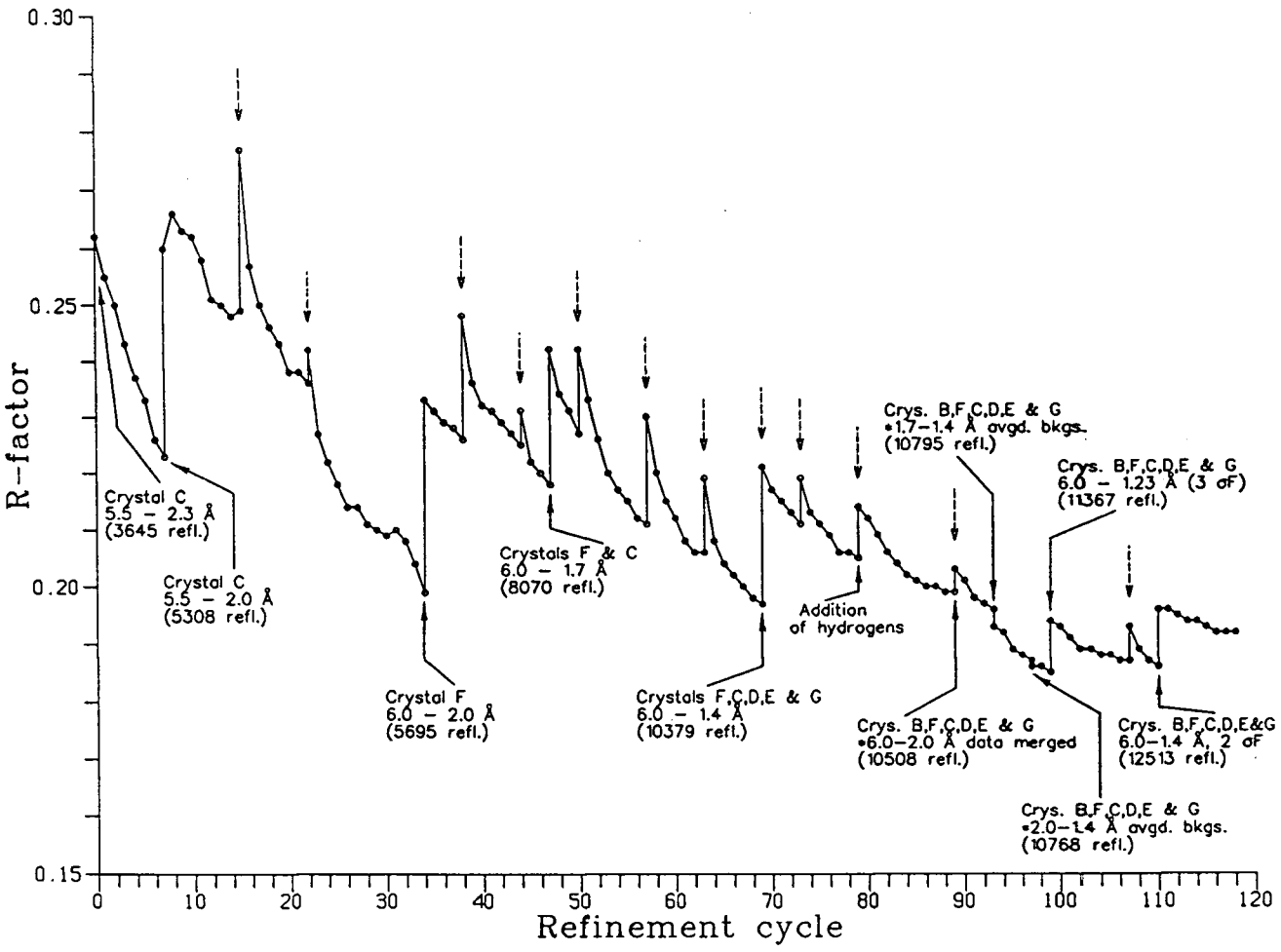

Figure III.2. Course of structural refinement of wild-type yeast iso-1-cytochrome *c* at high resolution. The plot shows the crystallographic R-factor as a function of the refinement cycle number. Points at which manual adjustments were made to the model are indicated by downward-pointing dashed arrows. Remarks have been included to indicate points at which changes were made in the selection of reflections used in the refinement process. Asterisks indicate points at which reprocessing of the reflection data occurred. Tables III.1 and III.2 provide additional documentation on the structure factor data sets used.

data used as a function of resolution is shown in Figure III.3. There is a significant falloff in the amount of data used beyond ~1.5 Å resolution. For the resolution range 1.4–1.23 Å, less than 7% of the theoretically available data has been included in the current structure refinement. Because most of the required corrections to the atomic model had been made during the earlier stage, in general, further manual adjustments to the atomic model were minor. Most corrections were restricted to the side chains of amino acids, with particularly large adjustments being made to Ser2, Thr8, Thr12, Gln16, Ser47, Asn62, Tml72, Arg91 and Glu103. In addition, a number of surface

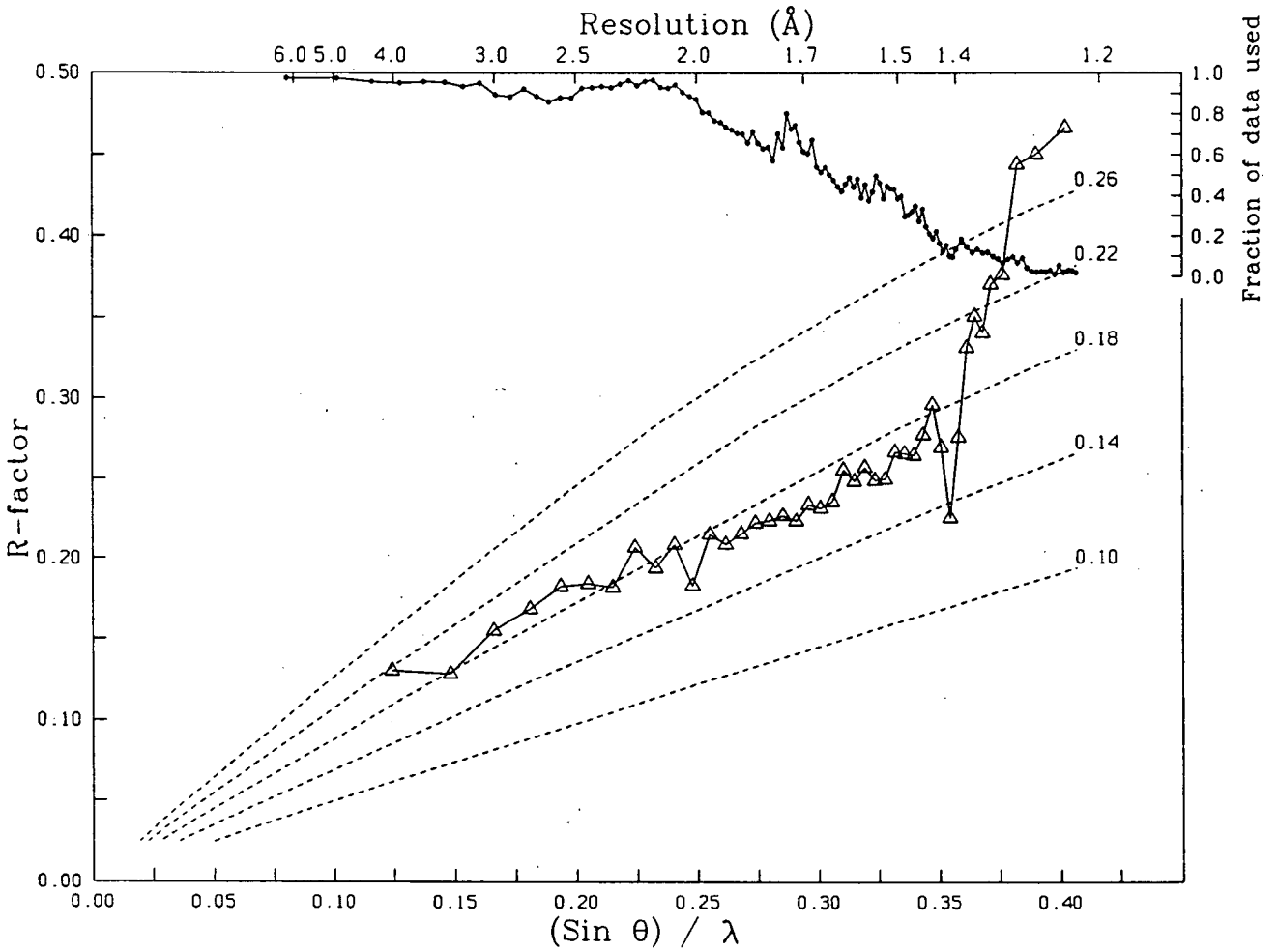

Figure III.3. Dependence on resolution of the R-factor and the percentage of available data used for the final refined structure of yeast iso-1-cytochrome *c*. In this analysis, reciprocal space was divided into shells according to $\sin\theta/\lambda$, each shell containing at least 150 reflections. For each shell, both the agreement between calculated and observed structure factors (Δ—Δ), and the percentage of data used (o—o, axis at top right) are shown. For the purpose of assessing the accuracy of the atomic coordinates of the yeast iso-1-cytochrome *c* structure, 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).

lysine and glutamic acid side chains, as well as the polypeptide chain backbone of residues -5 to -2, and 54 to 58 required frequent adjustments throughout the course of refinement. Many additional water molecules (116 occur in the final model) and also 1 sulfate ion (see discussion below, Section III.D.4) were located and incorporated into the refinement model. Hydrogen atoms were included in the atomic model when the structure of yeast iso-1-cytochrome *c* had been refined to an R factor of 0.205 at 1.4 Å.

4. Quality of the final structural model

The agreement with ideal stereochemistry in the final refined structure of yeast iso-1-cytochrome *c* is summarized in Table III.4. Overall, deviations from ideality are reasonably small for all classes of geometrical restraints. The overall R factor for the 12513 reflections used from 6.0 to 1.23 Å is 0.192. The agreement between the observed and calculated structure factors as a function of resolution is shown in Figure III.3. Curves representing the theoretical dependence of R factor on data resolution, assuming various values for the r.m.s. error in a set of atomic coordinates (Luzzati, 1952), are also drawn in Figure III.3. This analysis gives an estimate of 0.18 Å for the r.m.s. error in atomic positions in the refined structure of iso-1-cytochrome *c*. This value should be regarded as an over-estimate of the error, as it assumes that discrepancies between observed and calculated structure factors are due entirely to errors in the coordinate set, and neglects considerations such as errors in intensity measurements and the omission of the solvent continuum (Chambers and Stroud, 1979). The error in atomic positions is likely much lower than 0.18 Å for well-defined portions of the structure, but undoubtedly much higher for mobile surface side chains.

A separate estimate of the atomic coordinate error can be obtained from an unrestrained refinement of the final model (Read *et al.*, 1983). For yeast iso-1-cytochrome *c*, this analysis was carried out without hydrogen atoms present and resulted in a lowering of the R-factor from 0.202 to 0.184 in three cycles. Atomic positions were found to deviate by an overall r.m.s. value of 0.12 Å from those of the final restrained model.

Table III.4. Agreement with ideal stereochemistry in the final refined model of wild-type yeast iso-1-cytochrome *c* at 1.23 Å resolution

Class of restraint	R.m.s. deviation from ideality
1-2 bond distance	0.026 Å
1-3 angle distance	0.050 Å
1-4 planar distance	0.062 Å
1-2 bond involving hydrogen atom	0.011 Å
1-3 bond involving hydrogen atom	0.021 Å
Planar	0.024 Å
Chiral centre	0.179 Å^3
Non-bonded contact[a]	
single torsion	0.173 Å
multiple torsion	0.185 Å
possible hydrogen bond	0.261 Å
Staggered (±60°, 180°) torsion angle	19.3°
Planar torsion angle (0 or 180°)	4.3°
Temperature factor	
1-2 bond (main chain)	1.7 Å^2
1-3 angle (main chain)	2.3 Å^2
1-2 bond (side chain)	2.9 Å^2
1-3 angle (side chain)	4.4 Å^2
1-2 bond involving hydrogen atom	2.0 Å^2
1-3 bond involving hydrogen atom	2.9 Å^2

[a]The r.m.s. deviations from ideality for this class of restraint incorporate a reduction of 0.1 Å from the radius of each atom involved in a contact.

Figure III.4 shows the high quality of the electron density maps for iso-1-cytochrome *c*. In the electron density map of the heme group, dimples are apparent in the centres of the five-membered pyrrole rings and small bulges of electron density belonging to the hydrogen atoms of the methine and side chain methyl carbons are visible. Figure III.4b shows that a dimple is also present in the centre of the pyrrolidine ring of Pro71, and that electron density can be seen for hydrogen atoms of the side chains of Tyr67 and Leu68. The high quality of such electron density maps is also evident in the observation that for atoms with temperature factors lower than about 20 Å^2, carbon, nitrogen and oxygen atoms can be readily distinguished from each other by

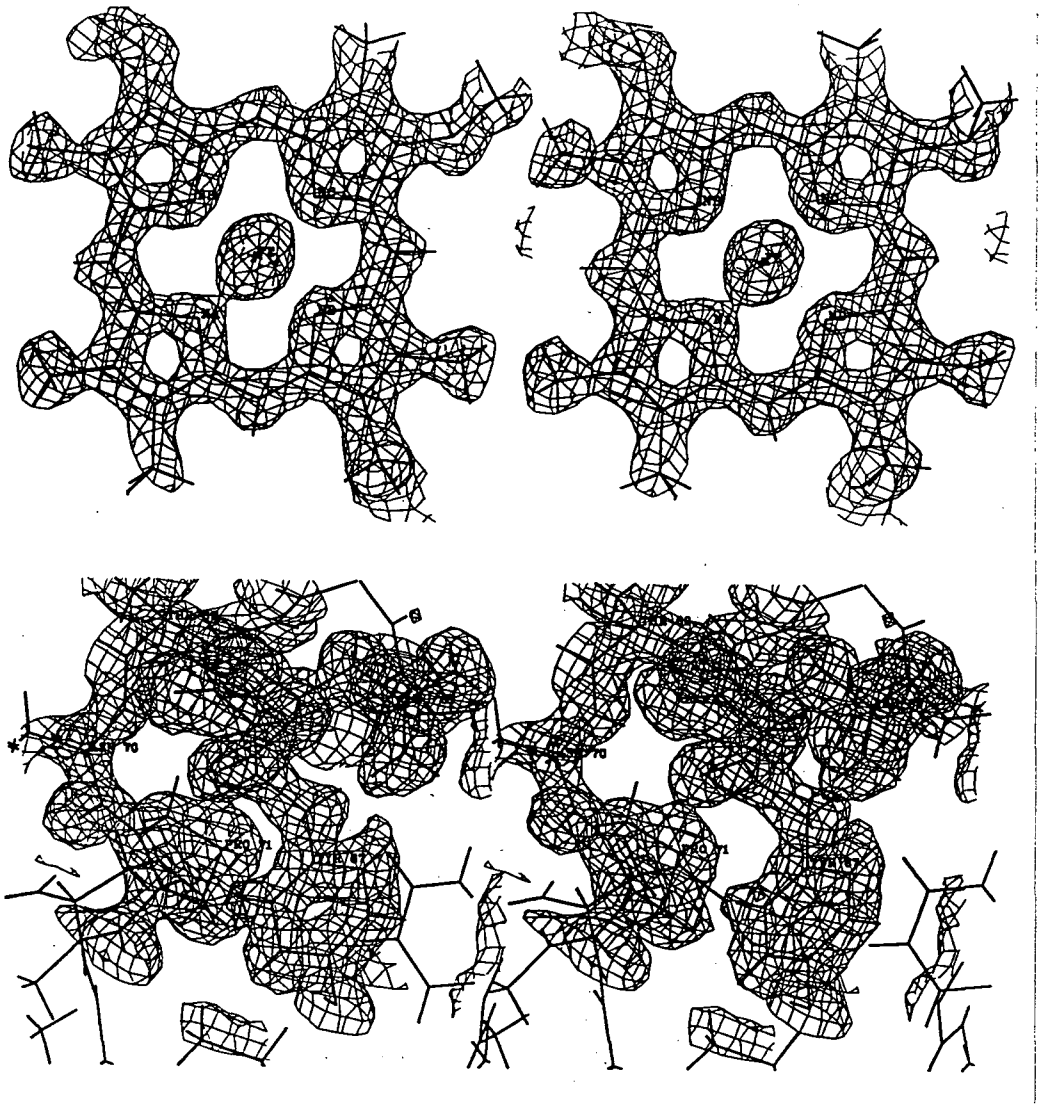

Figure III.4. Stereo diagrams of portions of the electron density map calculated for the final model of yeast iso-1-cytochrome *c*. The regions of the molecule shown are (*top*) the heme group, and (*bottom*) the polypeptide chain from residues 67 to 71. These maps were calculated with coefficients of the type $2F_o-F_c$, α_c and have been contoured at an electron density level of 0.5 e/Å^3.

the height of the electron density peak centred on each atom.

5. Effect of the inclusion of hydrogen atoms

An analysis of the effects of including hydrogen atoms in the refinement process was made by deleting these atoms from the final atomic model and running 6 further refinement cycles. The most notable trend indicated was a downward shift in the temperature factors of atoms having

one or more hydrogen atoms bonded to them (Table III.5). This suggests that for protein structures refined without the inclusion of hydrogen atoms, temperature factors of non-hydrogen atoms will tend to be systematically too low due to their compensating for electron density belonging to the missing hydrogen atom(s). The effect of inclusion of hydrogens on the agreement between calculated and observed structure factors is a decrease of 1.0% in the overall R-factor for data to 1.23 Å. The most significant improvement occurs at low to moderate resolution (> 2.0 Å), as expected since hydrogen atoms make little contribution to the high angle scattering (Stout and Jensen, 1968).

At the completion of refinement of the structure of yeast iso-1-cytochrome *c*, an assessment was made of the validity of including hydrogen atoms. The majority of the hydrogen atoms were found to be well-behaved, as they clearly reside in small lobes of electron density (see Figure III.4) and possess temperature factors similar to those of their parent heavy atoms. Furthermore, as Table III.4 shows, deviations of hydrogen atoms from their ideal positions with respect to their parent atoms is negligible. Problems that were apparent arise at heavy atoms

Table III.5. Effect of inclusion of hydrogen atoms on refined atomic temperature factors

Atom type	Average shift in temperature factor (in Å²) for atoms having the indicated number of bonded hydrogens			
	0	1	2	3
N	+0.42 (4)	-0.72 (104)		
CA		-0.58 (96)	-1.09 (12)	
C'	-0.12 (108)			
O	+0.14 (109)			
side chain C	+0.65 (76)	-0.21 (111)	-0.22 (165)	-0.27 (65)

Each entry represents the average shift in temperature factor, after removal of hydrogen atoms and refinement for an additional 6 cycles. Atoms have been grouped according to atom type and the number of bonded hydrogen atoms, with the number of atoms in each subclass given in parentheses. The C' type refers to carbonyl carbons.

whose electron density peaks are elongated in shape, due to either thermal motion or static disorder. At these sites, hydrogen atoms residing in the electron density smears had refined temperature factors that were much lower than would be expected. In addition, in cases where a hydrogen atom was free to rotate about its heavy atom (e.g. hydroxyls and primary amines), it sometimes refined into these smears of electron density rather than into apparently obvious hydrogen bonding positions.

B. Description of the conformation of the iso-1-cytochrome *c* molecule

1. Overall conformation

The yeast iso-1-cytochrome *c* molecule has the typical cytochrome *c* fold, with the polypeptide chain organized into a series of α-helices and fairly extended loop structures. The protein folding serves to envelop the heme prosthetic group within a hydrophobic pocket formed by a shell of polypeptide chain one to two residues thick. The polypeptide chain backbone, heme group and heme ligands of yeast iso-1-cytochrome *c* are shown in Figure III.5.

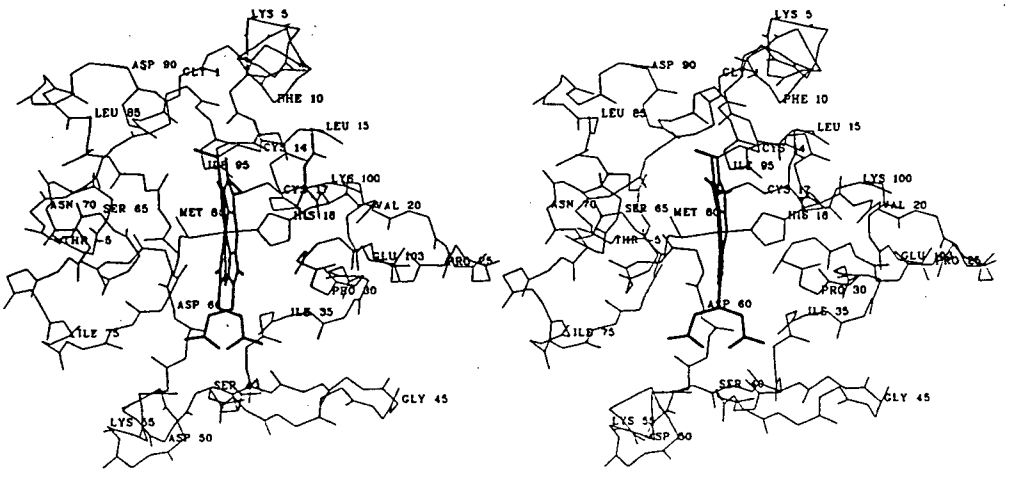

Figure III.5. Stereo diagram of the polypeptide chain backbone and heme group of yeast iso-1-cytochrome *c*. Main chain atoms are represented as thin lines and the heme group as thick lines. The heme thioether linkages formed by the side chains of Cys14 and Cys17, and the heme ligands formed by the side chains of His18 and Met80 are also shown. The α-carbons of every 5th residue, in addition to those of the N- and C-terminal residues, are labelled with their three-letter amino acid names and sequence numbers. The molecule is shown here in its standard orientation (see legend to Figure I.2). The heme group is seen edge-on, and that portion of the molecule containing the exposed heme edge directly faces the viewer.

2. Main chain torsion angles

A plot of all main chain torsion angles (ϕ and ψ) for iso-1-cytochrome *c* is shown in Figure III.6. Evident in this plot is the preponderence of residues with the α-helical conformation ($-60°$, $-40°$), and of glycines that belong to type II β-turns ($85°$, $5°$). All non-glycine residues fall within or near allowed regions, except for Lys27 ($-121°$, $-127°$). Lys27 is the first residue in a tight turn that has the conformation of a distorted reverse γ-turn (Milner-White *et al.*, 1988; Nemethy and Printz, 1972; Bandekar and Krimm, 1985), in which the amide group of residue *i*

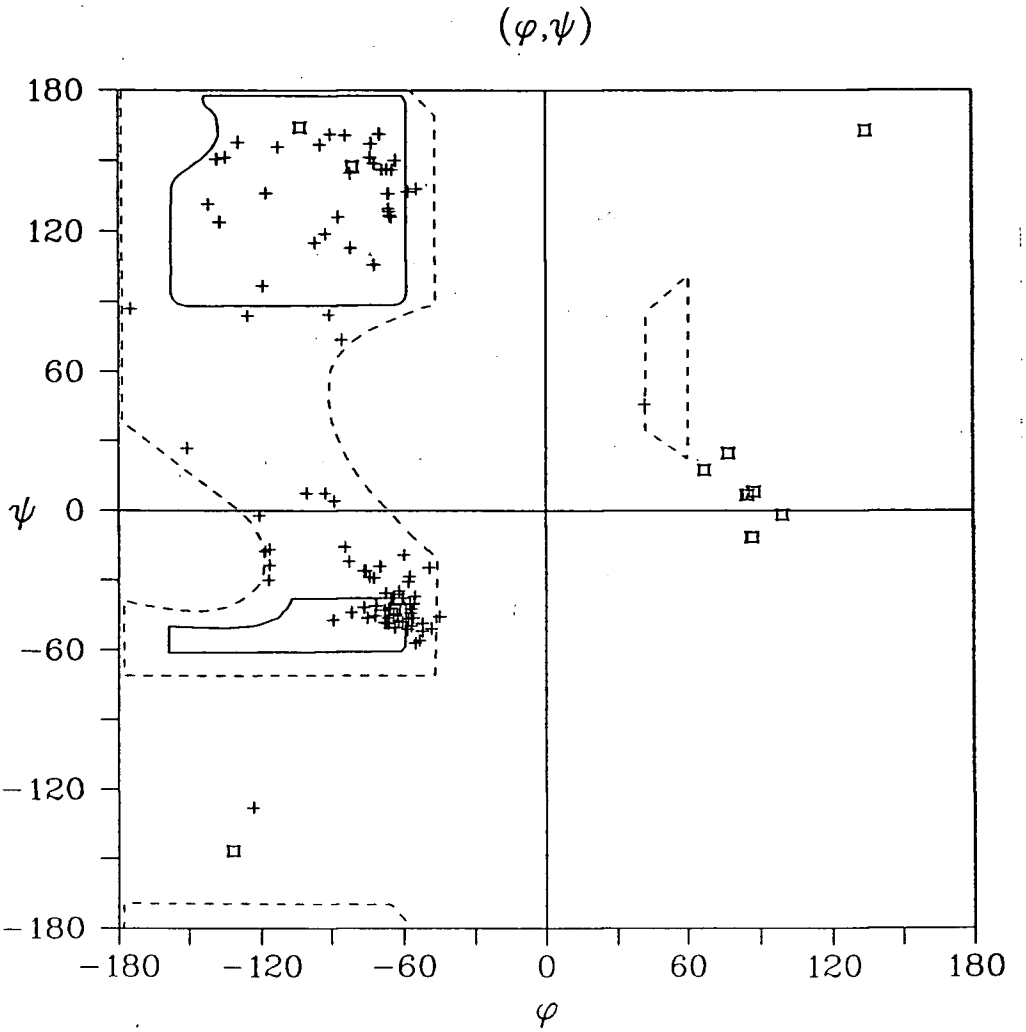

Figure III.6. Ramachandran plot for wild-type yeast iso-1-cytochrome *c*. The main chain ϕ and ψ dihedral torsion angles of glycine residues are denoted by the symbol □; those of all other residues are denoted by a +. Main chain conformations for an alanyl residue (having an N-CA-C angle, τ, of $112°$) in poly(L-alanine) that are fully allowed, and allowed assuming shortened permissible minimum interatomic contact distances (Ramakrishnan and Ramachandran, 1965) are enclosed by solid and dashed lines, respectively.

(Lys27) hydrogen bonds the carbonyl group of residue $i+2$ (Gly29) (see Figure III.7). The additional hydrogen bonds present, Gly29 N to Cys17 O and Lys27 NZ to Leu15 O, may serve to stabilize this strained conformation. Of note is the lowered stability of variant iso-1-proteins which possess side chains (Leu, Gln, Trp and Tyr) at position 27 (Hickey *et al.*, 1988) which are unable to form the latter interaction. The conformation observed for Asn56 (42°, 47°) is unusual in that both ϕ and ψ have values greater than 0°, but is not uncommon for asparagine residues (Matthews, 1977).

3. Secondary structure in yeast iso-1-cytochrome *c*

The secondary structural elements, assigned on the basis of the observed hydrogen-bonding patterns and main chain ϕ and ψ angles, present in yeast iso-1-cytochrome *c* are detailed in Table III.6. The average main chain torsion angles for all residues having an α-helical conformation are in good agreement with those observed in other highly refined protein structures (Sheriff *et al.*, 1987; McGregor *et al.*, 1987). However, as described below some distortions from ideal α-helical geometry do occur.

The N-terminal α-helix contains a kink at Ala7 and Thr8, with both the Ala7 O - Lys11 N and Thr8 N - Lys4 O hydrogen bonds being long. This distortion accommodates hydrogen

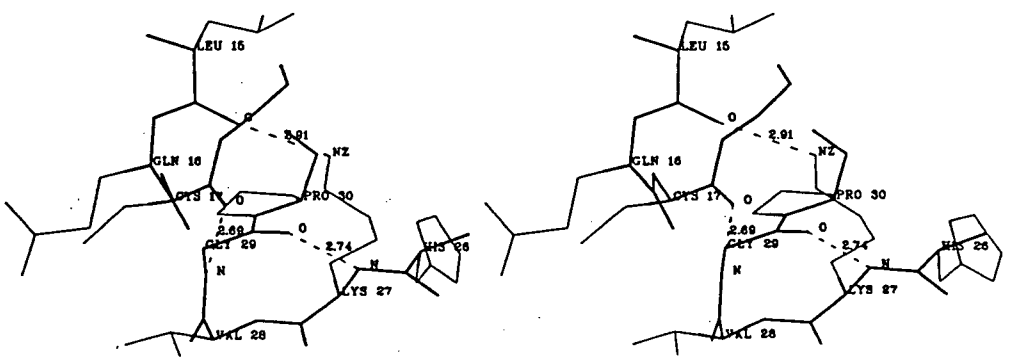

Figure III.7. Stereo diagram of the distorted γ-turn at residues 27 to 29 in yeast iso-1-cytochrome *c*. The polypeptide main chain is shown in thick lines, and side chains in thin lines. Hydrogen bonds occurring between Lys27 N and Gly29 O, Cys17 O and Gly29 N, and Lys27 NZ and Leu15 O are shown as broken lines, and their corresponding hydrogen bond lengths (in Å) are indicated.

Table III.6. Secondary structural elements present in yeast iso-1-cytochrome *c*

Secondary structural element	Residues involved	Main chain torsion angles[a] (°)
α-helix	3 - 12	(-64.7, -40.8)
β-turn (type II)	21 - 24	(-65, 128), (84, 8)
distorted γ-turn	27 - 29	(-123, -128), (-57, -49), (-103, 165)
β-turn (type II)	32 - 35	(-55, 138), (66, 19)
β-turn (type II)	35 - 38	(-64, 126), (86, -11)
β-turn (type II)[b]	43 - 46	(-82, 147), (87, 9)
α-helix	50 - 55	(-65.5, -37.8)
α-helix	61 - 69	(-63.6, -38.6)
α-helix	71 - 74	(-66.3, -40.0)
β-turn (type II)	75 - 78	(-66, 128), (99, 0)
α-helix	88 - 101	(-63.5, -42.2)
Overall:		
α-helix	43 residues	(-64.3, -40.3)
β-turn (type II)	5 turns	(-66.4, 133.4), (84.4, 5.0)

[a]The main chain torsion angles listed refer to the average (ϕ, ψ) angles for residues in each α-helical segment, to (ϕ, ψ)$_2$ and (ϕ, ψ)$_3$ for residues in each type II β-turn, and to (ϕ, ψ)$_1$, (ϕ, ψ)$_2$, and (ϕ, ψ)$_3$ for residues in the distorted γ-turn. Overall averages for the former two classes are also listed.

[b]Mediated through a water molecule.

bonding by the side chains of two nearby threonine residues to the carbonyl group of residues in the preceding turn of the helix (Thr8 OG1 - Lys5 O, and Thr12 OG1 - Thr8 O). The C-terminal end of this helix also possesses an unusual distortion in which the carbonyl groups of residues 12, 13 and 14 are directed markedly away from the helix axis, and the carbonyl group of residue 10 accepts a hydrogen bond from the amide groups of both residues 14 and 15. This distortion appears to occur in order to accommodate the formation of a thioether bond from Cys14 to the heme group.

The helix encompassing residues 61 to 69 is distorted towards a 3_{10} helix at both its N- and C- terminal ends. The carbonyl group of the first residue of this helix is involved in

hydrogen bonding with the $i+3$ and $i+4$ main chain amides (Asp60 O to Asn63 N and Met64 N). The last turn of this helix also assumes a 3_{10}-helix hydrogen bonding pattern (Thr69 N hydrogen bonds Glu66 O, and Asn70 N hydrogen bonds Tyr67 O). This latter distortion may be caused by Thr69, whose OG1 atom forms a hydrogen bond with Ser65 O. Also, in the middle of this helix, Glu66 N and Asn62 O do not form the expected hydrogen bond, whereas Asn63 O hydrogen bonds both Glu66 N and Tyr67 N.

A threonine residue also appears to cause a distortion in the C-terminal α-helix (residues 88 to 101), which is kinked at Thr96 to allow the side chain OG1 atom of this residue to hydrogen bond the carbonyl group of residue 92. This results in the amide nitrogens of both residues 96 and 97 being involved in bifurcated hydrogen bonds with the carbonyl groups of residues 92 and 93, and with those of residues 93 and 94 respectively. Beginning at residue 99, the C-terminal helix adopts main chain torsion angles characteristic of a 3_{10} helix. In this αC1 distortion (Baker and Hubbard, 1984), Lys100 N hydrogen bonds the carbonyl group of residue 97, and Ala101 N forms a bifurcated hydrogen bond to the carbonyl groups of residues 97 and 98. This allows the side chain of Lys99 to hydrogen bond to the carbonyl group of residue 96, and also facilitates the hydrogen-bonded interaction of the two most C-terminal residues with the rest of the molecule (Cys102 O hydrogen bonds Gly34 N, and the C-terminal carboxylate group of residue 103 forms a salt bridge to His33 NE2).

The other two helices present in yeast iso-1-cytochrome *c* (residues 50 to 55, and 71 to 74; see Table III.6) are short and thus contain only one or two main chain hydrogen bonds of the type that characterize an α-helix. Overall, the α-helices in iso-1-cytochrome *c* are typical of those found in other high resolution protein structures, in that their conformations do not strictly adhere to that of the idealized α-helix (Arnott and Dover, 1967). Particularly notable in the present study is that the occurrence of threonines within a helix causes substantial distortion of regular α-helical geometry, and in two of the helices, marks the end of these secondary structures. The propensity for threonine to cause distortions in α-helices has also been observed in other protein structures (McGregor *et al.*, 1987; Poulos *et al.*, 1987a).

Aside from the helical segments described, little other regular secondary structure occurs in the iso-1-cytochrome *c* molecule. Residues 37 to 40 and 57 to 59 form a very short two-stranded anti-parallel β-sheet (Figure III.8). This secondary structure element contains only three inter-strand main chain hydrogen bonds. There are five turns of the β type II class, which possess average main chain torsion angles typical of this class of secondary structure (see Table III.6; Richardson, 1981). The β-turn occurring at residues 43 to 46 is distorted, in that Ala43 O forms a hydrogen bond with a water molecule instead of with Tyr46 N as would be expected in a regular type II β-turn. This distortion is likely due to local interactions formed in the final folded state of the protein (Richardson, 1981), as it facilitates the hydrogen bonding of Glu44 O to His26 NE2, as well as the packing of the side chain of Tyr46 against the pyrrolidine ring of Pro30. The remaining regions of the polypeptide chain form irregular loop structures.

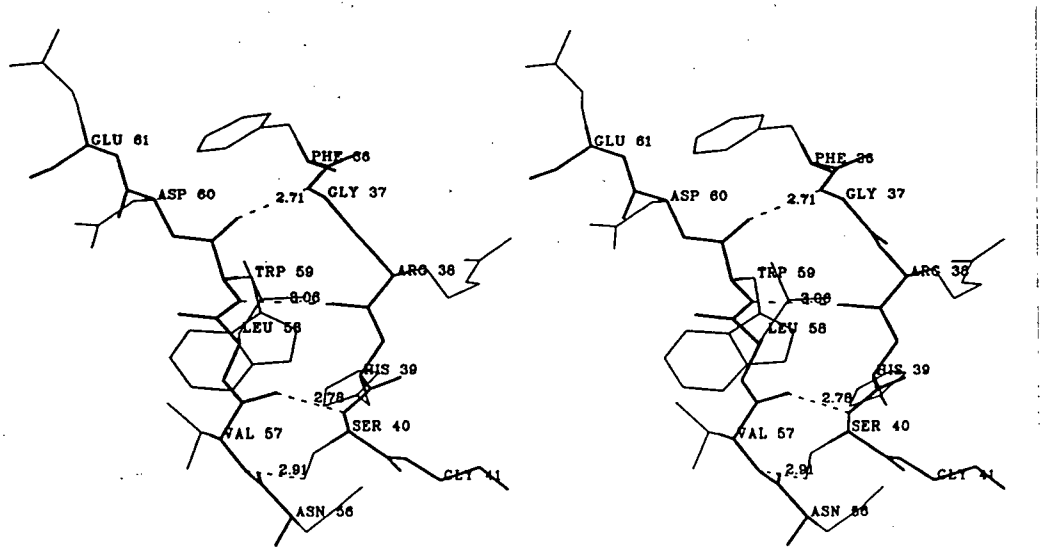

Figure III.8. Stereo diagram of the short segment of antiparallel β-sheet in yeast iso-1-cytochrome *c*. The two strands are formed by residues 37 to 40, and 57 to 59. The polypeptide main chain is shown as thick lines, and side chains as thin lines. The four interstrand hydrogen bonds that occur (Gly37 N - Trp59 O, Arg38 O - Trp59 N, Ser40 N - Val57 O, and Val57 N - Ser40 OG) are shown as thin, dashed lines and the bond distances (in Å) involved have been indicated.

4. Conformation of side chains

The positioning of all the side chains of yeast iso-1-cytochrome *c* is shown in Figure III.9. Figure III.10 shows the distribution of values for side chain dihedral torsion angles for those cases where a staggered conformation is expected (i.e. at sp^3 carbons; χ values of $\pm 60^\circ$, 180°). The relatively low value of 19.2° for the overall r.m.s. deviation of these torsion angles from their optimal staggered conformations is typical of well-refined protein structures. Yeast iso-1-cytochrome *c* has 16 lysine residues, some of which have poorly defined electron density for portions of their side chains. Excluding 29 torsion angle values (from 13 lysine residues) judged to be uncertain because they are determined by the positions of atoms beyond the last well-defined side chain atom (usually CD or CE), the r.m.s. deviation falls to 16.5°.

Several analyses of the preferred values for side chain dihedral angles in proteins have been recently published (Ponder and Richards, 1987; Summers *et al.*, 1987; McGregor *et al.*, 1987). The side chain conformations observed in iso-1-cytochrome *c* show similar trends to those noted in these studies. As Figure III.10 shows, the g^- conformation ($+60^\circ$) at $\chi 1$ is strongly avoided,

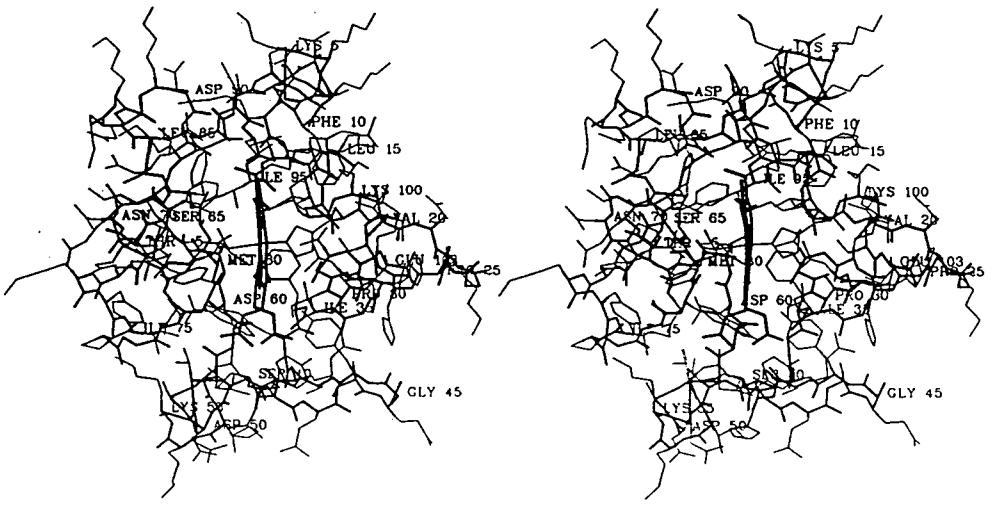

Figure III.9. Stereo diagram illustrating the positioning of side chains in yeast iso-1-cytochrome *c*. Drawn with thin lines are side chain groups; drawn with thick lines are the polypeptide chain backbone and also the heme group. Every 5th α-carbon has been labelled with its three-letter amino acid designation and sequence number. The orientation of view is the same as that in Figure III.5.

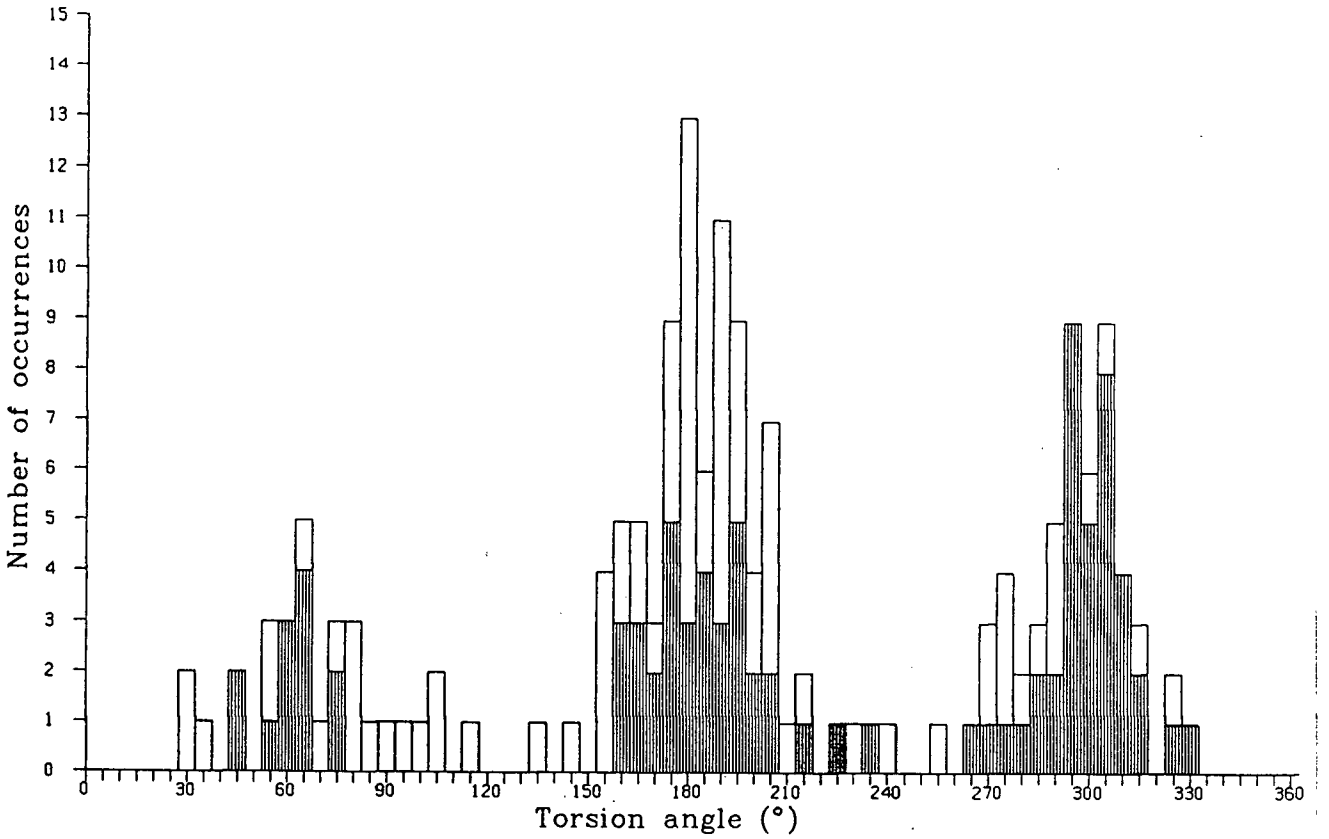

Figure III.10. Distribution of side chain dihedral torsion angles in yeast iso-1-cytochrome *c*. Only those angles for which a staggered conformation is expected to be most favored have been included (168 angles in total). The distribution of only the χ1 torsion angles, excluding those belonging to proline residues, (85 angles in total) is shown by the shaded bars.

accounting for only 12 of 85 χ1 angles present. For residues within *α*-helices, this avoidance is absolute except for Thr8 and Thr12 in the N-terminal *α*-helix, whose conformations are influenced by side chain hydrogen bonding interactions as described above. In general, threonine and serine are residues for which the g^- conformation is not disfavored with respect to g^+ and t (McGregor *et al.*, 1987). Six of the 12 side chains in iso-1-cytochrome *c* that have the g^- conformation at χ1 are threonines or serines. The remaining six side chains are all from glutamates and lysines. For these residues, χ2 usually strongly favors the t conformation if χ1 is g^- (McGregor *et al.*, 1987; Ponder and Richards, 1987), and this is the case in four of the six side chains in iso-1-cytochrome *c*. The χ1 torsion angles of aromatic residues are clustered closely about preferred values: 180° (r.m.s. $\Delta = 10°$) for all but 2 phenylalanines and tyrosines; and −60° (r.m.s. $\Delta = 4.6°$) for 3 of 4 histidines and the single tryptophan present. In 5 out of

7 occurrences, isoleucines and valines have the most preferred conformation with the lone hydrogen atom of the CB carbon being gauche to both the main chain N and C atoms. Of the sulfur-containing side chains, all three cysteines are g^+; and the two methionines, Met64 in an α-helix, and Met80 which forms a heme ligand, are tg^- and ttt, respectively. Overall, the conformations of the majority of the side chains in iso-1-cytochrome *c* are consistent with established torsional angle preferrences, and are in agreement with side chain rotamer templates compiled in other studies (Ponder and Richards, 1987; McGregor *et al.*, 1987).

5. Side chains with multiple conformations

Five side chains of yeast iso-1-cytochrome *c* are indicated by their appearance in electron density maps to exist in multiple conformations. The residues involved, and their alternative conformations and interactions are detailed in Table III.7. In the current structural determination, these side chains have been refined in only the conformation judged to have the greatest occupancy.

Ser47, Glu88 and Glu103 are on the surface of the protein molecule. The alternative conformations adopted by both Ser47 and Glu103 are fixed by interactions made to water molecules which are (more strongly) bound to neighboring protein molecules in the crystal lattice. Glu88 alternatively makes intramolecular interactions, through bridging water molecules, with one of two basic side chains: with Lys4 NZ via Wat221, and with Arg91 NH1 via Wat214.

Leu9 is partially buried, and in one of its conformations makes van der Waals contacts with another partially buried leucine (Leu85). The electron density belonging to Leu85 is suggestive that this side chain also undergoes dynamic movement, which may be associated with the alternative conformations adopted by Leu9. Leu58 is in a flexible region of the molecule and can interact with an adjacent molecule in the crystal lattice.

6. Poorly ordered regions of the iso-1-cytochrome *c* molecule

Presumably due to their high mobility, several regions of the iso-1-cytochrome *c* structure are poorly defined in electron density maps (see Table III.8). All such regions are on the surface

Table III.7. Side chains with multiple conformations in yeast iso-1-cytochrome *c*

Residue	Atoms affected	Conformation/ Estimated occupancy[a]	Torsion angles[b] (°)	Interactions made[c]
Leu9	CG, CD1, CD2	r / 0.7	-141, 67	CD2 contacts Ile95 O
		a / 0.3	-111, 123	CD2 contacts Leu85 CD1
Ser47	OG	r / 0.8	176	OG Hbond to Sul O2, Wat112
		a / 0.2	-57	OG Hbond to Sul O1, Wat210
Leu58	CG, CD1, CD2	r / 0.5	-169, -78	CD2 contacts Gly37 O
		a / 0.5	-177, 93	CD1 contacts His#39 ring; CD2 contacts Asp#60 CB
Glu88	CG, CD, OE1, OE2	r / 0.6	-172, -171, -21	OE2 Hbonds Wat221
		a / 0.4	-166, 167, 57	OE2 Hbonds Wat142 and Wat214
Glu103	CG, CD, OE1, OE2	r / 0.5	-163, 53, -133	CG contacts Lys#86 CB; OE2 Hbonds Lys#87 NZ
		a / 0.5	-95,-58,-57	OE2 Hbonds Wat131 and Wat182

[a]The 'r' designation refers to the conformation in which the side chain was refined and the 'a' to the alternate conformation as manually positioned in electron density maps. The occupancies of the two conformations have been estimated from the magnitudes of the respective electron density peaks.

[b]The torsion angles refer to $\chi 1$, $\chi 2$ and $\chi 3$, as applicable, for each conformation of the side chain.

[c]Where a '#' precedes a residue number, that residue belongs to a symmetry-related molecule in the crystal lattice. Note that 'contacts' refers to a van der Waal contact and that hydrogen bond has been abbreviated as Hbond.

of the molecule, have a high degree of solvent accessibility, and form no strong interactions with other protein groups. The accuracy of the coordinates of the affected atoms should be considered correspondingly low.

7. Heme conformation and interactions

The protoporphyrin IX heme prosthetic group of yeast iso-1-cytochrome *c* is buried within a hydrophobic pocket formed by the polypeptide chain. Only five atoms of the heme (CMC, CBC, CAC, CHD and CMD), which are positioned at the front edge of this group, are exposed to external solvent. The solvent-accessible area of the heme group is 48.8 $Å^2$, which represents just

Table III.8. Disordered regions of the yeast iso-1-cytochrome *c* structure

Residue	Appearance in electron density maps
Thr-5 to Glu-4	Main and side chain atoms: weak and fragmented
Phe-3	Main chain: weak
Lys4	CE onward: noisy
Arg13	CD onward: weak
Glu21	CG, CD: very weak.
Lys22	NZ: weak
Glu44	CG onward: very noisy and diffuse
Lys54 to Lys55	Main chain: weak. Lys54 side chain weak.
Asp60	Main chain: weak and fragmented
Glu61	CG onward: noisy and diffuse.
Glu66	Side chain carboxyl: weak
Leu85	CG, CD1, CD2: fairly strong, but density spread out
Lys86	CG onward: very weak
Lys100	CG onward: fragmented and diffuse

9.6% of the total surface area of this group (508 Å^2) and less than 1% of the total surface area of the cytochrome *c* molecule (4930 Å^2). The heme group is covalently attached to the polypeptide chain via two thioether bonds to the side chains of Cys14 and Cys17. The bond geometry of these thioether linkages is typical for carbon-sulfur single bonds, with the average bond distance being 1.79 Å, and the average CB-SG-C_{heme} angle being 101.2°.

The heme group itself is not absolutely planar, but is distorted into a saddle shape with the front and rear edges of the heme deviating toward the right, and the top and bottom edges toward the left of the molecule. Deviations of the planes of individual pyrrole rings with respect to the plane of the porphyrin ring as a whole, and with respect to the plane of the Fe-coordination pyrrole nitrogen plane, are detailed in Table III.9. All four rings show similar angular deviations with respect to the Fe-coordination sphere (i.e. the plane defined by the pyrrole nitrogen atoms), although the distortion is not one of a doming or ruffling of the porphyrin ring

Table III.9. Heme conformation, interactions and ligand geometry in wild-type yeast iso-1-cytochrome *c*

1. 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[a]

	Pyrrole N plane	Porphyrin ring plane
a. Pyrrole ring:		
A	9.4°	6.7°
B	11.1°	11.9°
C	8.8°	9.8°
D	8.2°	6.0°
b. Heme coordinate bond:		
Fe - His18 NE2 bond	2.1°	3.3°
Fe - Met80 SD bond	4.9°	7.5°

2. Heme iron coordination bond distances (Å)

His18 NE2	1.99
Met80 SD	2.35
Hem NA	1.97
Hem NB	2.00
Hem NC	1.99
Hem ND	2.00

3. Hydrogen bond interactions formed by heme propionic acid groups

Heme atom	Hydrogen bond partners and distances
O1A	Tyr48OH (2.83 Å), Wat121 (2.81 Å), Wat168 (2.85 Å)
O2A	Gly41N (3.21 Å), Asn52ND2 (3.34 Å), Trp59NE2 (3.09 Å)
O1D	Thr49OG1 (2.64 Å), Thr78OG1 (2.90 Å), Lys79N (3.17 Å)
O2D	Thr49N (2.94 Å)

[a]Each pyrrole ring plane 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 N plane is defined by only the 4 pyrrole nitrogens. These two latter planes differ in orientation by an angle of 2.6°.

(Scheidt, 1979). Pyrrole rings B and C are both involved in thioether linkages to the polypeptide chain, and show the greatest deviations with respect to the porphyrin ring plane as a whole.

These results suggest that the covalent bonds attaching the heme to the polypeptide chain are important factors in establishing the observed heme conformation. This is consistent with the atoms of pyrrole rings B and C having lower temperature factors (3.7 and 3.8 $Å^2$ on average, respectively) than atoms of pyrrole rings A and D (5.6 and 4.8 $Å^2$ on average, respectively).

The heme iron coordinate bond distances are also listed in Table III.9 and are typical of those found in small-molecule six-coordinate, iron-porphyrin complexes (Scheidt, 1979). The out-of-plane coordinate bonds involving the His18 and Met80 ligands show only slight deviations from being perpendicular to the pyrrole nitrogen plane. The heme iron atom is found to be displaced very slightly (by 0.03 Å) from this plane, in a direction toward the Met80 SD ligand atom. The imidazole ring of His18 lies almost perpendicular (99°) to the pyrrole nitrogen plane of the heme. Its orientation, in which a 46.5° angle is made between the imidazole ring and a vector joining the NA and NC pyrrole nitrogen atoms, is such that the CD2 and CE1 atoms make minimal contacts with the heme pyrrole nitrogens. As predicted by nuclear magnetic resonance and circular dichroism spectroscopies (Senn *et al.*, 1983), the ligation of the Met80 side chain to the heme iron has R chirality.

The two heme propionic acid groups are buried within the protein matrix, where they form a number of hydrogen bonds to internal polar groups of the protein (Figure III.11 and Table III.9). Thus while internally positioned, these propionic acid groups occupy a local environment which is clearly hydrophilic in character. In addition to the hydrogen bonds listed in Table III.9, an ionic interaction mediated by two bridging water molecules is formed between the carboxylate atom O1A of the heme propionate and the guanidinium group of Arg38.

C. Hydrogen bonding

The positions of hydrogen atoms in the structure of iso-1-cytochrome *c* were refined assuming shortened (X-ray) bond distances between the hydrogen atoms and their respective parent heavy atoms. However, for the analysis of hydrogen bonding, hydrogen atoms belonging to nitrogen or oxygen atoms were re-introduced with regular bond lengths to their parent atoms. A summary of all intramolecular hydrogen bonding interactions occurring in the structure of yeast

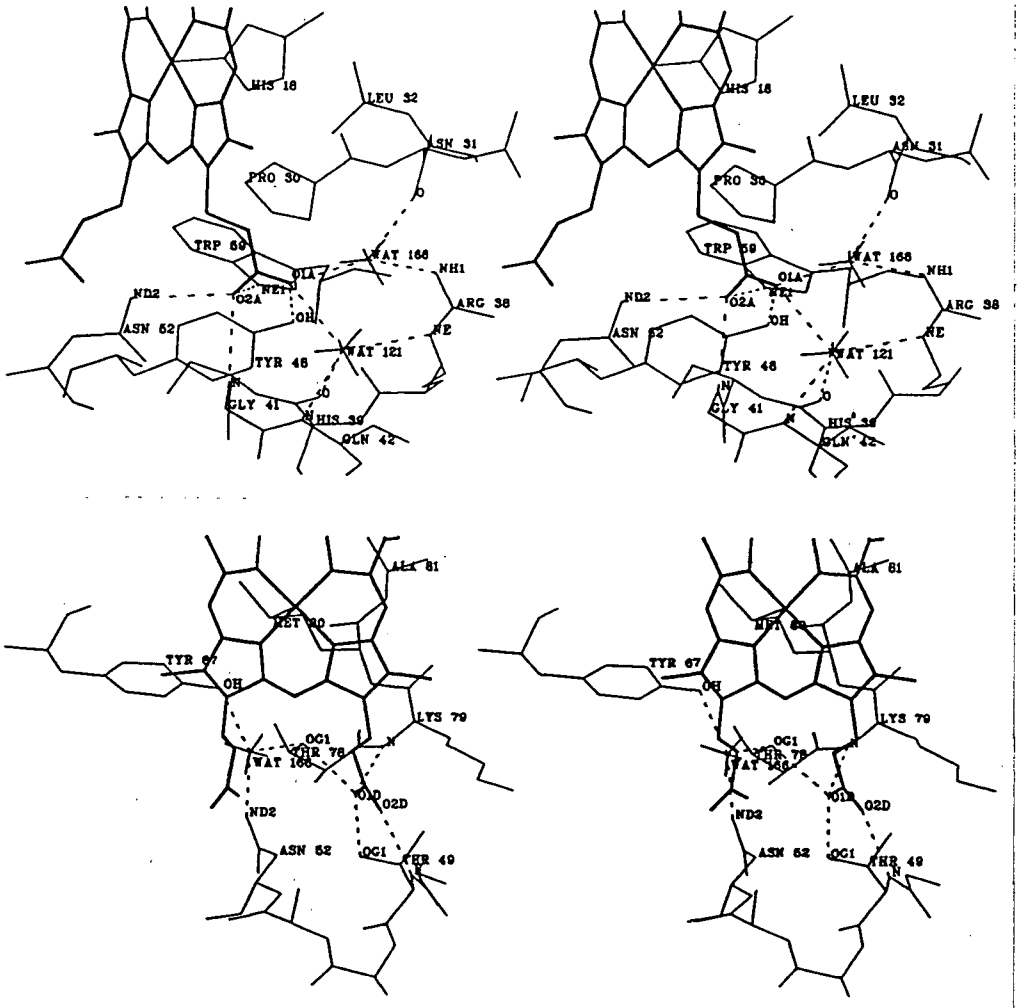

Figure III.11. Interactions formed by the heme propionic acid groups in wild-type yeast iso-1-cytochrome *c*. Shown are the local environment and hydrogen bonding interactions observed for the heme propionic acid group attached to (*top*) pyrrole ring A and to (*bottom*) pyrrole ring D. The heme and propionic acid groups are drawn with thick lines and the surrounding polypeptide chain residues with thinner lines. Hydrogen bonds are illustrated with thin dashed lines. Groups involved in forming hydrogen bonds to the heme propionyl carboxyl oxygen atoms, along with the hydrogen bond lengths observed, are detailed in Table III.9.

iso-1-cytochrome *c* is given in Table III.10, in which hydrogen bonds have been classified in terms of involvement of only main chain atoms, both main chain and side chain atoms, and only side chain atoms. The average geometrical parameters observed for these three classes of hydrogen bonds in the iso-1-cytochrome *c* structure are in excellent agreement with results obtained by Baker and Hubbard (1984) in their extensive analysis of hydrogen bonding in proteins.

Table III.10. Intramolecular hydrogen bonds[a] occurring in wild-type yeast iso-1-cytochrome *c*

1. Hydrogen bonds involving only main chain atoms

Gly6 N - Ser2 O	Ile35 N - Leu32 O	Ile75 N - Pro71 O
Ala7 N - Ala3 O	Gly37 N - Trp59 O	Thr78 N - Ile75 O
Thr8 N - Lys4 O	Arg38 N - Ile35 O	Phe82 N - Met80 O
Leu9 N - Lys5 O	Ser40 N - Val57 O	Leu85 N - Leu68 O
Phe10 N - Gly6 O	Ile53 N - Thr49 O	Arg91 N - Lys87 O
Lys11 N - Ala7 O	Lys54 N - Asp50 O	Asn92 N - Glu88 O
Thr12 N - Thr8 O	Lys55 N - Asn52 O	Asp93 N - Lys89 O
Arg13 N - Leu9 O	Trp59 N - Arg38 O	Leu94 N - Asp90 O
Cys14 N - Phe10 O	Met64 N - Asp60 O	Ile95 N - Arg91 O
Leu15 N - Phe10 O	Ser65 N - Glu61 O	Thr96 N - Asn92 O
Cys17 N - Cys14 O	Tyr67 N - Asn63 O	Tyr97 N - Asp93 O
His18 N - Cys14 O	Leu68 N - Met64 O	Leu98 N - Leu94 O
Gly24 N - Glu21 O	Thr69 N - Ser65 O	Lys99 N - Ile95 O
Lys27 N - Gly29 O	Thr69 N - Glu66 O	Lys100 N - Thr96 O
Gly29 N - Cys17 O	Asn70 N - Tyr67 O	Ala101 N - Tyr97 O
Leu32 N - Thr19 O	Tyr74 N - Asn70 O	Cys102 N - Leu98 O
Gly34 N - Cys102 O		

Averages: 2.12 (0.24) Å 2.99 (0.22) Å 155.0 (12.8)° 146.7 (16.1)° 142.0 (16.4)°

2. Hydrogen bonds involving main chain and side chain atoms

Phe-3 N - Thr-5 OG1	His33 N - Asn31 OD1	Asn63 N - Asp60 OD1
Gly1 N - Thr96 OG1	His33 ND1 - Val20 O	Thr69 OG1 - Ser65 O
Ser2 N - Asp93 OD1	His33 NE2 - Glu103 OXT	Lys73 N - Asn70 OD1
Lys5 N - Ser2 OG	Arg38 NH1 - His33 O	Lys79 N - Heml04 O1D
Thr8 OG1 - Lys5 O	Arg38 NH2 - His33 O	Lys79 NZ - Ser47 O
Thr12 OG1 - Thr8 O	Gly41 N - Heml04 O2A	Met80 N - Thr78 OG1
His18 ND1 - Pro30 O	Ala43 N - Tyr48 OH	Lys86 NZ - Thr69 O
Val20 N - Glu21 OE1	Tyr46 OH - Val28 O	Arg91 NH2 - Leu85 O
Glu21 N - Glu21 OE1	Thr49 N - Heml04 O2D	Arg91 NH2 - Lys86 O
His26 ND1 - Asn31 N	Asn52 N - Thr49 OG1	Thr96 OG1 - Asn92 O
His26 NE2 - Glu44 O	Lys55 NZ - Tyr74 O	Cys102 SG - Leu32 O
Lys27 NZ - Leu15 O	Val57 N - Ser40 OG	Cys102 SG - Leu98 O
Asn31 ND2 - Glu21 O	Asn62 N - Asp60 OD1	

Averages: 2.08 (0.27) Å 2.93 (0.25) Å 150.8 (15.7)° 133.6 (19.6)° 131.5 (20.0)°

3. Hydrogen bonds involving side chain atoms only

Thr-5 OG1 - Asn62 OD1	Asn52 ND2 - Heml04 O2A	Tyr67 OH - Met80 SD
Asn31 ND2 - Thr19 OG1	Trp59 NE1 - Heml04 O2A	Thr78 OG1 - Heml04 O1D
Tyr48 OH - Heml04 O1A	Asn63 ND2 - Asp60 OD2	Arg91 NE - Ser65 OG
Thr49 OG1 - Heml04 O1D		

Averages: 2.15 (0.21) Å 3.00 (0.26) Å 152.6 (16.4)° 115.6 (18.8)° 114.2 (21.2)°

Table III.10 (continued)

[a]For each potential hydrogen bond, the donor of the hydrogen atom is listed on the left, and the acceptor on the right.

The averages listed refer to H...A and D...A distances, and D–H...A, C–A...D and C–A...H angles respectively, where H denotes the hydrogen atom involved in the hydrogen bond, and A and D designate the acceptor and donor atoms. Standard deviations are given in parentheses.

Note that 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 $> 90°$.

A considerable portion of the hydrogen bonding observed in iso-1-cytochrome *c* occurs within secondary structures. Of the 49 hydrogen bonds in which both partners are main chain atoms, only 7 (14.3%) are not part of secondary structural elements. Also, 9 of 36 hydrogen bonds formed between main chain and side chain atoms are directly involved in secondary structures: four are made by intrahelical threonine side chains (discussed above), one is made by a serine in a β-sheet (Figure III.8), and four make capping interactions (Richardson and Richardson, 1988) at the N-termini of α-helices.

Of the 213 main chain groups that could potentially form hydrogen bonds, 96 make hydrogen bonds to other main chain atoms in the molecule. An additional 33 main chain groups make intramolecular hydrogen bonds to side chain atoms. In total, 61% of main chain atoms form intramolecular hydrogen bonds, and in only 23 of the 108 residues does a main chain atom not directly form such a hydrogen bond. In addition, 31 carbonyl oxygen atoms and 9 amide nitrogen atoms are each involved in two hydrogen bond interactions. Only 24 (11% of the total) main chain groups (10 amide nitrogens, and 14 carbonyl oxygens) do not form hydrogen bonds to other protein atoms or to solvent molecules.

The extent of hydrogen bonding involving side chain atoms is even greater. Yeast iso-1-cytochrome *c* possesses 62 polar side chains, which bear 91 atoms capable of participating in hydrogen bonding. Hydrogen bond interactions are not observed for only 2 side chains, and for only 6 polar atoms. The two amino acid residues found not to be involved in hydrogen bonding are both surface lysines (residues 4 and 100), whose side chains are directed into the solvent medium. Failure to observe hydrogen bonding in these cases is likely due to the substantial

mobility of these lysines (see Table III.8) and of their associated solvent molecules as well. Considering that side chain oxygen atoms are capable of participating in two (carboxyl and amide) or three (hydroxyl) hydrogen bonds each, and that many side chain nitrogen atoms can act as hydrogen bond donors to two (Gln NE2; AsnND2; and ArgNH1,2) or three (Lys NZ) acceptors, yeast iso-1-cytochrome *c* contains a total of 197 side chain hydrogen bonding functionalities. Of these, 131 (67%) are observed to participate in hydrogen bonding. Of the 119 hydrogen bonds formed by side chain atoms, 45 are intramolecular, with the remainder involving solvent molecules or other protein molecules in the crystal lattice.

Overall, 90.1% of all main chain and side chain polar groups form well-defined hydrogen bond interactions. The saturation of hydrogen bonding potential considering all possible hydrogen bonding functionalities is 69.4% (360 of 519). This high degree of hydrogen bonding present in iso-1-cytochrome *c* is typical of that found in other carefully refined protein structures (Baker and Hubbard, 1984), and undoubtedly contributes significantly to the stability of the folded state of this protein. The lack of fulfillment of hydrogen bonding potential that is present in the iso-1-cytochrome *c* structure is due in most cases (particularly for main chain atoms) to the fold of the protein rendering some potential sites inaccessible to favorable hydrogen bonding interactions.

D. Solvent structure

1. The complement of solvent molecules associated with iso-1-cytochrome *c*

The volume occupied by one molecule of iso-1-cytochrome *c* is estimated, by the method of Connolly (1983), to be 14610 $Å^3$. Therefore, ~8130 $Å^3$ of space (35.8% of the total) is left unoccupied in the asymmetric unit of the crystallographic unit cell. Assuming that this vacant space is occupied by water molecules, and that the density of the solvent medium is thus 1.0 g/cm^3, one would estimate that ~272 water molecules also occur in the asymmetric unit. During the structural refinement of yeast iso-1-cytochrome *c*, a total of 116 water molecules and 1 sulfate ion were identified. Of these, 6 water molecules occupy sites that are inaccessible to the probe sphere used to delimit the envelope enclosing the volume of the protein molecule (four of these six are buried in the protein interior, whereas two occupy depressions in the protein surface). The

remaining 110 waters and 1 sulfate ion represent 41% of the theoretical maximum calculated above, a percentage comparable to that found in other protein structures determined using crystals having a low solvent content (James and Sielecki, 1983; Blake *et al.*, 1983; Frey *et al.*, 1987; Morize *et al.*, 1987).

The unique complement of water molecules and sulfate anions associated with one molecule of iso-1-cytochrome *c* is illustrated in Figure III.12. The temperature factors of the water molecules range from 10.1 to 62.2 Å^2, with a mean of 37.1 Å^2. The majority of the water molecules populate the first solvation layer of the protein molecule: 93% are within 4 Å of the protein surface, and the overall average distance from each water molecule to the nearest protein atom is 3.1 Å.

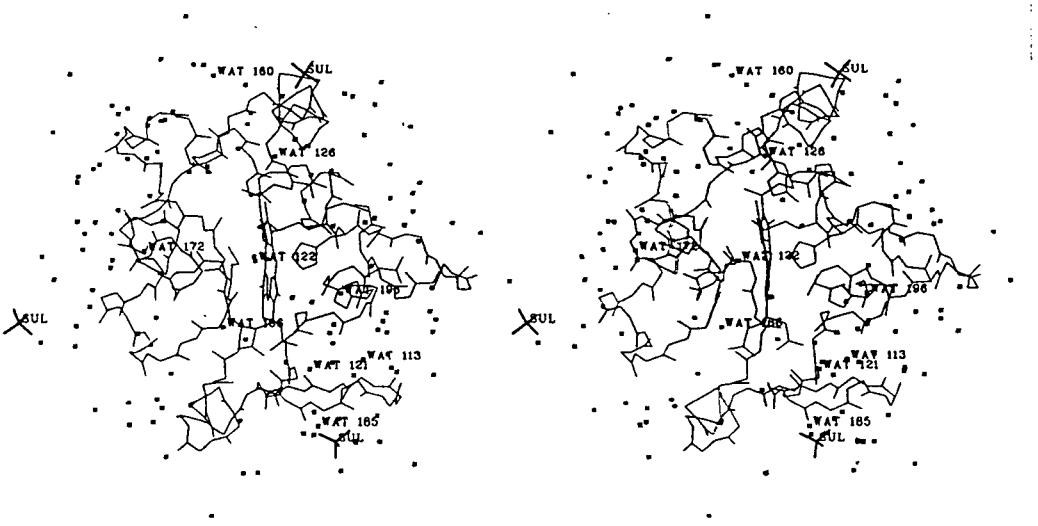

Figure III.12. Stereo diagram of the solvent molecules associated with a single molecule of wild-type iso-1-cytochrome *c*. Shown are 131 water molecules (crosses) and 3 sulfate anions (tetrahedrons). The asymmetric unit of the crystallographic unit cell contains only 116 water molecules and 1 sulfate ion, but an additional 15 water molecules and 2 sulfate ions that are involved in contacts with two (or more) symmetry-related protein molecules have also been included in this drawing. The sulfate anions (SUL), and the 9 conserved water molecules between yeast iso-1-, tuna and rice cytochromes *c* (see Table IV.6) have been labelled. The course of the polypeptide chain of the protein is drawn using only main chain atoms (thin lines).

2. Hydrogen bonding interactions

Solvation of polar protein groups plays an important role in stabilizing the folded conformation of a protein molecule (Edsall and McKenzie, 1983; Peters and Peters, 1985). As summarized in Table III.11, most of the characteristics of the hydrogen bonds made between protein atoms of iso-1-cytochrome *c* and the ordered solvent are typical of those found in other high resolution protein structures (Baker and Hubbard, 1984). In total, 265 hydrogen bonds are made by the 116 water molecules and one sulfate anion (2.3 per solvent molecule on average). Of the 116 water molecules, 76 (65.5%) interact directly with the protein molecule, through a total of 134 hydrogen bonds, or 1.76 hydrogen bonds each on average. Of the hydrogen bonds formed, 54% are to main chain atoms of the protein and 46% to side chain atoms. One anomaly in the pattern of hydrogen bonding interactions made by water molecules to yeast iso-1-cytochrome *c*

Table III.11. Summary of hydrogen bonds made by water molecules in the structure of wild-type iso-1-cytochrome *c*

Hydrogen bonding partner	Number of H bonds	Average distance[a] (Å)	Number of atoms involved[b]		Types of side chains involved
			Protein	Water	
Main chain N	25	3.07 (0.24)	25	21	--
Main chain O	46	2.88 (0.22)	37	31	--
Side chain N	29	2.89 (0.32)	24	26	8Lys, 3Arg, 2His 4Asn, 2Gln
Side chain O	34	2.84 (0.31)	25	28	2Thr, 3Ser, 2Tyr, 3Asp, 5Glu, 3Asn, 2Gln
Water molecule	71	2.76 (0.33)	--	81	--

[a]The standard deviation from the mean is given in parentheses.

[b]For each class of hydrogen bonding between water and protein (except for main chain N), the number of atoms of both water and protein involved will be smaller than the total number of hydrogen bonds. This is because some water molecules make hydrogen bonds to more than one protein atom of a given type, and some protein atoms of each type (except for main chain N) make hydrogen bonds to more than one water molecule.

relates to the preference for water molecules to interact with oxygen, as opposed to nitrogen, atoms of the protein. For the 13 protein structures considered by Baker and Hubbard (1984), the ratio of hydrogen bonds made to oxygen atoms versus to nitrogen atoms averaged 3:1 considering main chain atoms only, and 2.6:1 overall. The same ratios found for yeast iso-1-cytochrome *c* are 1.84:1 and 1.49:1, respectively. This can be accounted for partially by the large number of basic side chains present in iso-1-cytochrome *c* (the ratio of side chain O to N hydrogen bonding sites is only 1.1:1). However, this does not account for the observed increased participation of main chain amide nitrogens in acting as donors of hydrogen bonds to water molecules.

Previous structural studies of proteins have shown the associated water molecules are organized into networks (Teeter, 1984; Wright, 1987). If a water network is defined as a group of atoms, including at least two water molecules, that interact with each other through a continuous array of hydrogen bonds, most of the water molecules (73 of 116) associated with yeast iso-1-cytochrome *c* are observed to participate in forming 11 distinct networks (see Table III.12) that cover large sectors of the protein surface. On average, each of these networks is anchored to the protein surface by 5 hydrogen bonds. The largest network contains 18 water molecules and forms 14 hydrogen bonds to protein atoms. A further six small networks contain only two water molecules each and make zero to two hydrogen bonds to protein atoms. Overall, only six water molecules are neither directly hydrogen bonded to the protein, nor part of hydrogen bonding networks anchored to the protein molecule (4 of these make no contacts with any other atom). One instance of an ordering of water molecules into a pentagonal cluster (Teeter, 1984) is observed, and involves 4 water molecules and the OD1 atom of Asn31.

Of note is the fairly large number (35, or 30.2%) of water molecules found to make no hydrogen bonds to other water molecules. Disregarding the 4 that are completely internal, the remainder are of two main classes. The first form only one hydrogen bond to the protein molecule, and thus are only loosely held in position themselves and therefore unlikely to be able to order other solvent molecules around them. The second form several hydrogen bonds, and are involved in bridging polar protein groups. For these water molecules, the hydrogen bonding

Table III.12. Organization of water molecules into networks associated with yeast iso-1-cytochrome *c*.

Network[a]	Number of water molecules (and of hydrogen bonds made to protein)	Water molecules present[b]	Protein atoms forming hydrogen bonds
1	6 (4)	105 132 158 159 183 203	Arg13 NH1 Lys86 N Lys87 N Asp90 OD1
2	2 (1)	106 157	Gly37 O
3	18 (14)	107 113 115 123 126 128 143 156 165 167 181 186 188 189 198 210 212 220	Lys11 O Thr12 O (2) Leu15 N Gln16 N Gln16 OE1 Lys#27 O Gly#45 O (2) Ser#47 N Lys#79 NZ Gly83 N Gly84 N Gly84 O
4	6 (7)	108 133 174 191 194 211	Gln#16 NE2 Gly23 N Gly24 O Asn31 OD1 (2) Arg38 NH2 Glu44 N
5	5 (2)	109 130 141 187 200	Gln42 O (2)
6	4 (4)	111 131 163 182	Gly34 O Phe36 O (2) Arg38 N
7	3 (1)	114 116 148	Tyr97 OH
8	10 (10)	Sul 127 140 149 153 164 169 170 185 207	Gly41 O Ala43 O Tyr46 O (2) Tyr48 N Glu#66 O Asn#70 O Lys#73 O Lys#73 NZ (2)
9	3 (4)	119 195 218	Val20 N Glu21 OE1 Glu21 OE2 Lys27 NZ
10	12 (10)	120 150 161 173 177 178 179 202 205 208 209 214	Glu-4 O Glu-4 OE2 Glu-4 OE2 Ser65 OG (2) Glu66 OE1 Glu66 OE2 Arg91 NH1 (2) Asn92 OD1
11	2 (2)	124 #124	Trp59 O Trp#59 O
12	2 (0)	137 190	
13	3 (3)	138 139 219	Asp50 N Lys#99 NZ Asp50 OD1
14	2 (2)	146 193	Thr12 OG1 Lys#89 NZ
15	2 (0)	151 #151	
16	4 (2)	155 175 176 213	Gln42 NE2 Ala#51 N
17	2 (2)	217 #217	Asp90 OD2 Asp#90 OD2

[a]Water networks have been arbitrarily ordered according to the water molecule in each network having the lowest sequence number.

[b]Groups having numbers preceded by a '#' belong to an asymmetric unit other than the one containing the remainder of the water molecules and protein atoms present in the network.

potential is saturated by protein groups, or the protein itself occludes interaction with additional water molecules.

A trend clearly evident from Table III.13 is that the water molecules forming the greatest number of hydrogen bonds have the smallest mean-square positional displacements, as measured by observed atomic temperature factors. [It should also be noted that since occupancy and temperature factor are highly correlated (Edsall and McKenzie, 1983), these results must also be considered to be consistent with a greater occupancy at solvent sites at which there is the potential for the formation of a larger number of hydrogen bonds.] These data also show that the most tightly bound water molecules are those that interact directly with the protein molecule, and in general, this is particularly true when the interactions made are with main chain atoms. This is characteristic of the solvent structures of many other proteins (Blake *et al.*, 1983; Kamphuis *et al.*, 1984). The eight most tightly bound water molecules have an average B-factor of 13.3 $Å^2$ (a value lower than the overall average B-factor of protein atoms) and form an average of 3.6 hydrogen bonds each to the yeast iso-1-cytochrome *c* molecule. As illustrated in Figure III.13, the two water molecules observed to make 4 hydrogen bonds to main chain atoms have a

Table III.13. Temperature factors of water molecules bound to yeast iso-1-cytochrome *c* as a function of the number of hydrogen bonds made

Hydrogen bonds made specifically to:	Average temperature factor of water molecules making the indicated number of such hydrogen bonds (in $Å^2$)				
	0	1	2	3	4 or more
Any atoms (protein and solvent)	46.5 (4)	41.5 (29)	38.8 (37)	35.9 (27)	27.8 (19)
Any protein atom	44.8 (40)	36.1 (46)	35.3 (15)	27.0 (9)	16.4 (6)
Main chain atoms	41.2 (71)	33.5 (29)	31.2 (9)	23.8 (5)	12.9 (2)

Each entry in this table refers to the average temperature factor for the water molecules making the specified number of hydrogen bonds to atoms of the type given in the column at the left. The number of water molecules in each grouping is given in parentheses below the average temperature factor entry.

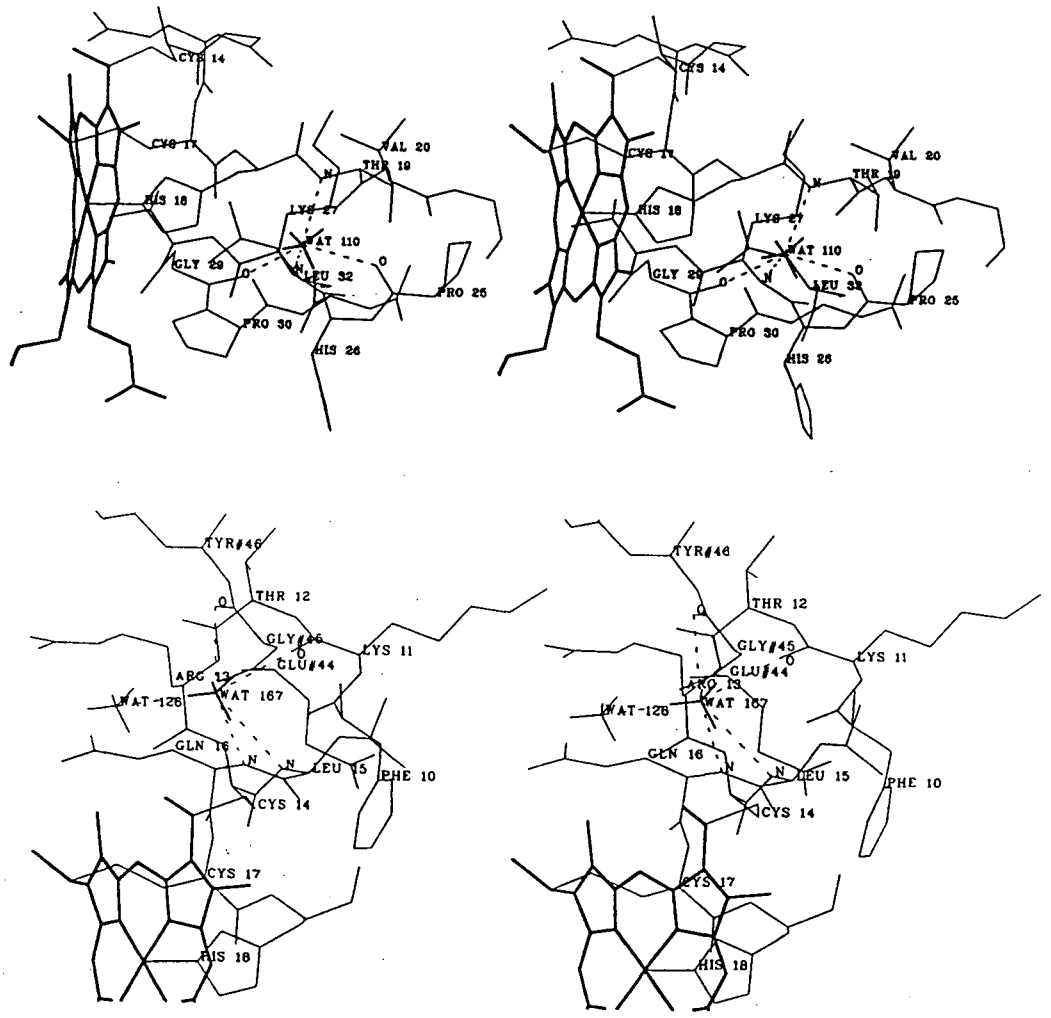

Figure III.13. Stereo diagrams of the local environments of the two water molecules forming four hydrogen bonds to main chain atoms of iso-1-cytochrome *c*. (*Top*) Wat110 is positioned within a loop of the polypeptide chain, and makes hydrogen bonds (dashed lines) to Thr19 N, Pro25 O, Lys27 N and Gly29 O. (*Bottom*) Wat167 sits in a small depression on the protein surface, and forms hydrogen bonds to Lys11 O, Leu15 N, Gln16 N, Gly#45 O and Wat128. (Residues belonging to symmetry-related molecules in the crystal lattice are labelled with a #.)

tetrahedral coordination sphere.

3. Internal water molecules

In total, there are 4 water molecules completely buried within the interior of yeast iso-1-cytochrome *c*. One of these, Wat110, is positioned adjacent to the His18 ligand and forms the hydrogen bonding interactions illustrated in Figure III.13a. The other three internal water molecules are in close proximity to the heme group. As shown in Figure III.11b, Wat166 is

hydrogen bonded to atoms belonging to three buried side chains (Asn52 ND2, Tyr67 OH, and Thr78 OG1). Wat121 and Wat168 together act as a bridge between the heme propionic acid oxygen O1A and the guanidinium group of Arg38 (see Figure III.11a). These water molecules also form hydrogen bonds to other internal main chain atoms (His39 O and Gln42 N with Wat121, and Asn31 O with Wat168).

4. Anion binding site

One well-ordered solvent molecule, which appeared in electron density maps as a peak having an unusually large volume and bearing tetrahedrally arranged lobes, and when refined as a water molecule had a temperature factor that fell to a value of 2.0 $Å^2$, was modelled as a sulfate ion. Although an ion undoubtedly lies at this position, its precise identity is somewhat uncertain. The crystallization solution contained both sulfate and phosphate ions, although the sulfate concentration was 40 times greater. At pH 6.2, a phosphate would be expected to exist as $H_2PO_4^-$ or HPO_4^{2-}, and thus would have different hydrogen bonding preferences from that of SO_4^{2-}. However, the nature of the hydrogen bonds made by the anion in this case does not allow distinguishing between sulfate and phosphate based on these considerations. During refinement, the behavior of the sulfate atoms was reasonable, with no significant deviations from restraint bond distances and with the sulfur and four oxygen atoms all having similar temperature factors ranging from 33.5 to 35.0 $Å^2$.

The electron density attributed to this sulfate ion and also the residues with which the sulfate ion forms hydrogen bond interactions are illustrated in Figure III.14. The sulfate is located at the N-terminal end of an α-helical segment of polypeptide chain (residues 3 to 12), a common site for anion binding (Hol *et al.*, 1985). Three of the oxygen atoms of the sulfate are oriented so as to individually hydrogen bond to Ser2 OG, Ala3 N, and Lys4 N. Interestingly, this system of hydrogen bonds is identical to those occurring at the N-termini of helices that have key roles in binding sulfate in sulfate-binding protein (Pflugrath and Quiocho, 1985) and in binding phosphate in flavodoxin (W.W. Smith *et al.*, 1977). The sulfate ion in yeast iso-1-cytochrome *c* is further held in place by hydrogen bonds from the side chains of Ser47 and Lys73, residues

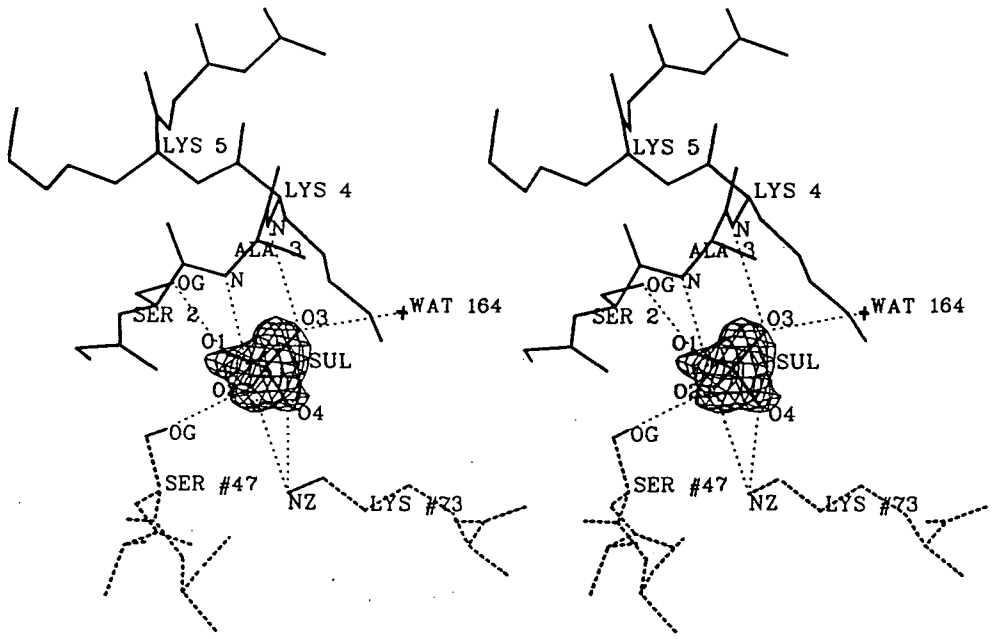

Figure III.14. Stereo diagram of the electron density and hydrogen bonding interactions of the sulfate ion bound to iso-1-cytochrome *c*. Electron density, from a map calculated with coefficients of the type $2F_o-F_c$, α_c, belonging to the sulfate ion is represented by the solid wire envelopes. The hydrogen bonds (thin dashed lines) formed by the sulfate ion are Sul O1 - Ser2 OG, Sul O2 - Ala3 N, Sul O3 - Lys4 N, Sul O3 - Wat164; and involving adjacent protein molecules, Sul O2 - Lys#73 NZ, Sul O2 - Ser#47 OG and Sul O3 - Lys#73 NZ. Amino acid residues originating from adjacent molecules in the crystal lattice have been drawn using thick dashed lines, and have residue-number labels preceded by a '#'.

which are from adjacent protein molecules in the crystal lattice.

Cytochrome *c* has been shown to possess an anion binding site, which when occupied affects the rate at which cytochrome *c* carries out electron transfer reactions with both small molecule redox reagents (e.g. dithionite and ferricyanide) and cytochrome oxidase (reviewed in Ferguson-Miller *et al.*, 1979; Pettigrew and Moore, 1987). Considering that it is chemical modification of those lysines located at the front of the cytochrome *c* molecule that has the greatest effect on these reactions, electron transfer is thought to occur via the front, exposed edge of the heme (Ferguson-Miller *et al.*, 1979). Because the site at which the sulfate ion is most strongly bound in yeast iso-1-cytochrome *c* occurs near the rear of the molecule (Figure III.12), it would seem unlikely that this site would directly influence these electron transfer reactions.

E. Polypeptide chain mobility and internal cavities

The average B-factors of main chain and side chain atoms of each residue in iso-1-cytochrome *c* are shown in Figure III.15. The overall average B for all non-hydrogen atoms is 16.5 Å^2. Refined crystallographic temperature factors can be used as a measure of the relative flexibility of different regions of a protein molecule (Ringe and Petsko, 1985). For yeast iso-1-cytochrome *c*, they clearly demonstrate the dominating influence of the heme group on

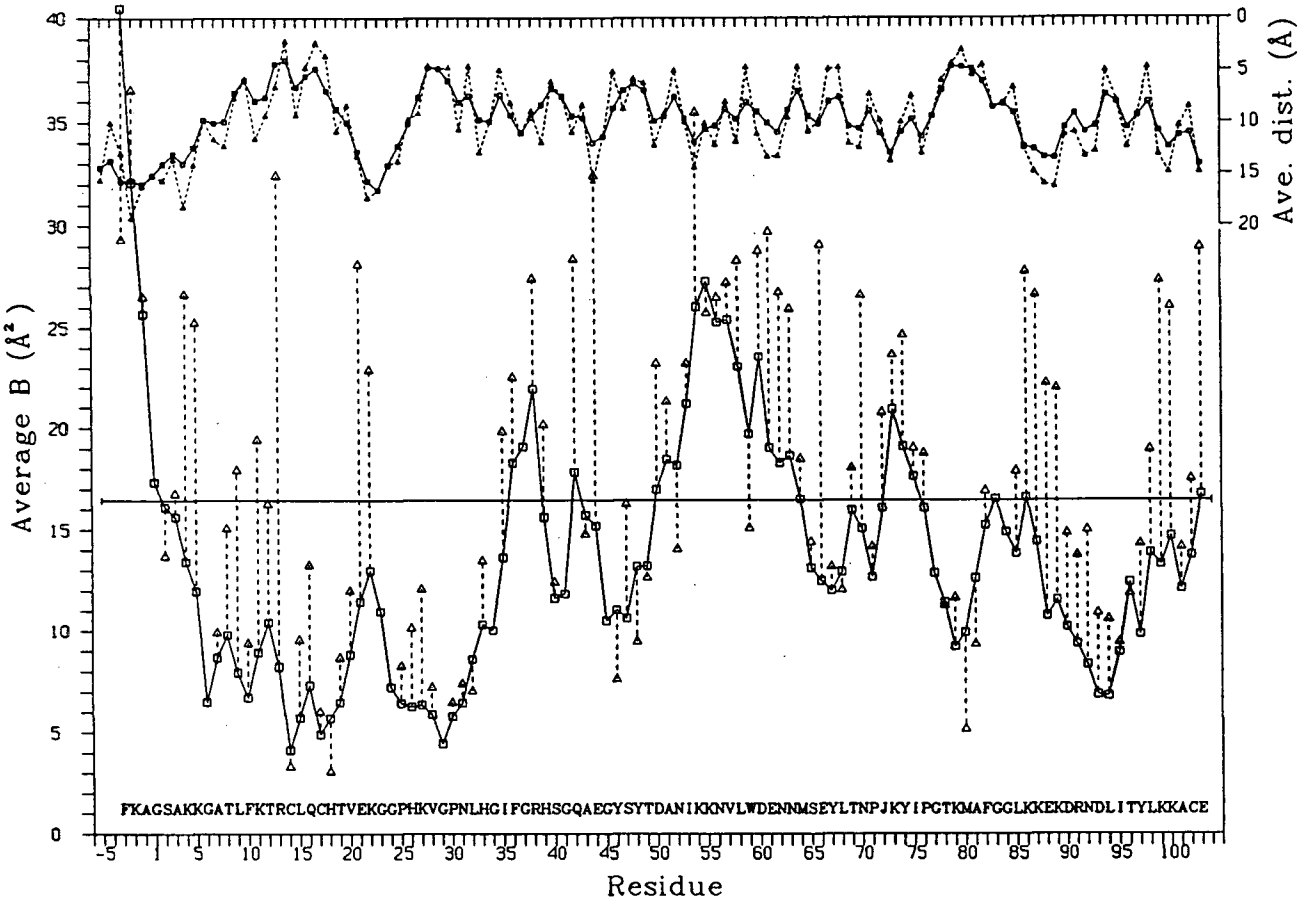

Figure III.15. Average temperature factors of the main chain and side chain atoms of each amino acid residue in wild-type yeast iso-1-cytochrome *c*. Values for the polypeptide chain backbone are represented as (¤——¤), and for side chain groups as (¤--Δ). The amino acid sequence of the protein is shown using the one-letter code above the axis indicating residue numbering. Note that the average temperature factors for the two N-terminal residues [Thr(-5): 53.6 Å^2 for main chain, and 54.7 Å^2 for side chain atoms; and Glu(-4) 47.9 Å^2 and 50.8 Å^2] are not shown. A further horizontal line defines the average B value (16.5 Å^2) of all protein atoms. The average closest distance to the porphyrin ring (for definition, see legend to Table III.9) of atoms belonging to the main chain (¤——¤) and side chain (Δ--Δ) of each residue is shown at the top of the Figure, according to the axis scale at the top right.

internal dynamics. The heme itself is tightly held within the molecule, having an average B-factor of 5.3 $Å^2$ for the 43 non-hydrogen atoms. The regions of the polypeptide chain backbone observed to have the greatest rigidity are those involved in forming the heme pocket. These include the heme-binding segment (Cys14-Leu15-Gln16-Cys17-His18), the right wall of the heme pocket (residues 28 to 32), the floor of the heme pocket (46 to 49), and the left wall containing the Met80 ligand (78 to 82). The three α-helical regions containing more than two turns (residues 3 to 12, 61 to 69, and 88 to 101) also have lowered mobility, and segments of these three helices also form part of the heme pocket. Localized maxima in molecular flexibility occur at segments of the polypeptide chain that are looped away from the heme and occupy surface positions. These include the N-terminal region at the rear (residues -5 to -1), a β-turn at the right periphery (21 to 25), another β-turn at the bottom rear (36 to 39), the left rear wall of the molecule between two α-helices (54 to 58), the extreme left periphery (73 to 75), and the C-terminus of the molecule (residue 103). For several residues, side chain atoms possess lower B-factors than main chain atoms. These residues either form covalent or coordinate bonds to the heme (Cys14, His18 and Met80), or are directed into the interior of the molecule and form part of the heme pocket (Leu32, Tyr46, Tyr48, Asn52, Trp59 and Leu68). Figure III.15 shows that for both main chain and side chain atoms of iso-1-cytochrome *c*, there is a strong correlation between close positioning to the heme group and low observed B-factors.

Other trends relating to the temperature factors of iso-1-cytochrome *c* are evident. For side chains of iso-1-cytochrome *c*, there is a strong positive correlation between solvent accessibility and B-factor. This consideration is also related to proximity to the heme group, as those side chains which are close to this buried moiety will have a low degree of solvent accessibility. The side chains with the highest B-factors, relative to the main chain atoms of the same residue, are all hydrophilic and are directed into the solvent medium. These include Lys4, Lys5, Arg13, Glu21, Lys22, Gln42, Glu44, Lys54, Glu66, Asn70, Lys86, Lys87, Lys99 and Lys100. In addition, the greater overall structural rigidity of the right side (residues 1 to 48, and 91 to the C-terminus) as compared to the left side (residues 49 to 90) of the molecule is apparent from

Figure III.15.

One further point of note is that the variation of B-factors along the polypeptide chain determined for iso-1-cytochrome *c* at 1.23 Å resolution is very similar to that at a lower resolution of 2.8 Å (Louie *et al.*, 1988a). This indicates that refinement of B-factors gives meaningful results even when only data of moderate resolution is available. However, at 2.8 Å resolution, the overall average B of protein atoms in the yeast iso-1-cytochrome *c* structure was lower, 13.8 $Å^2$. Temperature factors refined against X-ray data of moderate resolution tend to be systematically low, due to the absence of the influence of weak, high resolution structure factors (Ringe and Petsko, 1985), and to the effect of disordered solvent on low angle X-ray scattering (Stuart and Phillips, 1986). An additional factor is likely the absence of hydrogen atoms in the 2.8 Å model.

An assessment of the internal mobility of the yeast iso-1-cytochrome *c* molecule obtained crystallographically can be compared with that obtained using NMR spectroscopy (Williams, 1988; Pielak *et al.*, 1988b; Williams *et al.*, 1985c). Dynamic information derived from NMR data confirm the lack of atomic mobility in the interior of the cytochrome *c* molecule. The large chemical shift differences in the CD1 and CD2 methyl groups of those leucine side chains that are internally positioned are reflected in the well-defined appearance on electron density maps and the low temperature factors of these side chains. Particularly interesting is the much lower temperature factors for the side chains of aromatic residues in yeast iso-1-cytochrome *c* observed to have slow flip rates (Phe10, Tyr46, Tyr48, Trp59, Tyr67 and Tyr97; average B-factor, 11.7 $Å^2$) as compared to those for aromatic rings that undergo rapid flipping (Phe-3, Phe36, Tyr74 and Phe82; average B-factor, 23.5 $Å^2$).

Correlated with the observed low mobility of internal groups in yeast iso-1-cytochrome *c* is a particularly high packing efficiency. Of 893 non-hydrogen atoms, 307 could be considered to be part of its central core, if this is defined as being within half of the molecular radius (7 Å) from the heme group. In total, these internal atoms make 640 inter-residue van der Waals contacts (centre-centre distance $\leq$ 4.0 Å), or on average 2.1 contacts per atom. An assessment of

the number and size of internal cavities, using the method of Connolly (1983), reveals only one internal cavity large enough to accommodate a probe sphere 1.4 Å in radius. Groups enclosing this internal cavity include the non-polar side chains of Leu32, Met64 and Leu98, as well as the CHB atom of the heme.

F. Crystal packing

The packing observed in crystals of yeast iso-1-cytochrome *c* (see Figure III.16) can best be described as an interleaving along the crystallographic *c* axis of sheets of protein molecules, with each of the molecules centred near (< 5 Å from) a 4_3 screw axis. Each sheet has a thickness of 1/8 of the *c* axis and alternate sheets have the cytochrome *c* molecules close to either the 4_3 axis in the middle of the *a* or the *b* crystallographic cell edge. Individual molecules make packing contacts not with other molecules in the same sheet, but with molecules in sheets above and below their own: in one direction in *c*, intermolecular contacts are generated by the diagonal

Figure III.16. Stereo drawing of the packing of molecules in crystals of yeast iso-1-cytochrome *c*. The closest neighboring molecules to the origin molecule (thick lines) are shown. The iron atoms of each of the heme groups of the molecules shown have been labelled. Overlaid are the boundaries of the crystallographic unit cell about the origin molecule. In total a portion of four sheets of molecules, which run parallel to the *ab* plane, are shown. Only one molecule in the second sheet (the origin molecule) and in the topmost sheet are presented. For clarity, the rearmost molecule belonging to the first sheet, and the front most molecule of the third sheet are not drawn. The origin molecule as pictured interacts with the sheet below through contacts generated by the diagonal two-fold, and with the sheet above through contacts generated by the 2_1 axis parallel to *a*.

two-fold operation, and in the other direction, by the 2_1 screw axes parallel to *a* or *b*. Only a small number of contacts is made between molecules related by the 4_3 screw axis, i.e. between every second sheet of molecules. Therefore, packing parallel to the *c* axis has the appearance of columns of cytochrome *c* molecules stacked very loosely end-to-end, with adjacent columns being antiparallel and displaced by 1/8 *c*. In total, three symmetry operations are found to generate all packing contacts, and Table III.14 lists the intermolecular hydrogen bond and van der Waals contacts formed. Two adjacent protein molecules in the crystal lattice are joined by an additional 29 hydrogen bonds, formed through direct bridging interactions made by 16 water molecules. The high protein content of crystals of yeast iso-1-cytochrome *c* (estimated to be ~64%, see Section III.D.1) is consistent with the observed relatively tight packing of the protein molecules. The largest solvent channels present in the crystals are narrow columns that run parallel to the crystallographic *c* axis and alongside the loose stacks of cytochrome *c* molecules described above. Interestingly, the heme plane almost exactly bisects the angle between the *a* and *b* crystallographic axes, and is nearly parallel to the *c* axis (the porphyrin heme plane normal makes an angle of 41.5° with *a*, and 3.6° with *c*). Therefore, with the crystallographic symmetry elements present, each heme group in the crystal is either perpendicular to or parallel to every other heme group.

Crystals of yeast iso-1-cytochrome *c* have the shape of a tetragonal bipyramid, with the shortest dimension being along the crystallographic *c* axis (which also corresponds to the longest unit cell axis length). The equatorial edges of the bipyramid parallel the *ab* diagonals of the unit cell. Thus, the external morphology of the crystals suggests that the fastest crystal growth occurs along the *ab* diagonals, which is in agreement with the tightest intermolecular contacts being made by molecules related by the diagonal two-fold symmetry operations (Table III.14). It is therefore tempting to speculate that crystal formation involves first, the interaction of two molecules of cytochrome *c* as a dimer, and then the association of such dimers to form two-layered sheets. The slowness of crystal growth in the *c* direction would be due to the weaker associations between sheets whose molecules interact through operation of the 2_1 screw axes parallel to the *a* and *b* crystallographic axes.

Table III.14. Intermolecular packing contacts occurring in crystals of yeast iso-1-cytochrome *c*

1. Symmetry element: 4_3 screw axis

a. 1/2−y, 1/2+x, z−1/4		b. y−1/2, 1/2−x, 1/4+z	
Lys-2 CD - Lys79 O	3.12 vdW	Ser47 OG - Sul O2	2.88 Hb
Ala-1 O - Lys79 NZ	3.53 Hb		
Gly1 O - Ser47 OG	3.06 Hb		

2. Symmetry element: diagonal 2-fold axis

a. y, x, −z		b. y−1, x+1, −z	
Thr-5 N - Asn56 O	2.91 Hb	Thr8 CG2 - Lys9 O	3.32 vdW
Thr-5 CG2 - His39 CD2	3.18 vdW	Lys5 CE - Leu9 CD1	3.59 vdW
His39 NE2 - Glu61 OE1	2.77 Hb	Thr8 OG1 - Thr8 OG1	3.32 vdW
His39 NE2 - Glu61 OE2	3.21 Hb	Thr8 CG2 - Thr8 CG2	3.32 vdW
Leu58 CD1 - Asp60 CB	3.69 vdW	Lys87 CE - Lys87 CE	3.01 vdW

3. Symmetry element: diagonal 2_1 screw axis

a. y−1, x, −z		b. y, x+1, −z	
Lys-2 NZ - Lys22 O	3.55 Hb	Lys22 N - Glu88 OE1	3.11 Hb
Lys73 NZ - Sul O4	3.34 Hb	His33 CE1 - Glu88 OE1	2.66 vdW
Lys73 NZ - Sul O2	3.19 Hb		
Lys86 O - Glu103 CB	3.26 vdW		
Lys87 CA - Glu103 O	3.49 vdW		
Gly88 N - Glu103 OXT	3.04 Hb		

4. Symmetry element: 2_1 screw axis || *a*

a. 1/2+x, 1/2−y, 1/4−z		b. x−1/2, 3/2−y, 1/4−z	
Gln42 OE1 - Gly77 N	3.31 Hb	Gln16 OE1 - His26 N	3.20 Hb
Gln42 NE2 - Pro76 O	3.28 Hb	Gln16 NE2 - Gly24 O	2.50 Hb
Tyr48 O - Lys73 NZ	2.92 Hb	Lys27 NZ - Glu44 OE1	3.25 Hb
Asp50 OD1 - Asn70 ND2	2.94 Hb	Lys27 NZ - Glu44 OE2	3.47 Hb
Asp50 OD2 - Asn70 ND2	3.37 Hb		
Ile53 CG2 - Tml72 O	3.39 vdW		
Ile53 CD1 - Lys73 CD	3.32 vdW	c. 1/2+x, 3/2−y, 1/4−z	
Lys54 NZ - Asn70 OD1	3.34 Hb	Pro25 CB - Ala81 O	3.25 vdW
Lys54 NZ - Gly83 O	2.20 Hb	Pro25 CG - Ala81 O	3.29 vdW

The symmetry operations listed are those which when applied to the origin molecule generate the symmetry equivalent neighboring molecule. Following each contact is the length (in Å) of the interaction, and whether a hydrogen bond (Hb) or van der Waals contact (vdW) is made.

Yeast iso-1-cytochrome *c* is known to dimerize in solution (Bryant *et al.*, 1985), through the formation of an intermolecular disulfide bond involving Cys102. Since it was considered likely that the strongest intermolecular interactions occurring during the process of crystal nucleation involved two iso-1-cytochrome *c* molecules associating as a dimer, it was of interest to determine if these dimer interactions could promote the covalent dimerization of iso-1-cytochrome *c*, or if this process could actually occur within dimers present in crystals of this protein. As shown in Table III.14, dimer contacts made by a pair of iso-1-cytochrome *c* molecules in the crystal lattice are of two types. Those generated by the symmetry operation (y-1, x+1, z) involve the antiparallel association of the N-terminal helices of two adjacent molecules, whereas those generated by (y, x, -z) involve the packing of the rear, lower left of two molecules. However, neither dimer interaction observed in the crystal lattice of yeast cytochrome *c* appears to be related to the interaction between two iso-1-cytochrome *c* molecules which would be required for intermolecular disulfide bond formation to occur; two Cys102 residues of dimer-associated molecules are no closer than 17 Å in both forms of dimers.

G. Distinctive features of the yeast iso-1-cytochrome *c* molecule

In this section, several distinctive features of the yeast iso-1-cytochrome *c* molecule are described, in terms of their structures and their intramolecular interactions formed.

1. N-terminus

Typical of fungal cytochromes *c*, yeast iso-1-cytochrome *c* has an extension, 5 amino acid residues in length, at its N-terminus. Residues -4 to -1 have ϕ, ψ values characteristic of an extended conformation (Schulz and Schirmer, 1979). Gly1, the residue preceding the start of the N-terminal α-helix, adopts an unusual conformation ($\phi = -132°$, $\psi = -146°$; accessible only to glycine residues), causing the polypeptide chain to undergo a bend at this position. Thus, the N-terminal segment in yeast iso-1-cytochrome *c* is projected parallel to the rear surface of the molecule (see Figure III.5), making few interactions with the body of the molecule. The most prominent contacts are made by the side chain of Phe(-3), which occupies a small, surface pocket

formed by the atoms Glu61 CG and CD, Asn92 CA, and Ile95 CB and CG2. The lack of intramolecular interactions formed by the N-terminal segment is reflected in the high temperature factors of residues -5 to -1 (see Figure III.15).

Cytochrome *c* can be reversibly unfolded by a variety of treatments (extremes in pH, heat, the presence of urea or guanidine hydrochloride) (Saigo *et al.*, 1986; Bryant *et al.*, 1985). Comparative studies on the denaturation of homologous cytochromes *c* have shown that there is a good correlation between the relative susceptibilities of a given cytochrome *c* to each denaturant (Osheroff *et al.*, 1980; Saigo *et al.*, 1986; Bryant *et al.*, 1985). The high flexibility of the N-terminal extension in yeast iso-1-cytochrome *c*, and thus presumably the ease with which this region can be unfolded, may contribute to the lower stability under denaturing conditions of cytochromes *c* possessing this extension (Saigo, 1986; Margoliash and Schejter, 1966; Ochi *et al.*, 1983). An interesting point relating to the unfolding of cytochrome *c* is the specific stabilizing effect of certain anions (e.g. phosphate) (Osheroff *et al.*, 1980). In yeast iso-1-cytochrome *c*, the occupation of the anion binding site (see Section III.D.4 and Figure III.14) may a factor in stabilizing the folded state of the N-terminal region of the molecule.

2. Arginine 13

In all known primary sequences of eukaryotic cytochromes *c*, the side chain of residue 13 is basic. This residue has been proposed to play a role in stabilizing the heme crevice through forming a salt bridge to the acidic side chain (usually glutamic acid) of residue 90 on the opposite side of the crevice (Osheroff *et al.*, 1980). In yeast iso-1-cytochrome *c*, despite the occurrence of an apparently compensating substitution of an aspartate at position 90, which could serve to maintain an appropriate distance between the two groups involved in the ionic interaction (Osheroff *et al.*, 1980), a salt bridge is not formed. Apparently, because the aspartate side chain is branched at the CG atom, it is prevented from adopting a conformation directing it toward the side chain of Arg13 by steric conflicts with the intervening Leu9 and Leu85. Instead, the Arg13 side chain is directed more into the solvent medium; the guanidinium group is positioned directly above the side chain of Phe82, and forms hydrogen bonds to a number of water molecules. The absence of an

intramolecular salt bridge between residues 13 and 90 may be an additional factor in the increased susceptibility of yeast iso-1-cytochrome *c* to alkaline isomerization and denaturation by guanidine hydrochloride (Saigo, 1986; Osheroff *et al.*, 1980). Based on the structure of yeast iso-1-cytochrome *c*, two other roles for residue 13 can be suggested. These are the occlusion, by the aliphatic portion of the side chain, of solvent from the top of the heme crevice, coupled with the participation of the positively charged side chain group in intermolecular interactions with redox partners (Poulos and Finzel, 1984; Ferguson-Miller *et al.*, 1978). Notably, both Falk *et al.* (1981) and Satterlee *et al.* (1987) have previously observed from NMR studies that residue 13 influences the character of the heme environment of cytochrome *c*.

3. ϵ-N-Trimethyl lysine 72

The cytochromes *c* of fungi and plants are tri-methylated at the epsilon nitrogen of lysine 72 (Paik *et al.*, 1989). This residue is specifically-modified by the enzyme S-adenosylmethionine: protein-lysine N-methyltransferase (DiMaria *et al.*, 1979). Various roles, generally relating to involvement in forming intermolecular interactions, have been proposed for the tri-methylated lysine residue. Polastro *et al.* (1978) have shown that the methylated iso-1-cytochrome *c* species binds more tightly to mitochondria than does the unmethylated species, possibly through an increased affinity for cytochrome *c* oxidase (Holzschu *et al.*, 1987). Farooqui *et al.* (1981) have found that methylated iso-1-cytochrome *c* is degraded intracellularly at a decreased rate, likely as a result of the indirect protection from proteolytic attack effected by increased binding to mitochondria. Methylation of lysine 72 has been suggested to be required for transport of newly synthesized cytochrome *c* from the cytoplasm to the inner mitochondrial membrane (DiMaria *et al.*, 1979; Paik *et al.*, 1980). Finally, the presence of Tml72 has been shown to result in an increased binding affinity of iso-1-cytochrome *c* for cytochrome b_2 (Holzschu *et al.*, 1987). Despite these findings, Holzschu *et al.* (1987) have shown that replacement of Tml72 by an arginine residue has only minor effects on the properties of yeast iso-1-cytochrome *c*. Thus, the functional significance of trimethylation of the lysine at position 72 is not yet firmly established.

The electron density and environment of Tml72 in yeast iso-1-cytochrome *c* is shown in Figure III.17. The side chain of Tml72 has essentially a fully extended conformation, and projects directly toward the front of the molecule. The bulky and aliphatic trimethylated amine is positioned at the front face of iso-1-cytochrome *c*, and has one of its methyl groups packed against Ala81. Relative to other lysine side chains in iso-1-cytochrome *c*, the side chain of Tml72 is fairly rigidly positioned (average B-factor 20.7 Å^2), despite forming only the single intramolecular interaction.

The structural determination of yeast iso-1-cytochrome *c* coupled with other functional studies indicate two possible roles for trimethylation of lysine 72. First, the aliphatic portion of the side chain of Tml72, in occupying a position adjacent to the Met80 ligand, may contribute to maintaining the integrity of the heme pocket. This would be consistent with the lowered stability of a variant iso-1-cytochrome *c* in which trimethyllysine 72 is replaced by an arginine (Holzschu *et al.*, 1987). The less hydrophobic side chain of the latter residue would be unlikely to pack against residue 81 or to adopt the same rigid conformation observed for Tml72. Secondly, the

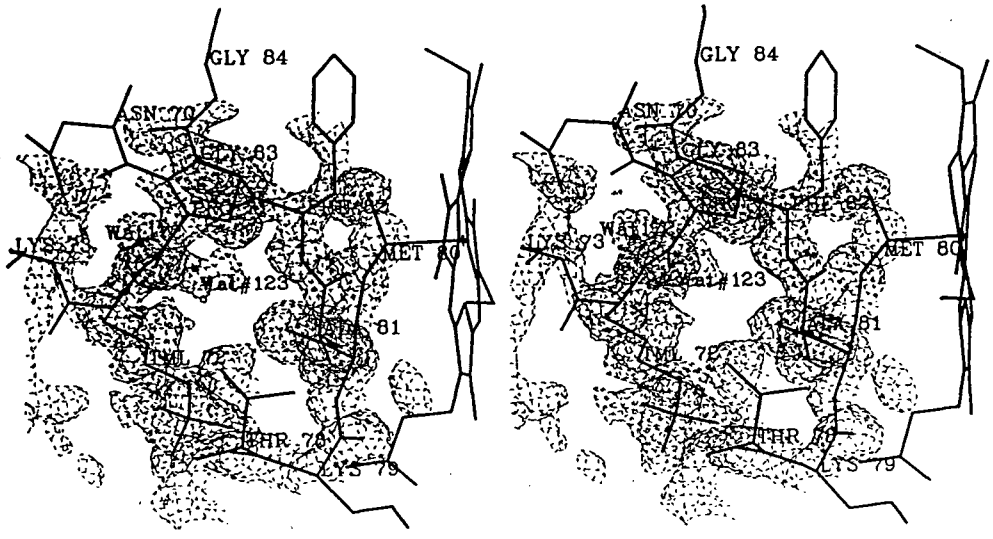

Figure III.17. Stereo diagram of the local environment of Tml72 in yeast iso-1-cytochrome *c*. Representative electron density (dashed line envelopes) in this region of the molecule is shown from a $2F_o-F_c$, α_c map. The direction of view is from the top of the molecule. The quaternary ammonium group of Tml72 occurs at the front face of the iso-1-cytochrome *c* molecule, which is located at the lower edge of this diagram.

trimethylated amine group may provide a hydrophobic patch for interaction with other proteins. Holzschu *et al.* (1987) have shown that the Arg72 variant of iso-1-cytochrome *c* has an ~4-fold higher K_m for binding to cytochrome b_2. Additional evidence is provided by NMR studies of the interaction between cytochrome b_5 and cytochrome *c* from the yeast *Candida krusei*, which show that the linewidth of the trimethyl group of Tml72 is considerably broadened upon bimolecular complexation (Eley and Moore, 1983). The indicated decrease in mobility of Tml72 suggests that this residue forms part of the protein-protein interface. The hydrophobic patch also appears to be conserved in rice and tuna cytochromes *c*: in rice cytochrome *c*, the conformation of Tml72 is similar to that in yeast; and in tuna cytochrome *c*, Lys72 is not methylated, but the occurrence of an isoleucine at position 81 (Ala in yeast, Val in rice) places a hydrophobic group at the corresponding spatial position. From an analysis of the primary sequences of eukaryotic cytochromes *c*, it is notable that whereas plant and fungal cytochromes *c* have a small side chain (Val or Ala) at residue 81, animal cytochromes *c*, which are not methylated, have a conserved isoleucine residue at this position. In addition, amino acid replacements (by Ala and Val) at position 81 of horse cytochrome *c*, generated using semi-synthesis, have been shown to affect the binding affinity of the variant proteins to cytochrome *c* oxidase (Boots and Tesser, 1987; see Table I.2).

4. Cysteine 102

In the structure of yeast iso-1-cytochrome *c*, Cys102 is positioned on the inward-facing side of the C-terminal α-helix. The sulfhydryl group of Cys102 is inaccessible to solvent, and occupies a predominantly hydrophobic pocket formed by the side chains of Val20, Leu32, Ile35, Phe36 and Leu98. The $\chi 1$ torsion angle of -55° places the SG atom in a position where it can potentially form a hydrogen bond to the carbonyl group of Leu98 or Leu32 (see Figure III.18).

Because the side chain of Cys102 is normally directed into the interior of the molecule, covalent dimerization of iso-1-cytochrome *c* via intermolecular disulfide bond formation (Bryant *et al.*, 1985) would require at least a partial unfolding of the C-terminal helices of the two molecules involved. Since the C-terminal helix is an important structural component of the heme pocket,

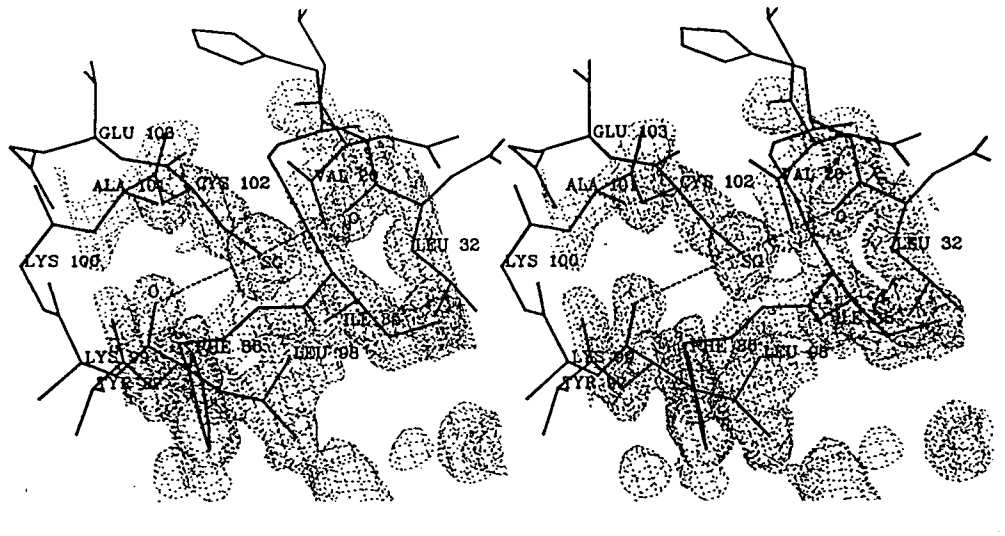

Figure III.18. Stereo diagram of the local environment of Cys102 in yeast iso-1-cytochrome *c*. Representative electron density (dashed line envelopes) in this region of the molecule is shown from a $2F_O-F_C$, α_C map. Cys102 occurs in the C-terminal α-helix, which is partially shown oriented vertically at the left of the diagram. The internal face of this helix faces the right in this diagram. Possible hydrogen bonds involving the sulhydryl group of Cys102 are indicated as thick dashed lines.

dimerization would be expected to destabilize the overall fold of the protein. These considerations are thus consistent with experiments which show that dimerization increases the susceptibility of iso-1-cytochrome *c* to the following: denaturation by heat, guanidine hydrochloride, urea, and acid (Bryant *et al.*, 1985); alkaline isomerization, and imidazole and azide binding to the heme iron (Saigo, 1986); and digestion by proteases (Motonaga *et al.*, 1965). From measurements of the same susceptibilities, Bryant *et al.* (1985) and Motonaga *et al.* (1965) have found that reaction of Cys102 with iodoacetate or thiosulfite also causes destabilization of iso-1-cytochrome *c*. Inspection of the structure of yeast iso-1-cytochrome *c* indicates that the pocket occupied by the SG atom of Cys102 is not large enough to accommodate many additional atoms, and thus derivitization of the sulfhydryl group with bulky substituents would likely perturb the native conformation of (at least) this region of the molecule. It is also notable that Moench and Satterlee (1989) have shown using NMR techniques that both dimerization and modification with 5,5'-dithiobis(2-nitrobenzoate) cause significant structural changes in the iso-1-cytochrome *c* molecule, which in turn perturb the heme environment.

H. Amino acid residues at which site-specific replacements have been made

Much of this dissertation describes structural studies on variant iso-1-cytochromes *c* in which the naturally occurring amino acid residues (Phe82, Arg38 and Cys102) have been specifically replaced. This section details the conformations and local environments of Phe82 and Arg38. Cys102 was discussed above.

1. Phenylalanine 82

Phe82 is part of the fairly extended 80's region of polypeptide chain which sweeps across the upper left face of the heme group. The side chain of Phe82 adopts a nearly ideal conformation, with torsion angles $\chi 1 = 172^{\circ}$ and $\chi 2 = -104^{\circ}$. The phenyl ring is packed against the heme group, and occupies a hydrophobic pocket that is located just below the surface of the protein molecule (Figure III.19). This pocket is formed by the upper edge of the heme group, the side chains of Leu68, Pro71, Met80 and Leu85, and the backbone of residues 83 and 84. The plane of the phenyl ring is nearly parallel to that of the heme group. The interplanar

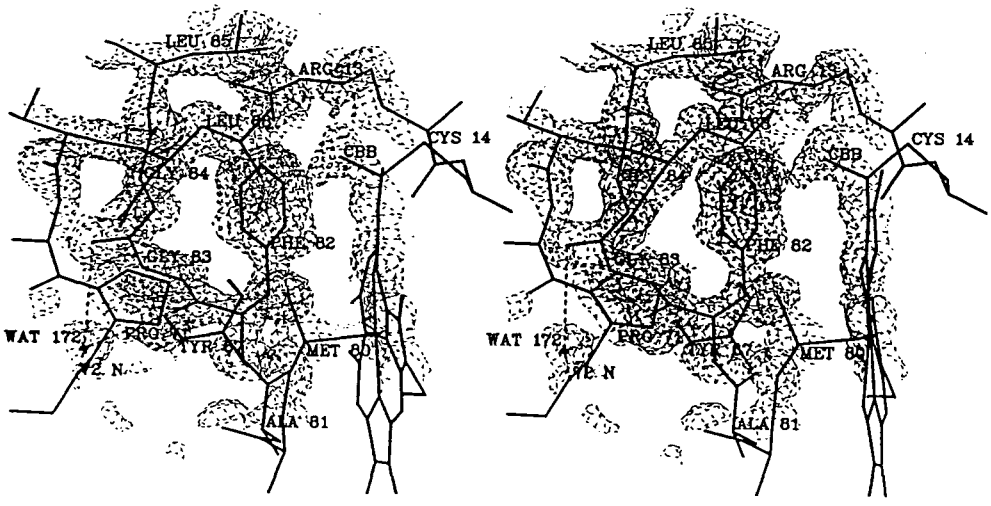

Figure III.19. Stereo diagram of the local environment of Phe82 in yeast iso-1-cytochrome *c*. Representative electron density (dashed line envelopes) in this region of the molecule is shown from a $2F_o-F_c$, α_c map. The direction of view is from the upper front of the molecule. Indicated as thick dashed lines are hydrogen bonds between Wat172 and the main chain amide nitrogen of residue 72 and the carbonyl groups of Phe82 and Gly83.

distance and angle are ~5 Å and ~23°, respectively. Van der Waals contacts that are formed by the phenyl ring within the hydrophobic pocket (CE1 - Gly84 CA: 3.4 Å; CZ - Hem CBB: 3.3 Å) likely promote the parallelism of the phenyl and heme groups. It is notable that it is the extended conformation of residues 83 and 84, in which the invariant Gly84 adopts a conformation inaccessible to residues having a side chain, which permits the close packing interactions made by the left face of the phenyl ring. The packing of the side chain of Phe82 adjacent and approximately parallel to the heme group is consistent with the proposed interaction of the π-electron systems of these two groups (Pielak *et al.*, 1986; Poulos and Kraut, 1981).

A notable feature of the local environment of the side chain of Phe82 is the absolute exclusion of polar groups. The phenyl ring itself is positioned to occlude the access of solvent to the heme crevice. The orientation of the peptide bond planes in the adjacent segment of polypeptide chain backbone (residues 81 through 84) is such that the carbonyl groups are directed away from the heme group. Thus, Phe82 clearly plays a structural role in maintaining the non-polarity of the upper left portion of the heme pocket.

2. Arginine 38

Arg38 occurs in a flexible, external loop positioned at the rear right of the floor of the heme pocket. Its side chain adopts a fairly extended conformation, such that the guanidinium group occupies a position where it can interact with the buried propionic acid group of heme pyrrole ring A. This interaction is mediated by two internal water molecules, which form bridging hydrogen bonds between the NH1 and NE atoms of Arg38, and the O1A atom of the propionic acid (see Figure III.20). The position of the guanidinium group is further fixed by hydrogen bonds between both the NH1 and NH2 atoms and the carbonyl group of His33. On the solvent exposed edge of the guanidinium group, NH2 forms a hydrogen bond to an external water molecule (Wat108). The extensive network of hydrogen bonds formed by this guanidinium group prompted Wallace to suggest that the function of Arg38 is to stabilize the conformation of the bottom loop of the protein molecule (see Table I.2 and references listed therein).

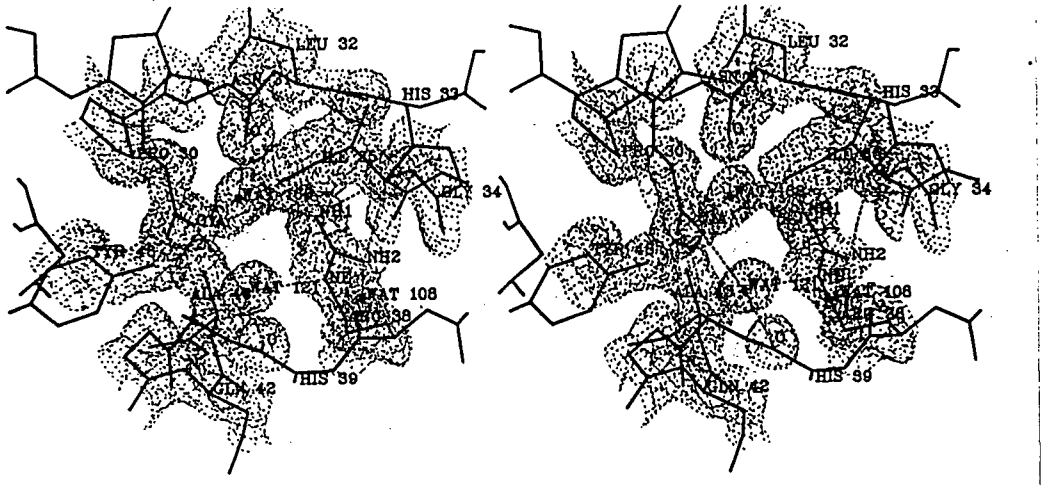

Figure III.20. Stereo diagram of the local environment of Arg38 in yeast iso-1-cytochrome *c*. Representative electron density (dashed line envelopes) in this region of the molecule has been shown from a $2F_o-F_c$, α_c map. The direction of view is from the solvent medium and toward the His18 face of the heme group. The hydrogen bonding network involving the guanidinium group of Arg38 and the internal water molecules Wat121 and Wat168 is represented as thick dashed lines.

IV. STRUCTURAL COMPARISON OF YEAST ISO-1-, TUNA AND RICE CYTOCHROMES *c*

The availability of well-refined structures for yeast iso-1-, tuna and rice cytochromes *c* permits a comparison of the structural features of three members of the cytochrome *c* family which have ~40% amino sequence divergence, and which have been determined by independent workers, in different crystal space groups. Key details of the structure determinations of these cytochromes *c* are listed in Table IV.1. An amino acid sequence alignment of these three proteins is shown in Figure IV.1. The numbers of amino acid sequence identities between pairs of these three cytochromes *c* are presented in Table IV.2.

Table IV.1. Details of the structure determinations of the various cytochromes *c*

Property	Yeast iso-1-	Tuna (reduced)	Tuna (oxidized)	Rice
Space group	$P4_32_12$	$P2_12_12$	$P4_3$	$P6_1$
No. of molecules per asymmetric unit	1	1	2	1
Unit cell dimensions (Å)				
a	36.46	37.33	74.42	43.78
b	36.46	87.10	74.42	43.78
c	136.86	34.44	36.30	110.05
Resolution (Å)	1.23	1.50	1.80	1.50
No. of reflections	12513	13840	16831	n.s.[b]
Method of refinement[a]	H.K.	J.L.	J.L.	H.K.
R-factor	0.192	0.173	0.208	0.200
Estimated coordinate accuracy (Å)	0.18	0.15	0.20	n.s.
R.m.s. deviations from ideal stereochemistry[c]				
bond distances (Å)	0.026	0.020	0.021	0.018
angle distances (Å)	0.050	0.048	0.036	0.040
ω torsion angles (°)	3.6	7.1	4.4	3.6
No. of solvent molecules	117	53	49	46

[a]H.K.: stereochemically restrained least-squares (Hendrickson and Konnert, 1981).
J.L.: simultaneous energy minimization and reciprocal space refinement (Jack and Levitt, 1978).

[b]n.s.: not specified by original authors.

[c]All deviations from ideality are with respect to the dictionary of ideal distances used in the refinement of yeast iso-1-cytochrome *c*.

Table IV.2. Matrix of structural equivalence and amino acid sequence identity for yeast iso-1-, tuna and rice cytochromes *c*

	Yeast iso-1-	Tuna (reduced)	Tuna (oxidized)	Rice
Yeast iso-1-	108	66 (64%)	66 (64%)	64 (59%)
Tuna (reduced)	0.54	103	103 (100%)	65 (63%)
Tuna (oxidized)	0.55	0.42	103	65 (63%)
Rice (oxidized)	0.51	0.62	0.59	111

The number of amino acid residues in each protein is given along the diagonal of the matrix. The upper triangular portion of the matrix indicates the number of identical residues found in common between each pair of proteins, as aligned in Figure IV.1. The number of identical residues as a percentage of the total number in the smaller protein is given in parentheses.

The lower triangular portion of the matrix indicates for each pair of proteins, the overall r.m.s. deviation (in Å) in the positions of corresponding main chain atoms of residues 1 to 103. Values for oxidized tuna cytochrome *c* are averages for the inner and outer molecules in the crystallographic asymmetric unit.

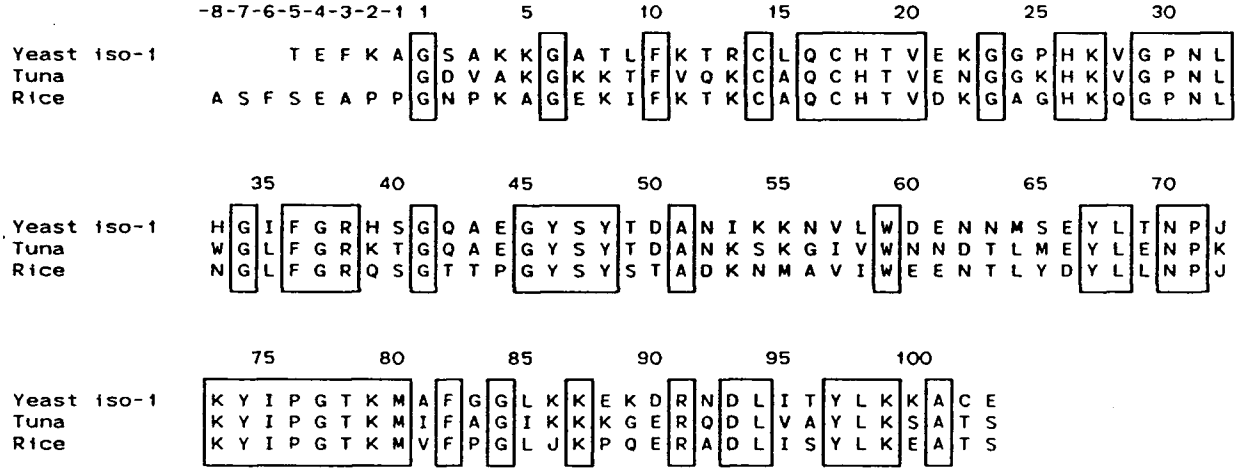

Figure IV.1. Primary sequence alignment of yeast iso-1-, tuna and rice cytochromes *c*. The amino acid sequences of yeast iso-1- (Smith *et al.*, 1979), tuna (Kreil, 1963, 1965), and rice (Mori and Morita, 1980) cytochromes *c* are shown aligned so as to maximize the sequence homology. Amino acid residues which occur identically in all three protein sequences are enclosed by boxes.

A. Overall structural conservation

1. Polypeptide backbone conformation

A comparison of the polypeptide backbone conformation of yeast iso-1-cytochrome *c* with those of tuna and rice cytochromes *c*, in terms of the average positional differences between main chain atoms of corresponding residues, is shown in Figure IV.2. Also, summarized in Table IV.2 are the overall r.m.s. differences in positions of common main chain atoms between individual pairs of cytochromes *c*. These data indicate that amongst these cytochromes *c*, there is a high degree of structural similarity. The overall main chain r.m.s. deviation between individual determinations of tuna ferricytochrome *c* (0.27 Å; Takano and Dickerson, 1981b) provides a baseline against which

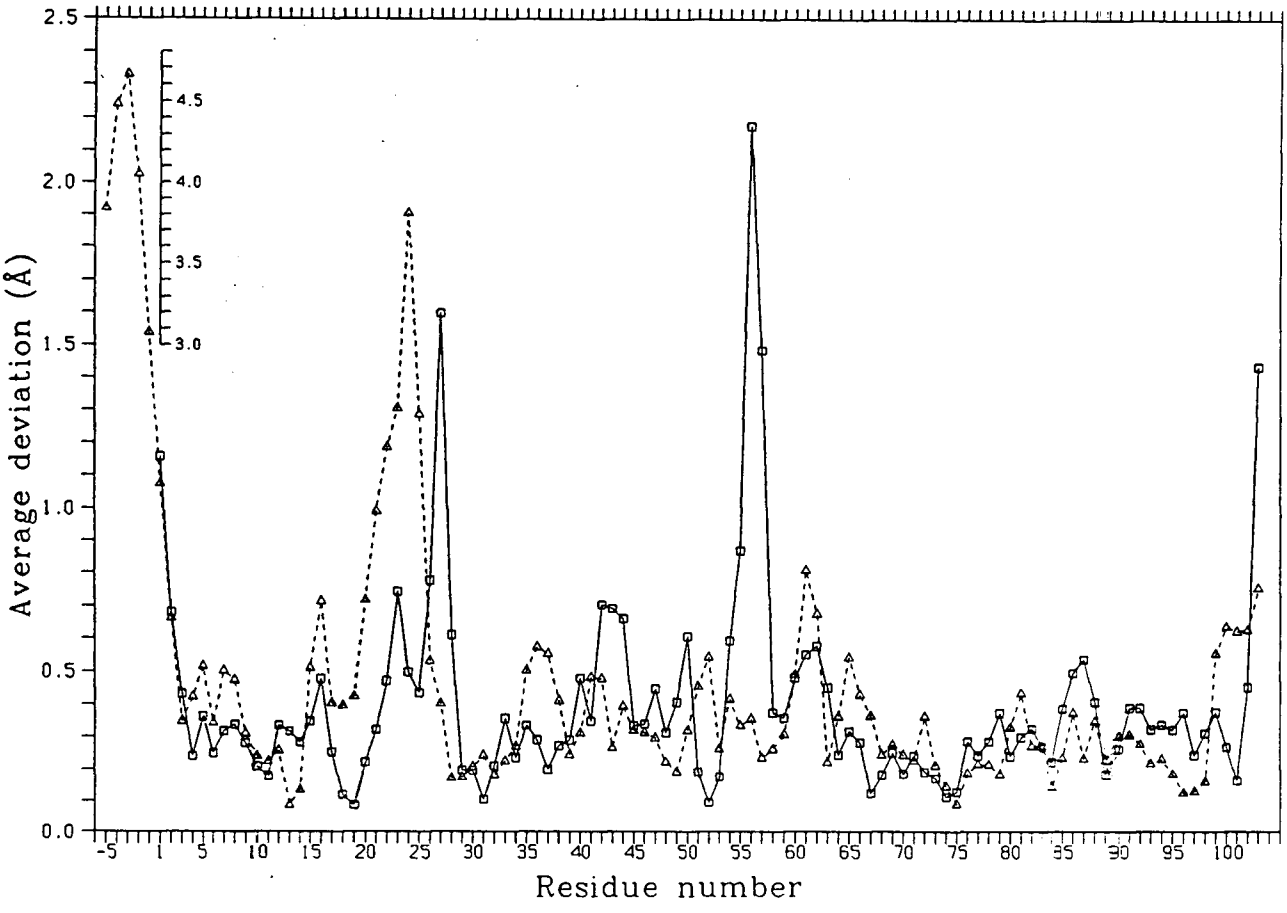

Figure IV.2. Comparison of the polypeptide backbone conformations of yeast iso-1-, tuna and rice cytochromes *c*. The plotted lines indicate the average deviations observed between the positions of main chain atoms in corresponding residues of yeast and tuna (□——□), and yeast and rice (Δ--Δ) cytochromes *c*. Differences in the positions of residues -5 to -1 in yeast and rice cytochromes *c* exceed 3 Å, and have been illustrated on the separate, inset scale.

the magnitude of differences between other pairs of cytochromes *c* can be assessed. Comparison of different species of cytochromes *c* yields values which are not greatly larger. The similarity in overall folding of yeast iso-1-, tuna and rice cytochromes *c* is also evident from a comparison of main chain dihedral torsion angles. Disregarding those phi and psi values affected by large localized differences in conformation (as discussed below), the overall r.m.s. difference between corresponding phi and psi angles are 9.9° and 9.4° respectively, between yeast and tuna, and 10.2° and 9.3° between yeast and rice cytochromes *c*.

Comparisons of several members of a single protein family (e.g. the serine proteases and the globins) have established that tertiary structure can be highly conserved despite the presence of only moderate amino acid sequence homology (Creighton, 1984). This is clearly the case for yeast iso-1-, tuna and rice cytochromes *c*. For these three proteins, this result is particularly unsurprising, since residues whose side chains are directed into the solvent medium account for the vast majority of sites at which variability in amino acid sequence occurs (Figure IV.3). Amino

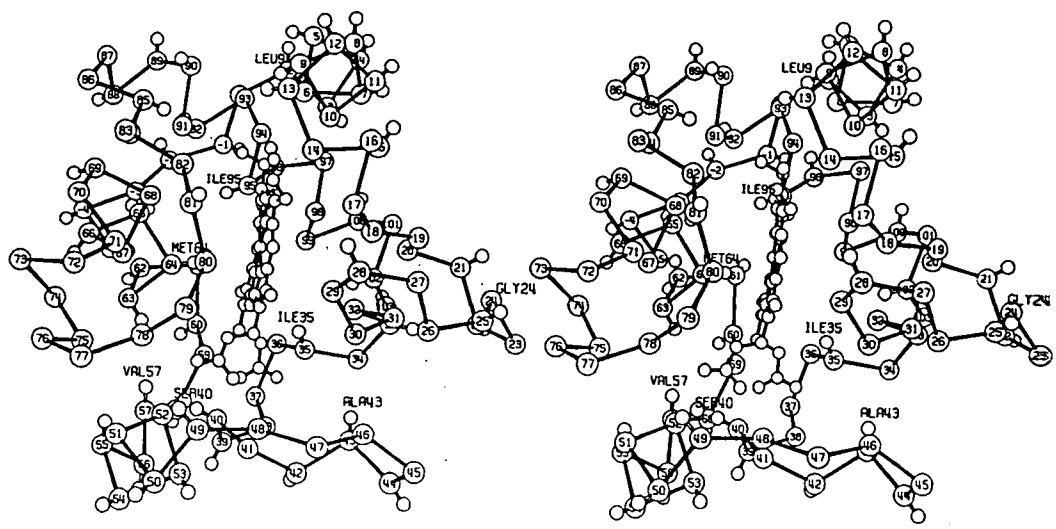

Figure IV.3. Stereo drawing showing the sites of variability in amino acid identity between yeast iso-1-, tuna and rice cytochromes *c*. The positions of the CB side chain carbons are shown for those amino acid residues whose identities vary between the three cytochromes *c*. This figure serves to indicate the orientation of the variant side chains with respect to the polypeptide backbone of the cytochrome *c* molecule, which is represented here as an α-carbon backbone. Internally directed variant side chains that cause conformational differences between the three proteins (see text for discussion) have been labelled.

acid substitutions which result in substantial differences in polypeptide chain conformation are described in Section IV.B.

Chothia and Lesk (1985) have determined that six regions of the cytochrome *c* molecule have an equivalent folding pattern in tuna cytochrome *c*, *R. rubrum* cytochrome c_2, *P. denitrificans* cytochrome c_{550}, and *P. aeruginosa* cytochrome c_{551}. These six segments of polypeptide chain (residues 5 to 11, 14 to 20, 27 to 33, 60 to 71, and 90 to 101) form the core structure of the cytochromes *c*, and all are involved in creating the heme pocket. It is notable that between yeast iso-1-, tuna and rice cytochromes *c*, the same six regions constitute the most spatially conserved segments of main chain (see Figure IV.2). In contrast, those regions of the polypeptide chain looped away from the heme group account for the majority of instances of localized conformational differences between the three proteins (see Figures III.15 and IV.2).

2. Side chain conformations

Between yeast iso-1-, tuna and rice cytochromes *c*, the agreement between side chain dihedral angles of corresponding residues is also very high. Table IV.3 shows a comparison of side chain conformations in these three cytochromes *c*, in terms of both the overall average differences in χ values and the distribution of difference values, for several subgroups of side chains. Between corresponding residues of both yeast iso-1- and tuna, and yeast iso-1- and rice cytochromes *c*, approximately 70% of all comparable side chain dihedral angles are equivalent [two side chain dihedral angles are considered equivalent if they differ by less than 40° (Summers *et al.*, 1987; Kamphuis *et al.*, 1985)]. The results for the correspondence of comparable $\chi 1$ angles shown in Table IV.3 can be directly compared with results compiled by Summers *et al.* (1987) in their analysis of seven pairs of proteins having various degrees of amino acid sequence homology. The fairly high sequence homology between these three cytochromes *c*, about 60% amino acid identity between yeast iso-1- and both tuna and rice cytochromes *c*, gives rise to just over 70% equivalence in $\chi 1$ angles, which agrees well with the data presented in Table 5 of Summers *et al.* (1987). Although the overall mean values for angular deviations presented in Table IV.3 obscure the details of differences in the orientations of individual side chains, they also point out

Table IV.3. Comparison of side chain torsion angles in yeast iso-1-cytochrome *c* with those found in the tuna (T) and rice (R) proteins

Side chain dihedral angles considered	Protein and no. of residues	Mean angular deviation[a] (°)	Number of angles					
			Total	Having deviations[b]				
				0–10	10–20	20–30	30–40	>40°
a) All comparable[c]	T 103	37.4	155	66	28	6	7	48
				0.43	0.18	0.04	0.05	0.31
	R 103	36.9	159	48	40	16	9	46
				0.30	0.25	0.10	0.06	0.29
b) Comparable[c] χ1	T 70	37.6	70	36	8	2	3	21
				0.51	0.11	0.03	0.04	0.30
	R 72	33.2	72	32	15	6	2	17
				0.44	0.21	0.08	0.03	0.24
c) Identical residues[d]	T 62	29.2	119	60	25	4	3	27
				0.50	0.21	0.03	0.03	0.23
	R 60	30.7	118	41	34	9	7	27
				0.35	0.29	0.08	0.06	0.23
d) Conserved residues in all three[d]	T 49	23.8	89	50	19	2	3	15
				0.56	0.21	0.02	0.03	0.17
	R 49	28.9	89	34	29	6	2	18
				0.38	0.33	0.07	0.02	0.20
e) Invariant residues in eukaryotic cyt. *c*[e]	T 21	19.7	45	26	11	1	1	6
				0.58	0.24	0.02	0.02	0.13
	R 21	25.3	45	17	14	5	1	8
				0.38	0.31	0.11	0.02	0.18
f) Residues that form heme pocket	T 18	9.1	35	28	5	0	0	2
				0.78	0.17	0.0	0.0	0.06
	R 18	17.3	35	18	11	2	0	4
				0.50	0.33	0.06	0.0	0.11

[a]This value refers to the mean of the absolute values of the individual angular deviations.

[b]Listed on the second line is the proportion of angles having deviations in each range.

[c]Individual χn (n=1,2,3,4 or 5) torsion angles in two given side chains were considered comparable if an equivalent pattern of branching is present beyond the third atom defining the torsion angle (e.g. Lysχ1 ≠ Valχ1; Lysχ2 ≠ Aspχ2; but Aspχ2 = Pheχ2).

[d]Identical residues between yeast and tuna, or yeast and rice cytochromes *c* are those for which the amino acid residue present in the tuna or rice sequence is the same as that found at the corresponding position in the yeast sequence. The conserved residues in these three proteins

Table IV.3 (continued)
have the same identity in the sequences of all of yeast iso-1-, tuna and rice cytochromes *c* (see Figure IV.1).

[e]The invariant residues in the primary structures of eukaryotic cytochrome *c* are listed in Table I.1.

the high degree of similarity in positioning of side chains. Considering all pairs of comparable side chain torsion angles between both yeast iso-1- and tuna, and yeast iso-1- and rice cytochromes *c*, the overall average difference in conformation is approximately 37°.

As was the case in the four pairs of related proteins having greater than 40% sequence homology considered by Summers *et al.* (1987), in yeast iso-1-, tuna and rice cytochromes *c* the side chains of residues whose identities show conservatism in the primary structure have a particularly high degree of conservation of conformation. Both the overall average deviations in side chain torsion angles, and the proportions of side chain angles showing only a small difference in conformation (Table IV.3) demonstrate that amino acid residues whose identities are the same in all three cytochromes *c* have more similar side chain conformations than those whose identities are the same in just yeast iso-1- and one of tuna or rice cytochrome *c*. Likewise, those residues which are invariant in the eukaryotic cytochromes *c* possess the greatest degree of conservation of side chain conformation. For these invariant residues (Arg38, Lys73 and Arg91 are excluded; see discussion below) the largest individual difference in the value of a side chain torsion angle is only 20°.

As Figure IV.4 illustrates, the heme groups of all three cytochromes *c* have very similar conformations, with comparable degrees of distortion from planarity, and essentially identical conformations of both propionic acid and vinyl side groups. Amino acid residues that interact closely with the heme group are highly conserved (18 of 23 are identical) and also display a high degree of structural conservation. Considering 31 side chain torsion angles in these 18 identical residues (those of Arg38 in all molecules, and of Leu94 in rice cytochrome *c* are excluded), the largest torsion angle difference that occurs between yeast iso-1-, and both rice and tuna

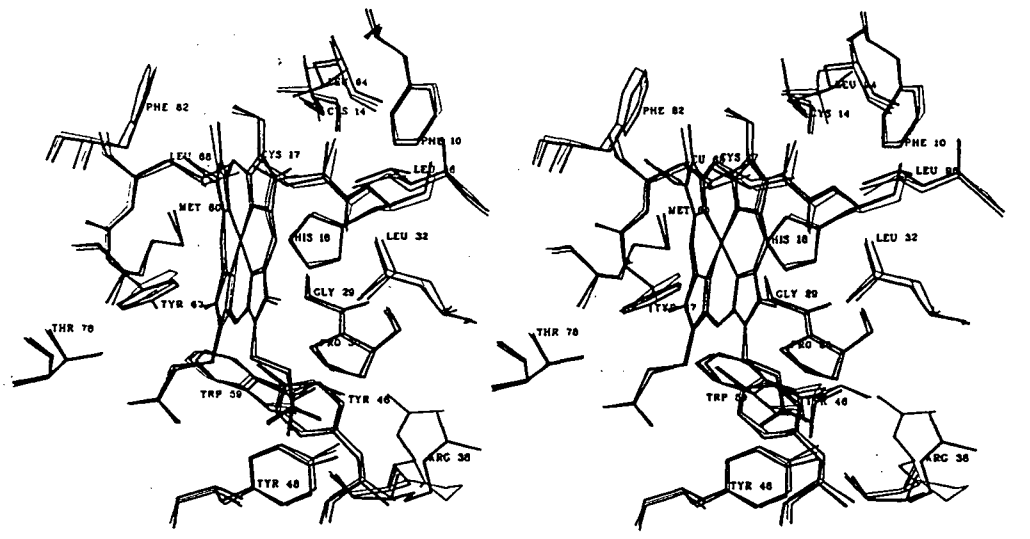

Figure IV.4. Conformations of the heme groups and of the side chains of conserved residues in the heme pockets of yeast iso-1- (thick lines), tuna (medium lines) and rice (thin lines) cytochromes *c*. Distinctly different conformations are observed only for Arg38 (all three proteins) and Leu94 (rice cytochrome *c*).

cytochromes *c* is only 17° (Table IV.3). Thus conservation of both the identity and conformation of side chains in the heme pocket, coupled with the similarity in folding of the polypeptide chain in the vicinity of the heme group (see discussion above), results in nearly identical heme environments in these three cytochromes *c*. Shellnut *et al.* (1981, 1979) have shown that the near-heme residues strongly influence the electronic distribution within the π-orbitals of the heme group. Thus it is likely that the highly conserved packing of protein groups within the heme crevice is an important determinant of the functional properties of the heme in cytochrome *c*.

3. Intramolecular hydrogen bonding

As would be expected from the strong similarity in both main chain and side chain conformations of yeast iso-1-, tuna and rice cytochromes *c*, intramolecular hydrogen bonding interactions are also highly conserved between these proteins. Among the three cytochromes *c*, the only deviations in the pattern of main chain hydrogen bonding from that occurring in the yeast protein (see Table III.10) are listed in Table IV.4. The most notable differences arise from unique conformations adopted by the polypeptide chain of tuna cytochrome *c* (more fully detailed in

Table IV.4. Differences between yeast iso-1-, tuna and rice cytochromes *c* in main chain hydrogen bonding

Hydrogen bond	Species of cytochrome *c*	Occurrence
27 N - 29 O	tuna	absent
28 N - 17 O	tuna	present
40 N - 57 O	tuna	absent
53 N - 50 O	rice	present
55 N - 51 O	rice	absent
55 N - 52 O	rice	present
56 N - 53 O	rice	present
66 N - 62 O	tuna,rice	present
69 N - 65 O	tuna,rice	absent
74 N - 71 O	rice	present
100 N - 96 O	rice	absent
101 N - 98 O	tuna	present
103 N -100 O	rice	present

The differences listed are with respect to the hydrogen bonding observed in yeast iso-1-cytochrome *c*, Table III.10. Only the main chain hydrogen bonding involving residues common to all three cytochromes *c* (residues 1 to 103) has been considered.

Section IV.B.1). The absence of the 27 N - 29 O hydrogen bond, and the occurrence instead of a 28 N - 17 O hydrogen bond, are associated with the absence of the γ-turn at residues 27 to 29. The lack of an antiparallel β-sheet interaction between residues 37 to 40 and 57 to 59 precludes the occurrence of the 40 N - 57 O hydrogen bond. The other variations in the pattern of hydrogen bonding are all minor, and arise from small distortions from regular hydrogen bonding within α-helical segments of polypeptide chain. These results further demonstrate that the conformation of an α-helix can be subtly influenced by its local environment in the protein molecule (see Section III.B.3).

A number of intramolecular hydrogen bonds involving both a main chain and a side chain atom are also conserved in all three cytochrome *c* molecules (see Table IV.5.). Additionally, the eight hydrogen bond interactions formed between the propionic acid oxygens of the heme group

Table IV.5. Conserved hydrogen bonds involving side chains in yeast iso-1-, tuna and rice cytochromes *c*[a]

2 N - Asp93 OD1	Tyr46 OH - 28 O
His18 ND1 - Pro30 O	52 N - Ser49 OG or Thr49 OG1
Asn31 ND2 - 21 O	Tyr67 OH - Met80 SD
His26 NE2 - 44 O	Lys79 NZ - Ser47 O
His33 N - Asn31 OD1	Met80 N - Thr78 OG1
43 N - Tyr48 OH	

[a]For those residues designated only by residue numbers, the identity of the amino acid is not conserved in all of yeast iso-1-, tuna and rice cytochromes *c*.

and protein atoms of the cytochrome *c* molecule (see Table III.9) are conserved in yeast iso-1-, tuna and rice cytochromes *c*. However, two notable differences in the interaction of the propionic acid groups do occur. In tuna cytochrome *c*, the side chain of Arg38 forms a direct hydrogen bond with the propionic acid group of pyrrole ring A (see Section IV.B.1). In rice cytochrome *c*, residue 52 is an aspartic acid, rather than an asparagine as in yeast iso-1- and tuna cytochromes *c*, and thus the hydrogen bond interaction must involve an acidic proton rather than the hydrogen atom of an amide nitrogen. In addition, the hydrogen bonding atom of the Asp52 side chain is the lower, rather than the upper atom.

B. Large conformational differences between yeast iso-1-, rice and tuna cytochromes *c*

1. Polypeptide chain backbone

Despite a high degree of structural similarity between these three proteins, the main chain conformation of yeast iso-1-cytochrome *c* differs substantially from those of rice and tuna cytochromes *c* at a number of locations (see Figure IV.2). In the majority of these instances, the differences observed are localized at sites of amino acid substitutions.

The surface β-turn that includes Gly23 has a distinct positioning in rice cytochrome *c* (see Figure IV.5). This apparently results from the substitution at residue 24 of an alanine for the glycine present in both the yeast and tuna proteins. A CB atom on residue 24 in yeast

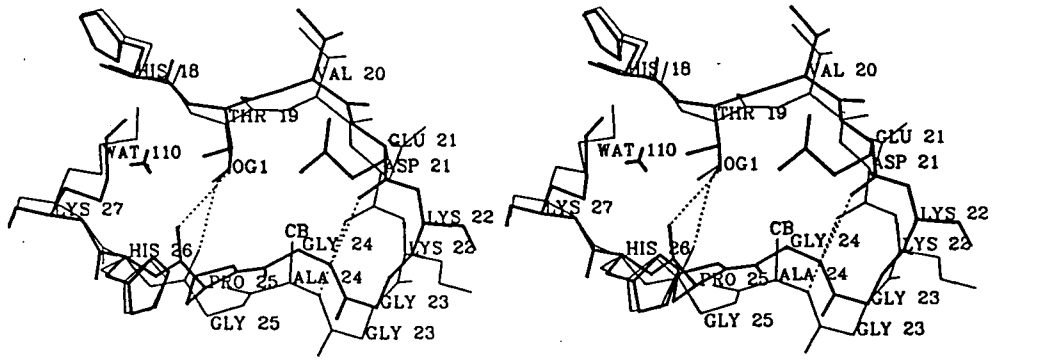

Figure IV.5. Differences in conformation between yeast iso-1- (thick lines) and rice (thin lines) cytochromes *c* at residues 20 to 25. A loop structure is formed by residues 18 through 27 in both proteins, and the hydrogen bonds (dashed lines) joining the two arms of the loop (24 N to 21 O, and Thr19 OG1 to 25 O) are conserved. However, the peptide bond linking residues 24 and 25 is of opposite orientation in the two proteins. Also, occurring only within the loop of the yeast protein is Wat110, which forms a hydrogen bond to the carbonyl group of Pro25.

cytochrome *c* would be positioned within 2.9 Å of both Glu21 O and Thr19 OG1. Thus the presence of Ala24 in rice cytochrome *c* appears to cause the β-turn to have a more open conformation, with the length of the hydrogen bond between 24 N and 21 O being 3.14 Å, as compared to 2.86 Å in yeast cytochrome *c*. Consequently, residues 24 to 26 assume a completely different conformation in the rice protein to allow the polypeptide chain beyond residue 26 to maintain close packing interactions with the rest of the molecule. These conformational changes are mediated by the ψ angle of Ala24 having a value ~180° different from that in the yeast and tuna proteins, as well as the presence of a glycine at position 25 (Pro in yeast; Lys in tuna) which adopts a conformation inaccessible to residues with a side chain.

Another example of a localized conformational difference is observed in the region of residue 56, which is displaced further into solvent in tuna cytochrome *c* (see Figure IV.6). This appears to arise from the replacement of two small side chains, Ser40 and Val57 in both yeast and rice cytochromes *c*, with bulkier ones, Thr40 and Ile57 in tuna cytochrome *c*. In the yeast protein, residues 37 to 40 interact closely with residues 57 to 59 in an antiparallel β-sheet type structure (Figure III.8). Consequently, if a threonine side chain were to occur at position 40, its CG2 atom would conflict sterically with the side chain of residue 57. In addition, the CD1 atom

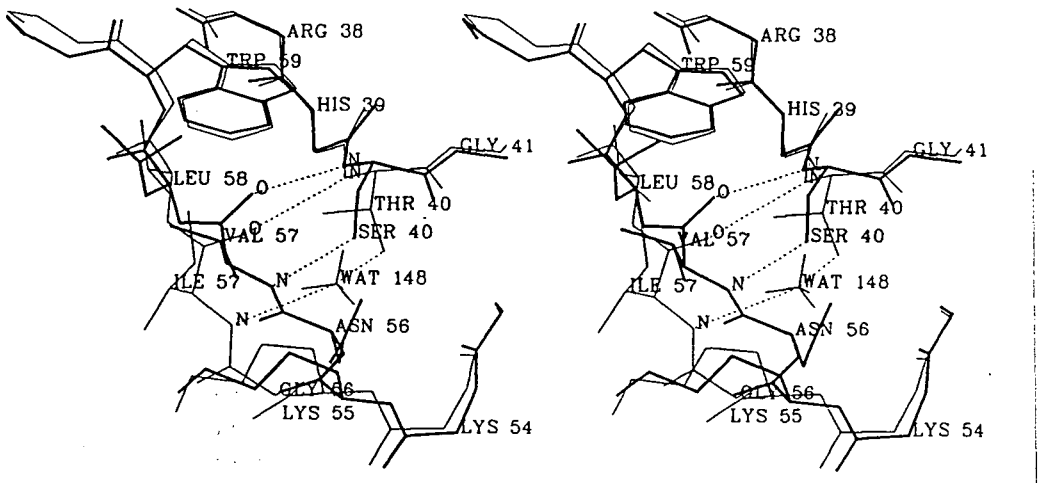

Figure IV.6. Differences in conformation between yeast iso-1- (thick lines) and tuna (thin lines) cytochromes *c* at residues 55 to 57. Residues 38 through 41, and 54 through 60 are shown for both proteins. Hydrogen bonds formed between these adjacent strands of polypeptide chain in the vicinity of residues 40 and 57 are shown as dashed lines. Note that the bridging Wat148 occurs only in the tuna protein.

of an Ile57 side chain would be positioned within 3.0 Å of the Trp59 indole ring. Thus, in tuna cytochrome *c*, the hydrogen bonds Ser40 N - Val57 O and Val57 N - Ser40 OG present in yeast cytochrome *c* cannot be formed, and instead Ile57 interacts with Thr40 through hydrogen bonds made via a bridging water molecule. Beyond this region, the conformation of the polypeptide chain is restored to that present in yeast and rice cytochromes *c*. This is made possible by Gly56 (asparagine and alanine in the yeast and rice proteins, respectively) adopting a conformation permitted only for glycine residues. The maintenance in all three proteins of an interaction between the side chain hydroxyl group of residue 40 and the main chain amide nitrogen of residue 57 is notable. Proudfoot *et al.* (1989) have shown that replacements of Lys, Val or Phe at position 40 destabilize horse cytochrome *c*, and suggest that a hydrogen bond formed by residue 40 is essential for maintaining the integrity of the bottom Ω-loop of the molecule.

The conformational differences observed about residues 23 to 25, and 56 to 57 emphasize the importance of glycines in providing to the polypeptide chain the necessary flexibility to accommodate the substitution of bulky side chains elsewhere in the cytochrome *c* molecule. This point is further emphasized by patterns of conservation of amino acid sequence in eukaryotic

cytochromes *c*. In the ~90 known primary sequences, where a residue with a side chain (usually alanine) occurs at position 24, it is almost always accompanied by a glycine at position 25; and where a threonine (instead of serine) occurs at position 40, and/or a residue larger than valine (most frequently isoleucine) occurs at position 57, a glycine is present at position 56.

Substantial differences in conformation are also observed at the polypeptide chain termini of yeast, tuna and rice cytochromes *c*. In all three cytochromes *c*, the N-terminus is only loosely associated with the rest of the molecule, and each N-terminal region has a unique conformation (Figure IV.2). The N-terminal extension characteristic of fungal and plant cytochromes *c* is positioned differently in yeast as compared to rice cytochrome *c*. In yeast cytochrome *c*, bulkier side chains at three positions (Phe vs. Ala-3, Glu vs. Pro88, and Asn vs. Ala92) result in the N-terminus being displaced further into the solvent medium, whereas in the rice protein this region is packed against the surface of the body of the molecule. As Figure IV.2 shows, the polypeptide chain conformations of the three proteins become equivalent only after residue 3. The C-terminus of tuna cytochrome *c* has a unique conformation due to the occurrence of a tryptophan residue at position 33. In yeast iso-1-cytochrome *c*, Glu103 OXT is hydrogen bonded to His33 NE2, whereas in tuna, the bulky Trp indole ring causes the C-terminal carboxylate to be rotated away from this side chain. That the N- and C-termini of these three cytochromes *c* differ greatly in conformation is indicative of their relatively unconstrained positioning on the surface of these molecules.

The unique conformation observed for tuna ferrocytochrome *c* at Lys27 is more difficult to rationalize. There are no amino acid substitutions in this region of the polypeptide chain with respect to the yeast protein, and in addition, in tuna ferricytochrome *c*, residue 27 has the same conformation as that in yeast iso-1-cytochrome *c*. It may be that whereas in the yeast protein, the strained conformation of residue 27 is stabilized by a Lys27 NZ - Leu15 O intramolecular hydrogen bond (see Figure III.7), in tuna ferrocytochrome *c*, a hydrogen bond between the side chain of Lys27 and an adjacent molecule in the crystal lattice serves to relieve this strained conformation. It should be noted that the occurrence of the same conformation of Lys27 in both

reduced (yeast iso-1-) and oxidized (tuna and rice) forms of cytochrome *c* brings into question the proposal of Rackovsky and Goldstein (1984) that this residue plays a role in redox-state dependent conformational switching.

2. Conserved amino acid side chains with variant conformations

Cytochrome *c* contains a large number of fairly conservative lysine residues. Of the 16 lysine residues present in yeast iso-1-cytochrome *c*, all but 4 also occur in at least one of tuna or rice cytochromes *c*. With the exception of Lys79, whose side chain is involved in forming a conserved hydrogen bond interaction between the left and right halves of the molecule (see Figure III.9), the side chains of the majority of these residues are not positioned similarly. The large average deviation (50.5°) in side chain torsion angles of the lysine side chains that are in common between yeast iso-1- and tuna or rice cytochrome *c* is not surprising, as these lysines are located on the surface of the molecule. Their mobility and lack of intramolecular hydrogen bond formation are consistent with their proposed role in the interaction with negatively charged groups of redox partners (Ferguson-Miller *et al.*, 1978).

Arg38 is invariant in the primary sequences of eukaryotic cytochromes *c* and is thought to act in stabilizing the buried, propionic acid group of pyrrole ring A. However, as Figure IV.7 shows, in each of yeast iso-1-, rice and tuna cytochromes *c*, the side chain of Arg38 has a distinct conformation and thus there exists a correspondingly different mode of interaction between the guanidinium group and the heme propionic acid. In yeast, the NH1 and NE atoms form hydrogen bonds to two water molecules that are in turn hydrogen bonded to the heme propionyl O1A atom (see Figure III.11a). In rice, the guanidinium group of Arg38 interacts end-on with the two bridging water molecules, with NH1 and NH2 donating hydrogen bonds. In tuna cytochrome *c*, Arg38 NE hydrogen bonds a lone bridging water molecule, while NH1 forms a direct hydrogen bond to the O1A heme atom. These differences appear to be due to amino acid substitutions in the local environment of Arg38 in the three cytochromes *c*. In yeast iso-1-cytochrome *c*, steric conflicts between the NH1 atom and the CG1 atom of Ile35, which replaces Leu35 occurring in tuna, prevent the guanidinium group of Arg38 from directly

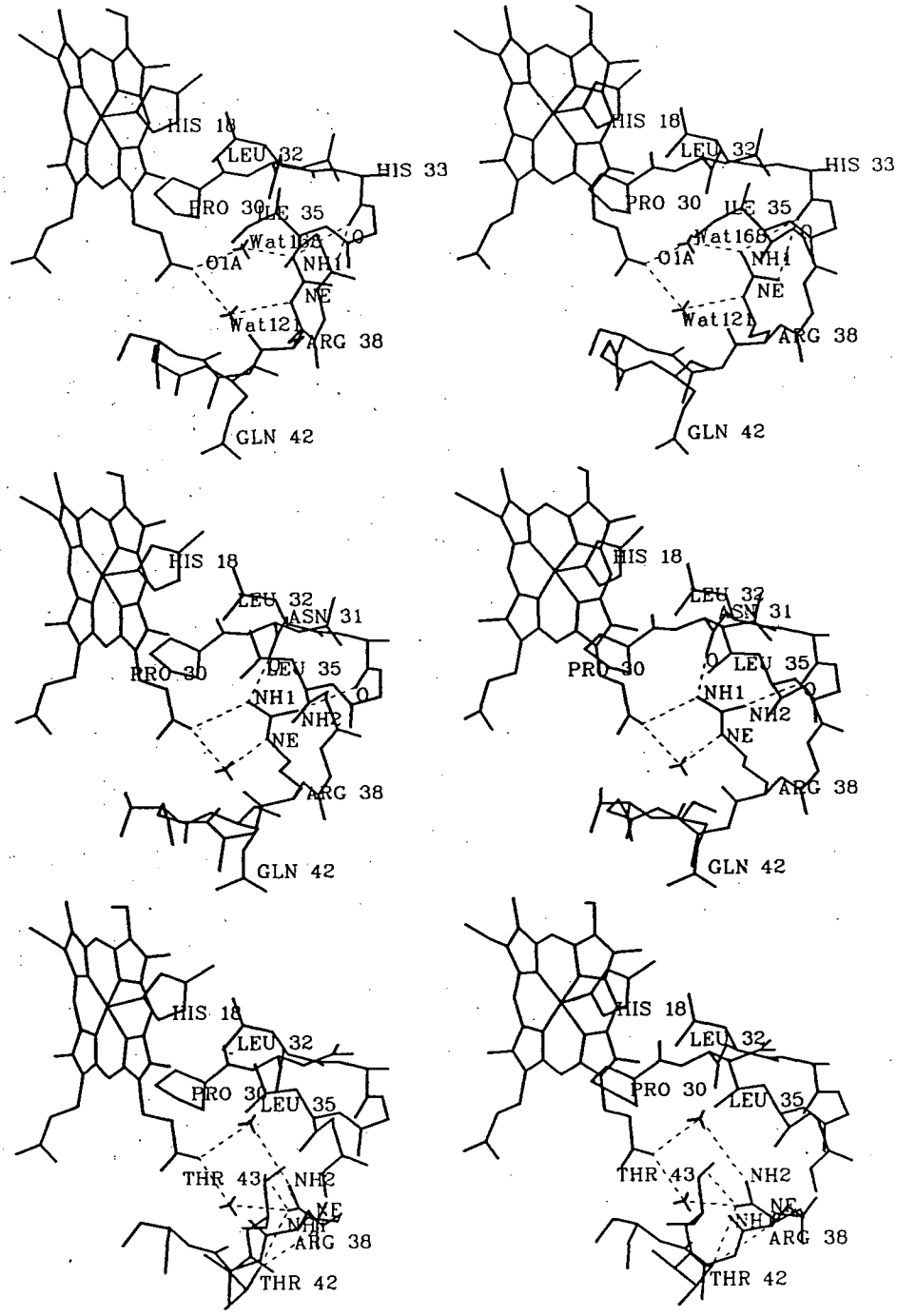

Figure IV.7. Comparison of the conformations of the Arg38 side chains in yeast iso-1- (*top*), tuna (*middle*) and rice (*bottom*) cytochromes *c*. Shown for each of the three protein molecules are Arg38, other protein groups in its direct vicinity, associated internal water molecules, and the hydrogen bonding network (dashed lines) involving its guanidinium group.

approaching the buried propionic acid group of the heme. In rice cytochrome *c*, the even more external positioning of the guanidinium group of Arg38 seems to be due to the unique hydrogen bonds formed between Arg38 NH2 and the OG1 atoms of two nearby surface residues, Thr42 and Thr43 (Gln42 and Ala43 in both yeast and tuna cytochromes *c*).

In the case of the side chain of invariant Arg91, the guanidinium group occupies a similar position in all three protein molecules, but the alkyl portion takes distinct paths (see Figure IV.8). Takano and Dickerson (1981a) have suggested that the role of Arg91 is to hydrogen bond and thereby stabilize the conformation of the fairly extended 80's region of the polypeptide chain. However, the different orientations of the Arg91 guanidinium group require this to be accomplished by different means in each molecule: in tuna cytochrome *c*, hydrogen bonds are formed between NE and Ile85 O, and NH2 and Lys86 O; in yeast between NH2 and both Leu85 O and Lys86 O; and in rice between NH1 and Lys86 O only. The conformation of the Arg91 side chain is influenced by nearby amino acid residues, particularly residue 65 (Met in tuna): in yeast, the NE atom is hydrogen bonded to Ser65 OG; and in rice cytochrome *c*, the guanidinium group is stacked against the aromatic ring of Tyr65. In addition, in tuna cytochrome *c* the side chain of

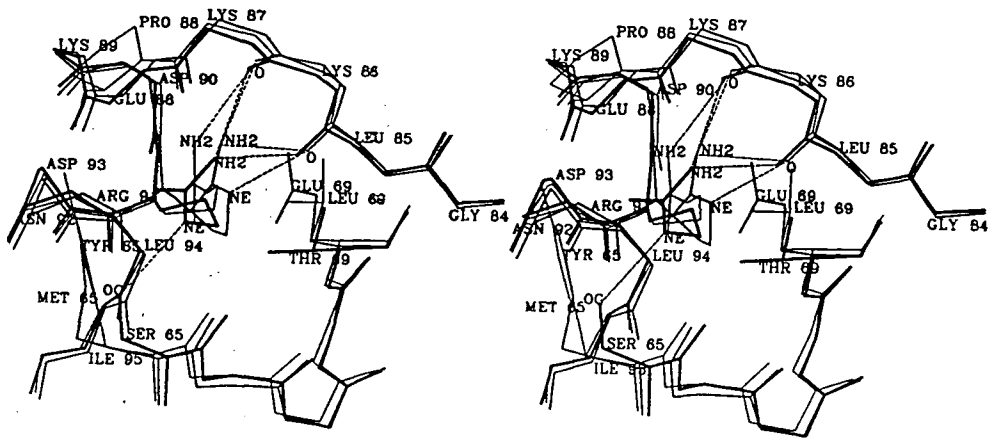

Figure IV.8. Comparison of the conformations of the Arg91 side chains in yeast iso-1- (thick lines), tuna (medium lines) and rice (thin lines) cytochromes *c*. Arg91, along with the polypeptide backbone in its direct vicinity and the amino acid side chains that appear to influence its conformation are shown for each molecule. Hydrogen bonds formed by the guanidinium group of Arg91 are shown as dashed lines.

Arg91 appears to be stacked against the side chain carboyxlate group of Glu69 (Thr and Leu in yeast iso-1- and rice cytochromes *c*, respectively). It may be of significance in rice cytochrome *c*, in which Arg91 forms only one hydrogen bond to the 80's region, that the presence of prolines at residues 83 and 88 could provide additional structural rigidity to the 80's loop. Interestingly, replacement of Ala83 by Pro in horse cytochrome *c* stabilizes the protein against both alkaline and thermal denaturation (Wallace *et al.*, 1989).

The conformation of the highly conserved residue Leu94 differs in the rice protein from that in yeast and tuna cytochrome *c* by a 120° rotation in the χ2 side chain torsion angle. In all three proteins, Leu94 is packed in the hydrophobic heme pocket, making close contacts with the side chains of residues 9, 68, 85 and 98 (all leucines in yeast iso-1-cytochrome *c*). The two conformations observed for Leu94 likely represent alternate means of achieving maximal packing efficiency. In yeast cytochrome *c*, both CD atoms of Leu94 are directed toward the CB carbon atom of Leu9. A similar conformation is observed in the tuna protein, in which the side chain of Leu9 is replaced by the CG1 atom of the shorter threonine side chain. However, in the rice protein, steric constraints arising from the replacement of Leu9 by an isoleucine results in an alternate conformation for Leu94, in which the CD2 side chain atom is packed against Leu98.

These results demonstrate that conserved or invariant residues do not necessarily carry out their role in an identical manner in different members of a protein family, and that there is flexibility in the formation of their intramolecular interactions. In addition, the intricate interplay between the conserved residue and compensatory changes nearby in the protein environment make difficult the assignment of definitive roles to conserved residues based on the inspection of just one member of the protein family. Nonetheless, the different modes by which a conserved residue can carry out its role provides clues as to the specific function of that residue.

C. Oxidation-state dependent conformational differences

With the completion of the high resolution structure of reduced yeast iso-1-cytochrome *c*, there are now available accurate structural data for two cytochromes *c* in the reduced state (yeast, tuna) and two in the oxidized state (tuna, rice). The conformational differences between tuna

ferro- and ferricytochromes *c* have been documented by Takano and Dickerson (1981a,b). The availability of the yeast and rice structures allows for a further assessment of oxidation-state dependent structural changes in cytochrome *c*.

Comparison of the superposed structures of all four proteins shows that the pattern of differences between the reduced and oxidized cytochromes *c* are similar to those observed considering only the two redox forms of tuna cytochrome *c*. Overall, the oxidation-state dependent structural changes are very small, with no significant alterations in main chain conformation. As observed by Takano and Dickerson (1981b), the largest changes occur at the lower left of the heme pocket (see Figure IV.9). These changes appear to be centred around an internal water molecule conserved in yeast, tuna and rice cytochromes *c* (Wat166 in yeast iso-1-cytochrome *c*; see Figure III.11 and Table IV.7). In the reduced yeast iso-1- and tuna

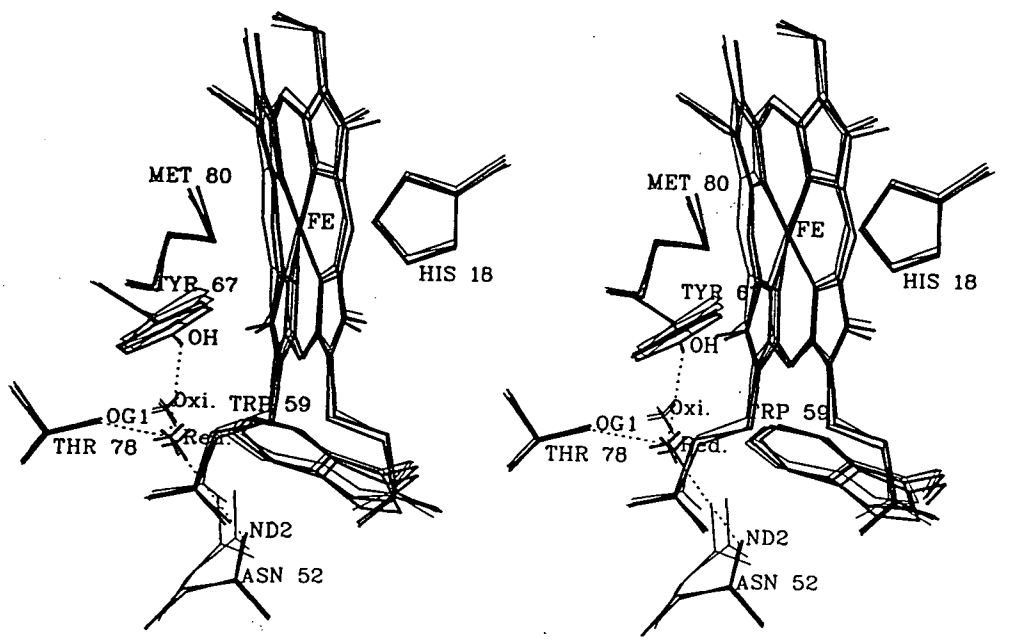

Figure IV.9. Oxidation-state dependent structural differences in cytochrome *c*. Comparison of the structures of reduced yeast iso-1- and tuna cytochromes *c* (thick lines) with those of oxidized rice and tuna cytochromes *c* (thin lines) shows that the most substantial structural changes accompanying oxidation/reduction occur in the vicinity of the buried water molecule located below Met80 in the heme crevice. Significant shifts also occur in the positions of two of the side chains, Tyr67 and Asn52 (Asp in rice cytochrome *c*), to which this water molecule is hydrogen bonded; and in the position of the side chain of Trp59.

cytochrome *c* structures, this water molecule is positioned ~6.5 Å from the heme iron atom. In the oxidized forms of tuna and rice cytochromes *c*, this water molecule has undergone a movement of about 0.9 Å and is positioned ~0.6 Å closer to the heme iron. The side chain of Asn52 (Asp in rice), which is hydrogen bonded to this water molecule, moves by almost 1 Å. The side chains of Tyr67 and Trp59 shift toward the rear of the molecule (by about 0.4 Å). This series of conformational changes likely occurs in response to, or alternatively in order to stabilize, the increased positive charge centred at the heme iron in the oxidized form of the cytochrome *c* molecule (Takano and Dickerson, 1981b; and see later discussions).

Other smaller shifts observed by Takano and Dickerson (1981b) are not unequivocally evident when the four cytochromes *c* are considered. At this point, it cannot be established whether such small conformational changes are less evident when yeast and rice cytochromes *c* are included in the comparisons, due to the greater overall structural dissimilarity arising from amino acid sequence differences; or the differences as observed between the two redox forms of tuna cytochrome *c* are larger than real, as a result of limitations in the accuracy of the structure determinations and refinement procedures. Work in progress on the high resolution structure of the oxidized form of yeast iso-1-cytochrome *c*, which crystallizes isomorphously with the reduced form, will provide additional insight on oxidation-state dependent conformational changes.

D. Crystal packing

Since those eukaryotic cytochromes *c* whose structures have been determined are structurally similar and all have been crystallized from concentrated ammonium sulfate solutions, the various crystal forms might be expected to exhibit related crystal packing. This has been shown to be the case for the oxidized forms of tuna and horse cytochromes *c* (Swanson *et al.*, 1977), which have 82% sequence identity. However, between cytochromes *c* possessing greater sequence divergence, the nature of the intermolecular contacts occurring in the crystalline forms are generally not conserved. A likely explanation for this poor conservation is that the precise packing contacts are in most cases made by the side chains of surface residues, which are frequently variant in both amino acid identity and spatial positioning. One notable aspect of crystal packing interactions that

does tend to be conserved in the various crystal forms is the specific regions on the surface of the cytochrome *c* molecule that participate in forming intermolecular contacts. In all crystal forms of cytochrome *c*, the segment of polypeptide chain from residues 54 to 60 is involved in tight intermolecular contacts, although the region of the symmetry-related molecule with which it interacts varies. Other regions of the cytochrome *c* molecule commonly involved in crystal contacts are centred at residues 8, 26 and 45.

E. Internal mobility

Crystallographic temperature factors can be used to compare the internal mobility of the yeast iso-1-cytochrome *c* molecule with that of tuna and rice cytochromes *c*. Figure IV.10 shows for each of the three cytochromes *c* the extent to which the B-factors of individual residues along the polypeptide chain differ from the overall average B of the protein molecule. Several observations can be made from a comparison of the three profiles. First, there is good agreement among the three cytochromes *c* of those regions of the polypeptide chain possessing markedly lower or higher flexibility than that of the adjacent polypeptide chain. The most rigid regions in rice and tuna cytochromes *c* are identical to those described above for yeast cytochrome *c* (see Section III.E). In addition, the three molecules have in common several regions of polypeptide chain, centred at residues 23, 38, 43, 55, 73, 83 and 103, which have a high degree of mobility. Nonetheless, differences in flexibility between the three cytochromes *c* are apparent at a number of locations in the polypeptide backbone. The most significant of these and possible explanations are summarized in Table IV.6. Localized differences in molecular flexibility between the three cytochromes *c* appear to be caused by two factors: the effects of amino acid substitutions and differences in crystal packing interactions. A comparison of Figures IV.10 and IV.2 also shows that the regions of the polypeptide chain at which the flexibility of yeast iso-1-cytochrome *c* differs from that of tuna or rice cytochrome *c* correspond to the regions at which yeast iso-1-cytochrome *c* differs in conformation from these other proteins. This result indicates that conserved intramolecular interactions are responsible for the observed similarities in molecular flexibility between the three cytochromes *c*.

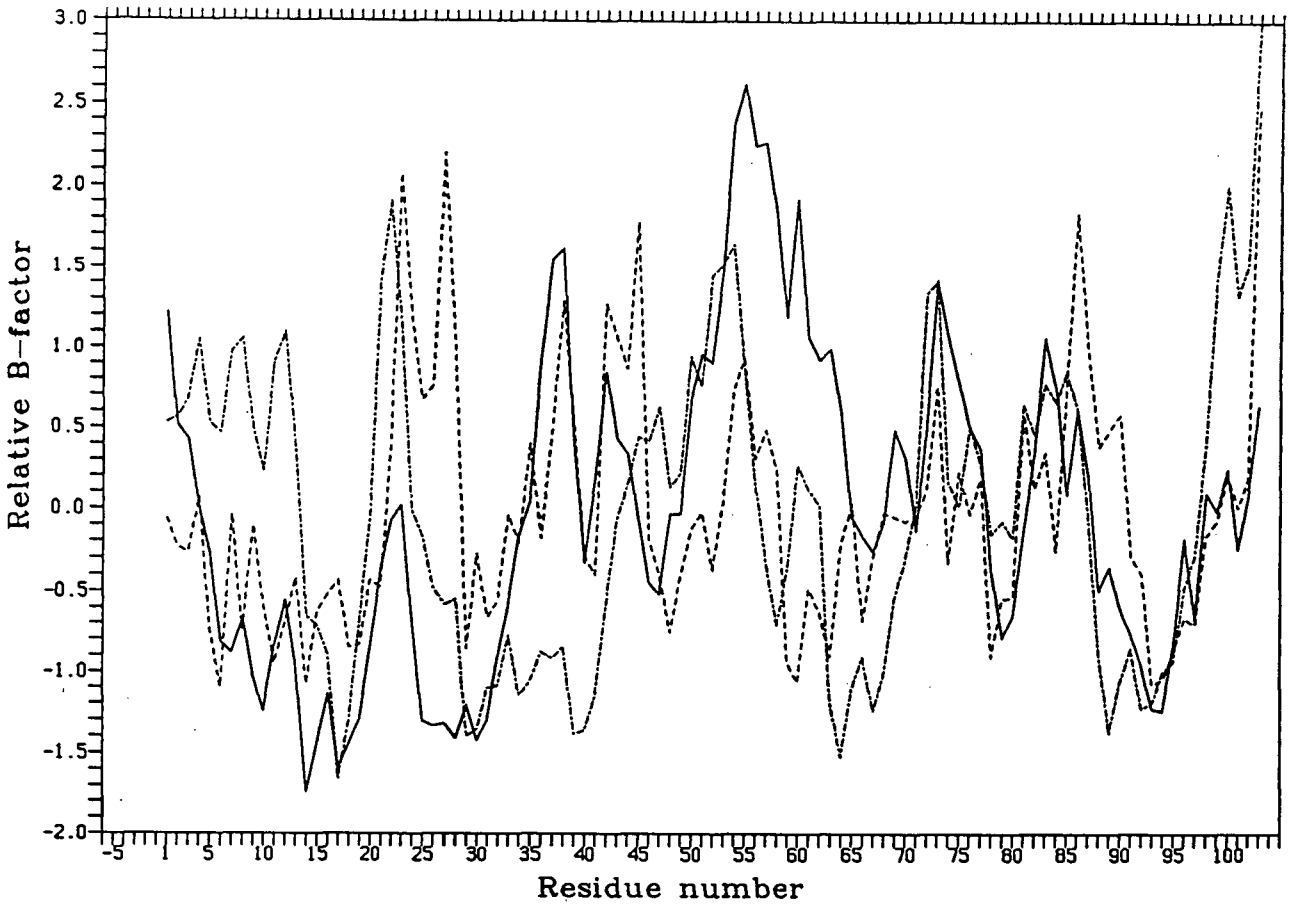

Figure IV.10. Comparison of the polypeptide backbone flexibilities of yeast iso-1- (——), tuna (– – – –) and rice (– · – · –) cytochromes *c*. The plots represent comparisons of the crystallographic temperature factors observed for common residues (1 to 103) in the three proteins. To facilitate comparisons, each protein was treated in the following manner: first, the overall mean, and standard deviation from the mean, of the temperature factors of all main chain atoms were determined; then, the overall mean value was subtracted from the average temperature factor of the main chain atoms of each residue, and these resultant values were divided by the standard deviation. The objective of this analysis is to correct for differences between the three cytochrome *c* structures in the degree of lattice disorder (Ringe and Petsko, 1985), as well as in the degree of variation in the temperature factor values in each structure, which may be affected by the tightness of restraints used during the respective refinements. The overall means and standard deviations for the main chain temperature factors of yeast iso-1-, tuna, and rice cytochromes *c* are 13.4 ± 5.3, 15.3 ± 3.9, and 9.6 ± 2.4, respectively.

Considering the good agreement between the temperature factor profiles determined for yeast iso-1-, tuna and rice cytochromes *c*, it is thus also of note that the results pertaining to the internal mobility of the cytochrome *c* molecule obtained crystallographically are consistent with those obtained using the methods of NMR (see previous discussion, Section III.E) and molecular dynamics. As discussed by Northrup *et al.* (1980, 1981), the r.m.s. atomic fluctuations of individual

Table IV.6. Differences in mobility between the polypeptide backbones of yeast iso-1-, tuna and rice cytochromes *c*

Mole-cule	Residues	Diff-erence	Possible explanation
Rice	5-13	high	Lacks lattice contacts, centred at residues 5 and 8, present in both yeast iso-1- and tuna cytochromes *c*. Also, hydrophilic side chains at Glu7 and Lys8 induce greater main chain mobility (Karplus and Schulz, 1985). In yeast iso-1-cytochrome *c*, rigidity of this helical segment may be further increased by the presence of two intra-helical hydrogen bonds formed by side chains of threonine residues.
Yeast	23-27	low	Pro25 adds rigidity. Also, lattice contact centred at residue 25 present.
Tuna	26-28	high	Altered conformation of this region of main chain, which makes few hydrogen bond interactions with rest of molecule.
Rice	34-39	low	Stabilized by crystal lattice contacts centred at residues 37 to 39.
Rice	42-44	low	Pro44 adds rigidity.
Tuna	50-53	low	Stabilized by crystal lattice contacts centred at residue 50.
Yeast	55-63	high	Flexibility induced by two pairs of adjacent residues with hydrophilic side chains (Lys54-Lys55, and Asp60-Glu61). Also lacks threonine side chain at position 63, which in tuna and rice cytochromes *c*, stabilizes the conformation of the start of an α-helix by forming an intrahelical hydrogen bond to the carbonyl group of residue 60.
Rice	64-67	low	Forms intramolecular interactions with extension at N-terminus of the polypeptide chain.
Rice	88-89	low	Pro88 adds rigidity.
Tuna	89-90	high	Lys88 and Gly89 induce increased main chain mobility.
Yeast	100-103	low	Position of C-terminus fixed by intramolecular salt bridge to His33 imidazole, and by crystal lattice interactions.

residues indicated by molecular dynamics calculations are in good agreement with X-ray diffraction results. Considering that the folds of yeast iso-1-, tuna and rice cytochromes *c* are very similar, it can be concluded that these results as a whole confirm that the dynamics of internal movements of cytochrome *c* are inherent in the overall three-dimensional structure of this protein molecule.

F. Internal packing efficiency

A structural feature notably conserved in yeast iso-1-, tuna and rice cytochromes *c* is a high packing efficiency in the heme pocket. Between these three cytochromes *c*, a number of

examples are observed in which compensating changes occur at a site spatially adjacent to the site of the introduction of a small side chain. These changes involve the substitution of a large side chain, or alternatively the altered conformation of a side chain. In the yeast protein, the shorter Ile35 side chain (leucine in tuna and rice) is compensated for by the longer side chain of Met64 (Leu64 in tuna and rice cytochrome *c*). In the tuna protein, neighboring the small side chain of Val95 is the internally oriented, hydrophobic side chain of Met65. In contrast, the side chains of Ser65 in yeast iso-1- and Tyr65 in rice cytochrome *c* are directed more into the external solvent medium, in order to accommodate an isoleucine residue at position 95. Two other examples, involving Arg38 and Leu94, were described earlier.

G. Solvent structure

1. Bound water molecules

A study was made to determine if the high degree of homology observed between the tertiary structures of yeast iso-1-, tuna and rice cytochromes *c* is reflected in a correspondingly high degree of homology between their respective solvent structures. Of the water molecules located in the present structure determination of yeast iso-1-cytochrome *c*, 28 were found to have counterparts in tuna cytochrome *c* and 25 in rice cytochrome *c*. However, only nine water molecules were found to occupy sites common to all three structures. Table IV.7 lists these conserved water molecules and the hydrogen bonding interactions formed by each. Table IV.7 also shows that the B-factors of the conserved water molecules are lower than average, undoubtedly reflecting the fact that these water molecules form a greater than average number of hydrogen bonds to the protein. These water molecules are likely to be of particular importance in stabilizing the folded state of the cytochrome *c* molecule.

The unexpectedly poor overall homology in the water structures of yeast iso-1-, tuna and rice cytochromes *c* prompted an inspection, in each individual protein molecule, of the sites corresponding to those occupied by a water molecule in one or both of the other protein molecules. Particular attention was paid to sites at which a water molecule is hydrogen bonded to a large number of main chain or invariant side chain groups, and to sites occupied by a water

Table IV.7. Conserved water molecules in yeast iso-1-, tuna and rice cytochromes *c*

Water	B-factor (Å^2)	Protein atoms involved in hydrogen bonding
113	21.7	Ser47 N
121	17.2	39 O, 42 N, Hem O1A, Arg38 guanidino
122	11.1	Lys79 O, 81 N
126	32.8	Gln16 OE1
160	24.3	Asp93 OD2
166	12.2	Asn(Asp)52 ND2(OD2), Tyr67 OH, Thr78 OG1
172	28.4	Asn70 OD1, Lys(Tml)72 N, Phe82 O
185	29.9	Tyr46 O, Tyr48 N
196	25.3	Phe36 N, 102 O

The water molecule designations and the temperature factors listed refer to the yeast iso-1-cytochrome *c* structure. The hydrogen bond interactions listed are those found to be homologous in all three cytochromes *c*. (For those residues designated only by residue numbers, the identity of the amino acid is not conserved in all of yeast iso-1-, tuna and rice cytochromes *c*.) The overall average B factor for these water molecules is 22.6 Å^2.

molecule in two of the three cytochromes *c*. It was found that in the majority of cases, the absence of a water molecule could be attributed to the occlusion of the binding site by either the altered conformation of an adjacent homologous amino acid side chain, or a substitution at a nearby residue in the protein molecule. Only in a few cases was the absence or presence of a water molecule influenced by packing interactions in the crystal lattice. An additional consideration in the low overall homology in solvent structure is the lesser extent to which the solvent structures of tuna and rice cytochromes *c* were characterized (see Table IV.1).

Of the four completely internal water molecules present in yeast iso-1-cytochrome *c* (see Section III.D.3), only Wat121 and 166 have counterparts in both tuna and rice cytochromes *c*. In the rice protein, Wat110 is absent because the altered conformation of the nearby β-turn affects the position of the carbonyl group of residue 25, which participates in coordinating this water molecule in yeast and tuna cytochromes *c* (see Figure IV.5). In tuna cytochrome *c*, because Arg38 interacts directly with the heme propionic acid group, the bridging Wat168 (see Figure IV.7) is absent. The overall average positional deviation of the two homologous internal water molecules is

0.49 Å.

2. Anion binding site

Yeast iso-1-, rice and tuna cytochromes *c* were all crystallized from concentrated ammonium sulfate solutions, and the start of the N-terminal α-helix in the latter two proteins has a conformation similar to that in yeast iso-1-cytochrome *c*. Nonetheless, the sulfate binding site observed in yeast iso-1-cytochrome *c* (see Section III.D.4) is not occupied by a sulfate ion in either the tuna or rice proteins. Structural explanations for this difference may be the replacement of Ser2 by a negatively charged aspartate in tuna cytochrome *c*, and the occurrence of a proline at residue 3 in rice cytochrome *c* which precludes the participation of the amide group of this residue in hydrogen bonding. A further consideration is that hydrogen bonding interactions to the fourth sulfate oxygen atom, which are provided through packing contacts in yeast iso-1-cytochrome *c*, may be required to sequester a sulfate ion.

V. SERINE 82 VARIANT OF YEAST ISO-1-CYTOCHROME *c*

A. Details of the structure determination

1. X-ray diffraction data

Diffraction data to 2.8 Å resolution, 2959 reflections in total, for the Ser82 variant of iso-1-cytochrome *c* were collected from a single crystal. Individual backgrounds were used in the processing of diffraction intensities. For this set of measured structure factors, a scale factor of 38.2 was indicated by a linear rescale against all common structure factors from the data set for wild-type iso-1-cytochrome *c*. The R-factor between the two data sets, $[\sum_{hkl} |F_o(\text{wild-type}) - F_o(\text{Ser82})|] / \sum_{hkl} |F_o(\text{wild-type})$, was 0.286.

2. Difference map

A difference map (Figure V.1) was calculated using the coefficients F_o(wild-type) - F_o(Ser82), and phases derived from a wild-type iso-1-cytochrome *c* atomic model which had been refined at 2.0 Å resolution to an R-factor of 0.199. This difference map confirmed the

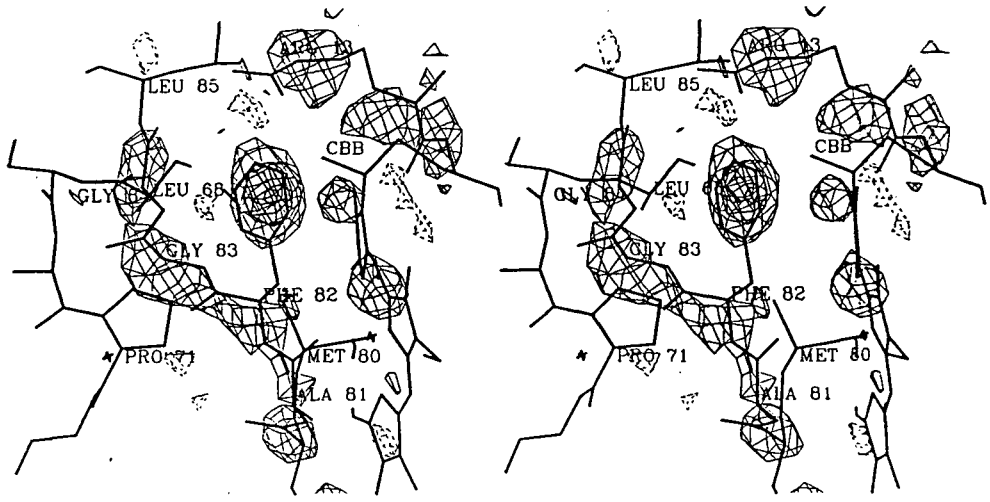

Figure V.1. Stereo drawing of the F_o(wild-type) - F_o(Ser82) difference map in the vicinity of residue 82. The atomic skeleton of wild-type iso-1-cytochrome *c* determined at 2.0 Å resolution is superimposed on the difference electron density envelopes. Positive density (solid lines) has been contoured at two levels, and negative density (dashed lines) at a single level.

replacement of Phe82, as only one large peak was observed, corresponding to a volume of positive density enveloping the position occupied by the phenyl ring of Phe82 in the wild-type molecule. Smaller positive peaks located adjacent to the polypeptide backbone of residues 81 to 84 indicated that in the Ser82 variant, small conformational rearrangements may occur in the vicinity of the substitution site. No other significant positive or negative peaks were evident.

3. Refinement of the structure of the Ser82 variant

Since the difference map indicated that conformational rearrangements were small and localized to a few specific regions of the iso-1-cytochrome *c* molecule, an initial model for the Ser82 variant was constructed from the 2.0 Å resolution wild-type structure. The phenylalanine side chain at position 82 was replaced with a serine side chain having the same $\chi 1$ torsion angle.

Table V.1. Agreement with ideal stereochemistry in the final refined model of the Ser82 variant of yeast iso-1-cytochrome *c* at 2.8 Å resolution

Class of restraint	R.m.s. deviation from ideality
1-2 bond distance	0.014 Å
1-3 angle distance	0.036 Å
1-4 planar distance	0.043 Å
Planar	0.014 Å
Chiral centre	0.160 Å^3
Non-bonded contact[a]	
single torsion	0.216 Å
multiple torsion	0.217 Å
possible hydrogen bond	0.231 Å
Staggered ($\pm 60^\circ$, 180°) torsion angle	28.8°
Planar torsion angle (0 or 180°)	2.8°
Temperature factor	
1-2 bond (main chain)	0.5 Å^2
1-3 angle (main chain)	0.8 Å^2
1-2 bond (side chain)	0.4 Å^2
1-3 angle (side chain)	0.8 Å^2

[a]The r.m.s. deviations from ideality for this class of restraint incorporate a reduction of 0.1 Å from the radius of each atom involved in a contact.

The original wild-type iso-1-cytochrome *c* model included 62 water molecules, which were incorporated into the initial refinement model for the Ser82 variant.

The Ser82 variant was refined against diffraction data with $F \geq 2\sigma_F$ in the resolution range 5.5 to 2.8 Å (1645 reflections in total, representing 75% of the available data). The course of refinement included two rounds of manual adjustments to the atomic model, the identification of an additional 10 water molecules, and 30 cycles of least-squares refinement. The manual adjustments required were minor. For the final refined structure of the Ser82 variant of iso-1-cytochrome *c*, the R-factor is 0.170, and the agreement with ideal stereochemistry is as shown in Table V.1.

B. Description of structural changes

The substitution of a serine at position 82 is accommodated without any gross conformational rearrangements in the native iso-1-cytochrome *c* fold. Figure V.2 shows a comparison of the Ser82 variant and the wild-type protein in terms of the average positional difference between main chain atoms and the differences in conformational torsion angles of corresponding residues. The overall r.m.s. difference between corresponding main chain atoms is 0.28 Å; the overall r.m.s. differences between phi and psi angles are 8.1° and 9.2°, respectively. The heme groups of the Ser82 variant and wild-type iso-1-cytochromes *c* do not differ significantly in either overall positioning or conformation, with the mean difference in atomic positions of the 43 heme atoms being 0.17 Å.

Figure V.2 shows that significant structural perturbations do occur in the direct vicinity of the amino acid substitution. In addition, the occurrence of a serine residue at position 82 gives rise to subtle conformational changes remote from the substitution site.

1. Vicinity of the substitution site

The conformational differences between the wild-type and Ser82 variant iso-1-cytochromes *c* occurring in the direct vicinity of the substitution site are shown in Figure V.3. The replacement of the large bulky phenylalanine side chain with a smaller serine side chain

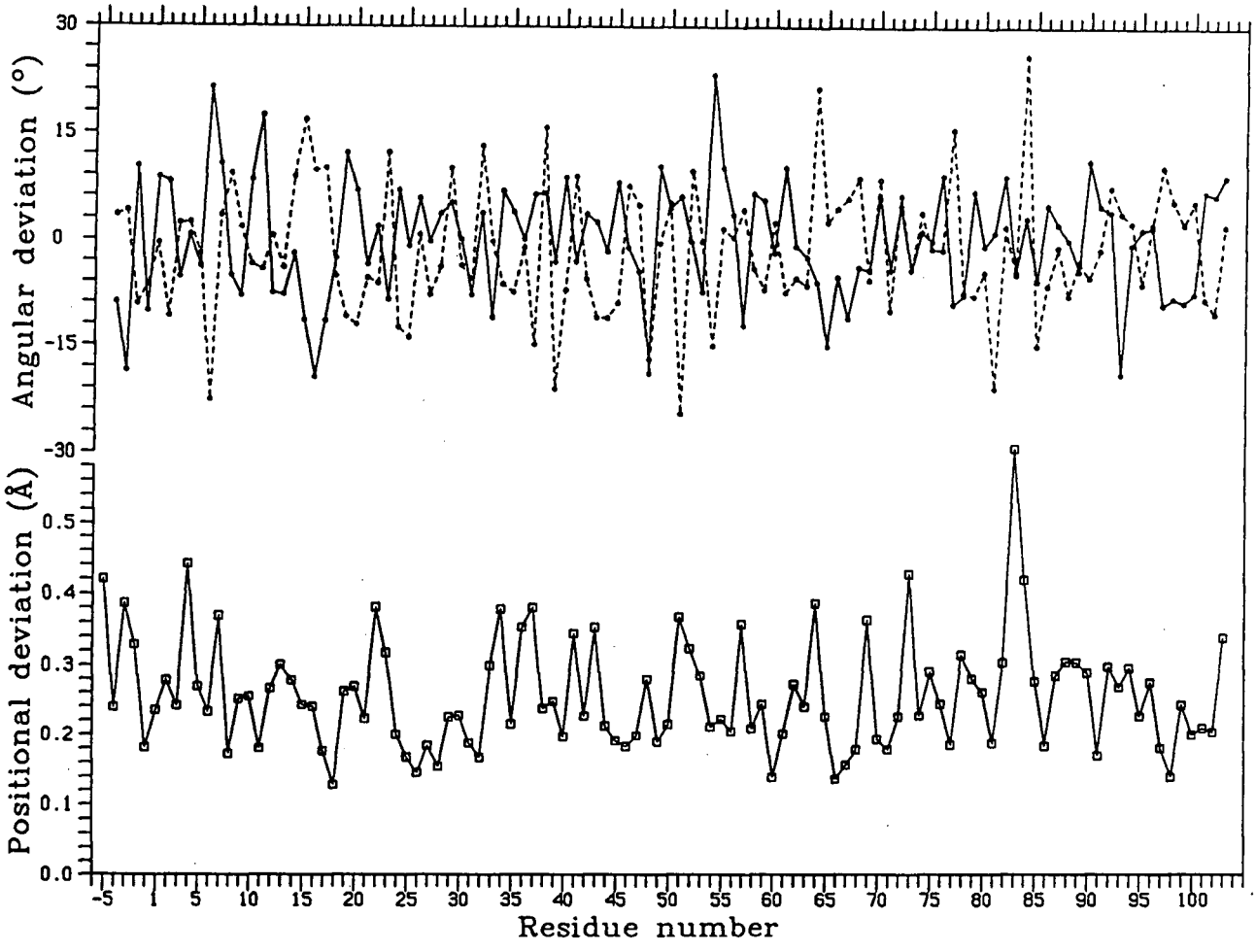

Figure V.2. Comparison of the polypeptide backbone conformations of the Ser82 variant and wild-type iso-1-cytochromes *c*. The lower plot indicates the average deviation observed between the positions of main chain atoms in corresponding residues of the two proteins. The upper plot indicates the deviations in the main chain conformational torsion angles ϕ (——) and ψ (---) of corresponding residues.

results in the polypeptide chain at residues 83 and 84, formerly packed against the left face of the phenyl group of Phe82, shifting downward and toward the pocket occupied by this group in the wild-type protein. This shift is mediated largely by alterations of $+21^\circ$ and -26° in the ψ angles of Ala81 and Gly84, respectively.

The position of the side chain of Arg13 also differs in the two proteins. In wild-type iso-1-cytochrome *c*, the alkyl portion of this side chain packs against the external edge of the phenyl ring of Phe82, and the guanidinium group is hydrogen bonded to water molecules. In the

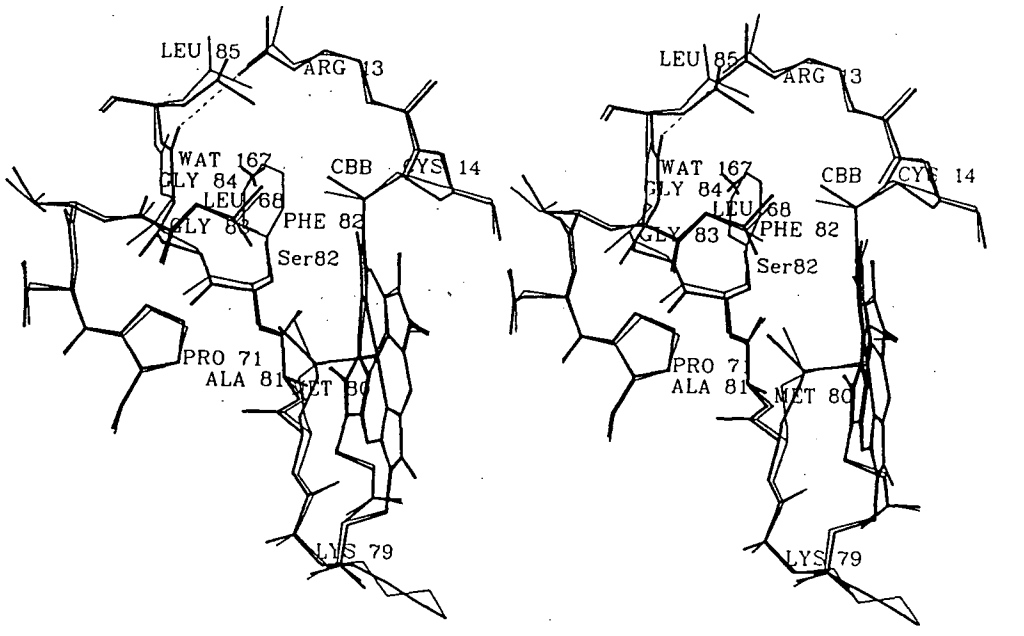

Figure V.3. Stereo drawing of the region around residue 82 in the Ser82 variant (thick lines) and wild-type (thin lines) iso-1-cytochromes *c*. The heme groups and the polypeptide chain of residues 13 through 14, 68 through 71, and 79 through 85 of the two proteins are shown superimposed. Two unique structural features in the Ser82 variant, the water molecule 167 hydrogen bonded to Ser82 OG, and the hydrogen bond between Arg13 NH1 and Gly84 O (dashed line), are also shown.

Ser82 variant, the guanidinium group of Arg13 is oriented to form a hydrogen bond with the carbonyl group of residue 84. This interaction is made possible by two factors. First, the smaller side chain of Ser82 allows the side chain of Arg13 to adopt a conformation which in the wild-type protein would generate prohibitive steric conflicts with the phenyl ring of Phe82. Second, the inward movement of residues 83 and 84 places the 84 O atom closer to Arg13. Indeed, it is possible that the observed shift in position of Gly84 is due in part to the hydrogen bond interaction formed with Arg13. That the side chain of Arg13 in the Ser82 variant forms an intramolecular interaction is reflected in its better defined appearance in electron density maps, and significantly lower atomic temperature factors (average 18 $Å^2$, as compared to 25 $Å^2$ in the wild-type protein).

2. Solvent accessibility to the heme crevice

In the wild-type protein, the phenyl ring of Phe82 plays an important structural role in excluding solvent from the heme pocket. Thus, in the Ser82 variant, the occurrence of the smaller serine side chain has created a channel allowing access of solvent down into the heme pocket (Figure V.4). The mouth of the channel is formed by the hydroxyl group of Ser82, the main chain of residues 83 and 84, the side chain of Leu85, and the CBB atom of the heme group, whereas the floor is formed by the side chain of Leu68 and the CE atom of the Met80 heme ligand. Although one water molecule (Wat167), hydrogen bonded to Ser82 OG, is positioned at the entrance to the channel, none are observed within the channel. This is consistent with the apolarity of the groups lining the internal portion of the channel. The exposure of the heme to the solvent medium is greatly increased in the Ser82 variant, with the total surface area directly accessible to solvent being 70.0 $Å^2$, as compared to 48.8 $Å^2$ in wild-type iso-1-cytochrome *c*.

In the Ser82 variant, the presence of the solvent channel exposes to solvent the CBB atom of the heme, and the side chains of Leu68 and Met80, three groups which are completely inaccessible to solvent in wild-type iso-1-cytochrome *c*. It is notable that in the Ser82 variant, these hydrophobic groups have shifted positionally, apparently in order to minimize their contact with solvent. The CBB heme atom has moved 0.8 Å toward a non-polar region at the rear of the molecule. This movement represents the only significant difference between atomic positions of the heme groups of the Ser82 variant and wild-type iso-1-cytochromes *c*, and is facilitated primarily by an adjustment in the conformation of the adjacent thioether linkage to Cys14. At the floor of the solvent channel, the Leu68 CD1 and CD2, and Met80 CE methyl groups have undergone shifts of 0.7, 1.2 and 0.7 Å, respectively, further down toward the interior of the iso-1-cytochrome *c* molecule. The exposure of these normally buried groups also further demonstrates the marked increase in solvent accessibility to the heme pocket that has occurred in the Ser82 variant.

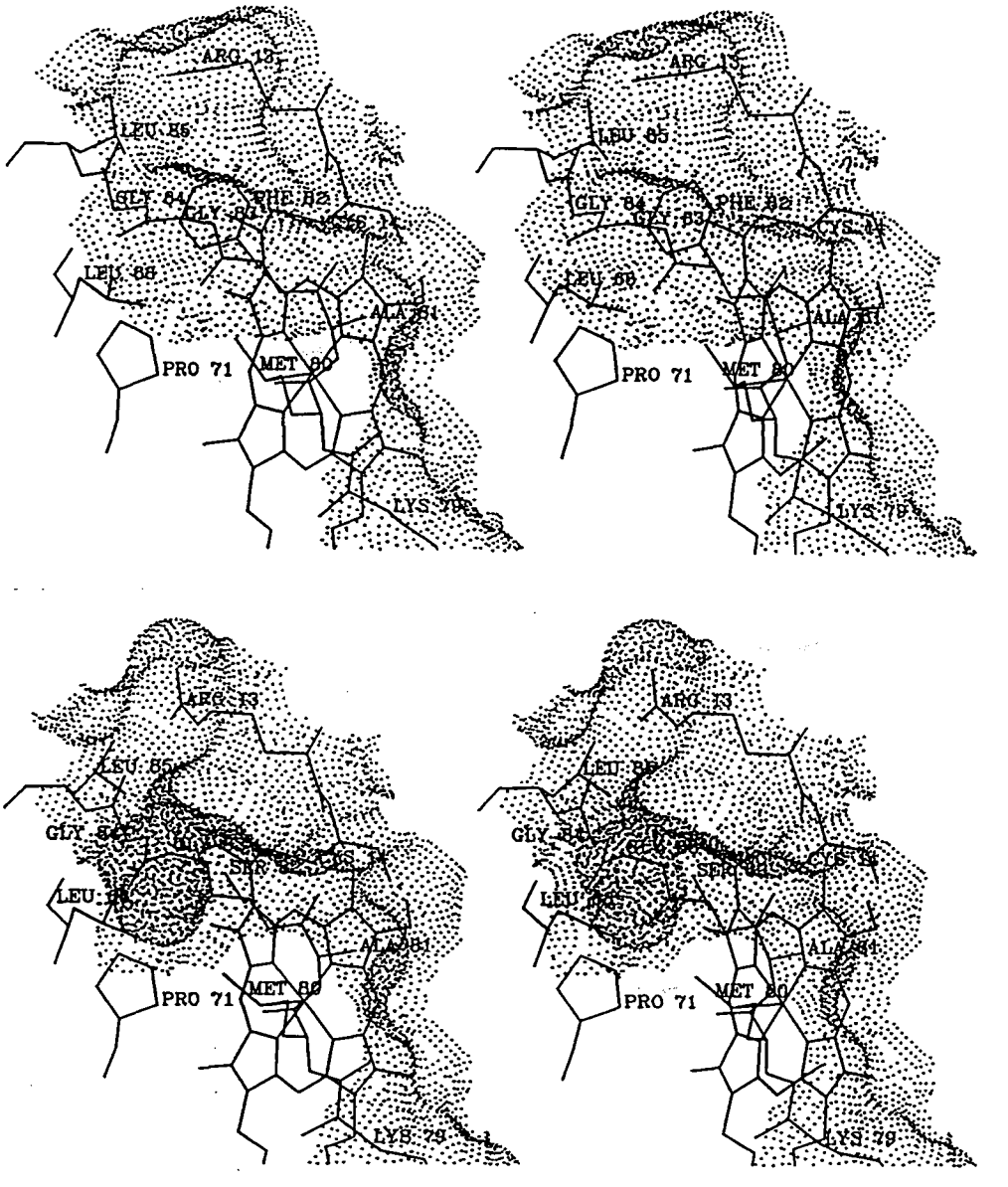

Figure V.4. Comparison of the molecular surface in the region of residue 82 in wild-type (*top*) and Ser82 variant (*bottom*) iso-1-cytochromes *c*. The dots outline those regions of the surface of the protein molecules which are accessible to solvent, and were calculated (Connolly, 1983) using a probe sphere of radius 1.4 Å. The directions of view are oblique to the Met80 face of the heme groups. In the Ser82 variant, a channel extends from the surface of the protein, immediately behind the main chain of residues Gly83 and Gly84, down to the side chains of Leu68 and Met80.

3. Remote from the substitution site

A number of conformational differences between the Ser82 variant and wild-type iso-1-cytochromes *c* occur around Trp59 and Wat166. These groups have positions near the

propionyl carboxyl group of the heme pyrrole ring A. This carboxyl group is located at the bottom of the heme pocket, and ~13 Å from the side chain of Ser82. The hydrogen bond interactions formed by this group in wild-type iso-1-cytochrome *c* are detailed in Figure III.11.

The position of the Trp59 side chain in the Ser82 variant differs with respect to that in the wild-type molecule by 0.75 Å overall, with shifts greater than 1 Å occurring for the CZ2, CZ3 and CH2 atoms (see Figure V.5). This difference is particularly notable considering the well-defined electron density for the aromatic rings in both structures, and the relatively small overall average shift (0.2 Å) observed for the side chain atoms of the other five aromatic residues occurring in the heme pocket. The movement of the Trp59 indole ring arises primarily from a ~15° rotation in the χ1 torsion angle of this side chain, and results in the loss of the hydrogen bond between the NE1 atom and the O2A propionic acid oxygen (the O2A-Trp59 NE1 distance is increased from 3.0 Å in the wild-type protein to 3.7 Å in the Ser82 variant). The propionic acid group of heme pyrrole ring A is otherwise unaffected by the shift in Trp59 and retains the positioning and hydrogen bonding interactions occurring in wild-type iso-1-cytochrome *c*. The movement of Trp59 also appears to induce the concerted movement of two adjacent side chains. The side chain of Met64, which is packed closely against the indole ring of Trp59, is displaced by ~0.7 Å. And the phenyl ring of Phe36, which is in turn packed adjacent to the side chain of Met64, undergoes a shift of ~0.5 Å. It is notable that a perturbation near 280 nm in the UV absorption spectrum of the Ser82 variant also suggests that Trp59 differs in conformation between the wild-type and Ser82 variant iso-1-cytochromes *c* (Pielak *et al.*, 1985, 1986).

The internal water molecule Wat166 and the side chain of Asn52, positioned on the frontward-facing side of the propionic acid group of pyrrole ring A, are also involved in a series of concerted movements (Figure V.5). In the Ser82 variant, Wat166 moves 1.4 Å in a direction almost directly toward the heme iron atom, thus decreasing the distance between these groups from 6.4 to 5.2 Å. Occurring in concert with this positional shift is the loss of the hydrogen bond between Wat166 and Asn52 ND2, and the movement of the side chain of Asn52 closer to the O1A atom of the heme propionic acid. The overall shift observed for the side chain atoms of

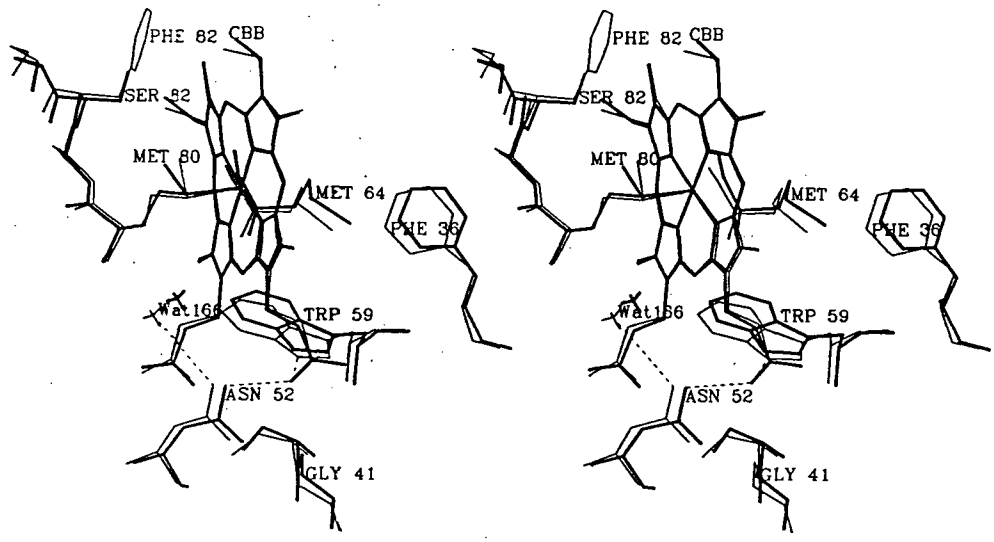

Figure V.5. Stereo drawing showing the structural differences between the Ser82 variant (thick lines) and wild-type (thin lines) iso-1-cytochromes *c* in the region around the heme propionic acid group of pyrrole ring A. The heme groups, the internal water molecule Wat166, and residues 36, 41, 52, 59, 64, and 80 through 83 of the two proteins are shown superimposed. Those hydrogen bonds which differ between the two proteins in this region of the molecule have been shown as dashed lines.

Asn52 is ~0.5 Å. In wild-type iso-1-cytochrome *c*, Wat166 forms additional hydrogen bonds to the side chains of Tyr67 and Thr78 (see Figure III.11). In the Ser82 variant, these hydrogen bonds are retained, and positional shifts are not observed for either of these latter residues.

C. Correlation of structural changes with differences in functional properties

1. Solvent accessibility to the heme pocket and polarity of the heme environment

a. Reduction potential

The reduction potential of the Ser82 variant of iso-1-cytochrome *c* is observed to be 255 mV, whereas that of the wild-type protein is 290 mV (Rafferty *et al.*, 1990). Structural differences between these proteins provide an explanation for the observed ~40 mV decrease in reduction potential of the Ser82 variant. The presence of the solvent channel in the Ser82 variant substantially increases the direct exposure of the heme group to solvent. In addition, the polarity of the heme environment of the Ser82 variant would be expected to be significantly increased by the accessibility of solvent to the heme pocket, the presence of a hydrophilic side chain at position

82, and the loss of the insulating phenyl group, which in the wild-type molecule shields the heme group from both the guanidinium group of Arg13 and the external solvent medium. Both inaccessibility of the heme group to solvent (Stellwagen, 1978), and non-polarity of the heme environment (Kassner, 1972) have been proposed to play significant roles in establishing a high reduction potential in cytochrome *c*. These results clearly demonstrate that Phe82 plays a major role in forming the heme environment of cytochrome *c*. Furthermore, they emphasize the significant influence of the polarity of the heme environment on the reduction potential of cytochrome *c*.

b. Reactivity of iso-1-cytochrome *c* with small molecule electron transfer reagents

Numerous small molecule redox reagents are able to carry out electron transfer reactions with cytochrome *c*. Experimental evidence indicates that these reactions occur via an outer sphere mechanism, and involve electron transfer occurring directly through the exposed edge of the heme group (reviewed in Ferguson-Miller *et al.*, 1979). Experiments measuring the rate of reaction of variants of iso-1-cytochrome *c* with the small molecule reductant $Fe(EDTA)^{2-}$ have shown that the Ser82 variant is reduced about twice as fast as the wild-type protein (14.8×10^4 vs. 7.2×10^4 M^{-1} s^{-1}) (Rafferty *et al.*, 1990). These results are consistent with the increased exposure to the solvent medium observed for the heme group of the Ser82 variant.

2. Stability of the heme crevice

At alkaline pH, cytochrome *c* is known to undergo a conformational change in which coordination of Met80 to the heme iron is lost (Pearce *et al.*, 1989). The pH at which this unfolding occurs provides a measure of the stability of the heme crevice (Osheroff *et al.*, 1980). The pK_a observed for the Ser82 variant (7.7) is significantly lower than that observed for the wild-type protein (8.5) (Pearce *et al.*, 1989), indicating that the stability of the heme crevice is decreased in the Ser82 variant. A decreased stability in the conformation of the polypeptide chain surrounding the heme is also inferred from the similiarity between the Soret CD spectrum of the Ser82 variant and that of oxidized iso-1-cytochrome *c* in which the integrity of the heme

environment has been disrupted by high temperature (Pielak *et al.*, 1986).

Pearce *et al.* (1989) have shown that of the two factors which determine the pK_a for the alkaline transition, it is the dissociation constant of the (unidentified) ionizing group that is most greatly affected by substitutions at Phe82. Somewhat unexpectedly, the equilibrium constant for the conformational rearrangement is little affected, and actually lowered in the case of the Ser82 variant. However, it is conceivable that an intramolecular interaction occurring in the native fold of the iso-1-cytochrome *c* molecule acts to maintain the pK_a of the ionizing group. In this case, factors which lower the stability of the native conformation could cause the observed shift in pK_a of the sensitive group. In the Ser82 variant, several structural perturbations occur that may effect such a lowering in stability. The replacement of the phenyl ring, which in the wild-type molecule is tightly packed within a hydrophobic pocket (see Section III.H.1), with a solvent channel gives rise to two destabilizing effects: a decrease in packing efficiency, and a disruption of the van der Waals forces maintaining the native conformation of the left wall of the heme pocket. In addition, both the exposure to solvent of normally buried, non-polar groups, and the introduction of the polar hydroxyl group of Ser82 into a hydrophobic environment would also be expected to destabilize the Ser82 variant. It is notable that these perturbations occur in a region of the molecule that has been proposed to play a role in stabilizing the heme crevice (Osheroff *et al.*, 1980).

3. Electron transfer activity

Phe82 has been proposed to play a direct role in the electron transfer reaction between yeast cytochrome *c* and yeast cytochrome *c* peroxidase. Liang *et al.* (1987, 1988) have determined that in the complex between iso-1-cytochrome *c* and zinc-substituted cytochrome *c* peroxidase, the rate of reduction of Zn^+-cytochrome *c* peroxidase by cytochrome *c* is 10^4 times slower for the Ser82 variant than for the wild-type protein.

These authors have attributed this lower electron transfer rate to the lack of an aromatic side chain at position 82. Interactions between the π-electron system of the aromatic group at this site and that of the heme group are proposed to account for the enhancement of electron

transfer occurring in the wild-type molecule. This interpretation is supported by the structural analysis of the Ser82 variant, which shows that structural perturbation at the putative electron transfer site is essentially restricted to a simple replacement of an aromatic group with a solvent channel. It is clear that in travelling from cytochrome *c* to cytochrome *c* peroxidase, an electron which might normally encounter the phenyl ring of Phe82 would encounter a substantially different medium in the Ser82 variant, where the intervening space would consist of a polar hydroxyl group, a solvent channel, and likely one or more entrapped solvent molecules (see Figure V.3). Indeed, it is possible that in the Ser82 variant, the altered character of this pathway results in electron transfer occurring via an entirely different mechanism (Marcus and Sutin, 1985; Liang *et al.*, 1987). Although the lack of the optimal medium for electron movement would appear to be the most obvious explanation for the impaired electron transfer activity of the Ser82 variant, it must also be considered that the other observed conformational differences or a perturbed electron distribution in the delocalized π-electron system of the heme (see discussion below) may contribute to the lowered ability of this protein to donate an electron to Zn^{+}-cytochrome *c* peroxidase. An additional note is that the absence of gross conformational differences between the wild-type and Ser82 variant iso-1-cytochromes *c* in the region of the molecule thought to interface with CCP suggests these two proteins can form a similar complex with CCP (Liang *et al.*, 1987, 1988). This is consistent with the similar rates observed for the photo-initiated electron transfer from ^{3}Zn-cytochrome *c* peroxidase to iso-1-cytochrome *c* for both the Ser82 variant and the wild-type protein (Liang *et al.*, 1987, 1988).

D. Possible explanation for remote conformational changes

In comparison to wild-type iso-1-cytochrome *c*, the Ser82 variant has a 50mV lesser reduction potential, which correlates with this protein's increased solvent accessibility to the heme and increased polarity in heme environment. The lower reduction potential indicates that the Ser82 variant has a decreased ability to stabilize the reduced state of the heme iron (i.e. a net neutral charge at the heme centre). This would be expected to create in the Ser82 variant, an increased electro-positive character at the heme iron, which may in turn be responsible for the

conformational changes occurring remote from the substitution site in this protein. Thus, the shifts observed at the internal water molecule Wat166 and at Asn52 may result from the attraction of Wat166 to the increased positive charge at the heme iron. The positioning of Trp59 may be affected by the electrostatic interaction between the heme iron and the propionic acid group of heme pyrrole ring A (Moore, 1983). In the Ser82 variant, the strengthened interaction between the electronegative carboxyl group and the increased electropositive charge on the heme iron may render the carboxyl group less able to act as an electron donor in a hydrogen bond with Trp59 NE1. The proposal that the conformational shifts observed in the Ser82 variant occur in response to increased electropositivity at the heme iron is supported by their strong similiarity to shifts observed between tuna ferri- and ferrocytochrome *c* (see Section IV.C). It can thus also be reasoned that the structural differences between the Ser82 variant and wild-type iso-1-cytochromes *c* represent the initial conformational changes occurring as reduced iso-1-cytochrome *c* is oxidized during an electron-transfer event.

An alternative explanation for the observed shift in position of Trp59 in the Ser82 variant arises from the proposal of Satterlee *et al.* (1987) that the effects of conformational changes occurring in cytochrome *c* in the vicinity of Phe82 can be transmitted, via a molecular orbital of the heme group, to the propionic acid group of pyrrole ring A. Shellnutt *et al.* (1979) have also emphasized the aromatic and electronegative groups occurring in the neighborhood of the heme are important determinants of the distribution of electron density within the heme π-orbitals. Thus in the Ser82 variant, it can be argued that the increased solvent accessibility and polarity occurring at the top of the heme pocket leads to a redistribution of electrons within the delocalized π-electron system of the heme. If this redistribution were to lower the electronegativity of the propionic acid group of pyrrole ring A, in a manner similar to that described above it may result in the loss or lengthening of the hydrogen bond to Trp59.

Considering that the propionic acid group of pyrrole ring A appears to be central to the conformational shifts occurring between the Ser82 variant and wild-type iso-1-cytochromes *c*, and that these shifts may represent the initial movements occurring during the oxidation of the

iso-1-cytochrome *c* molecule, it is particularly notable that Moore (1983) has proposed that this propionic acid group, through electrostatic interactions with the heme iron, is involved in mediating the oxidation-state dependent conformational changes in cytochrome *c*.

VI. GLYCINE 82 VARIANT OF YEAST ISO-1-CYTOCHROME *c*

Two proteins containing a glycine at position 82 in iso-1-cytochrome *c* were studied. In one, the naturally occurring Cys102 was present, whereas in the other this residue was replaced by a threonine. The two variants will be referred to as Gly82(Cys102) and Gly82(Thr102). Of particular significance is that the structure determinations were of the Gly82(Cys102) variant in its reduced state, and of the Gly82(Thr102) variant in its oxidized state.

A. Details of the structure determination

1. X-ray diffraction data

a. Gly82(Cys102) variant

For the Gly82(Cys102) variant of iso-1-cytochrome *c*, diffraction data to 2.6 Å resolution (3305 reflections in total) were collected from a single crystal. Individual backgrounds were used in the processing of diffraction intensities. The scale factor for the resultant set of structure factors was determined to be 44.6, using a linear rescale against all common structure factors from the data set for wild-type iso-1-cytochrome *c*. The R-factor between the two data sets was 0.415.

b. Gly82(Thr102) variant

Diffraction data for the Gly82(Thr102) variant were measured from two crystals, with one crystal (B, 8587 reflections) covering the resolution range ∞ to 2.0 Å and the other (C, 10733 reflections) ∞ to 1.76 Å. Individual backgrounds were used in the processing of both sets of diffraction intensities. The scale factors for the two sets of structure factors were determined, using the method of Thiessen and Levy (1973), to be 20.2 for B and 17.0 for C. The R-factor for structure factors in common between the data set for wild-type iso-1-cytochrome *c* and sets B and C of the Gly82(Thr102) variant were 0.27 and 0.28, respectively.

A complete data set to 1.76 Å resolution was obtained by merging the two independent sets of structure factors. The merging R-factor for the 6822 replicately measured reflections in the resolution range ∞ to 2.0 Å was 0.105.

2. Difference maps

For each of the Gly82 variants, difference maps (Figure VI.1) were calculated using the coefficients F_o(wild-type) - F_o(Gly82), and phases derived from a wild-type iso-1-cytochrome *c* atomic model which had been refined at 1.7 Å resolution to an R-factor of 0.197. The features in the two maps were similar but not identical. Interpreted with respect to the superimposed structure of wild-type iso-1-cytochrome *c*, both maps showed numerous significant electron density

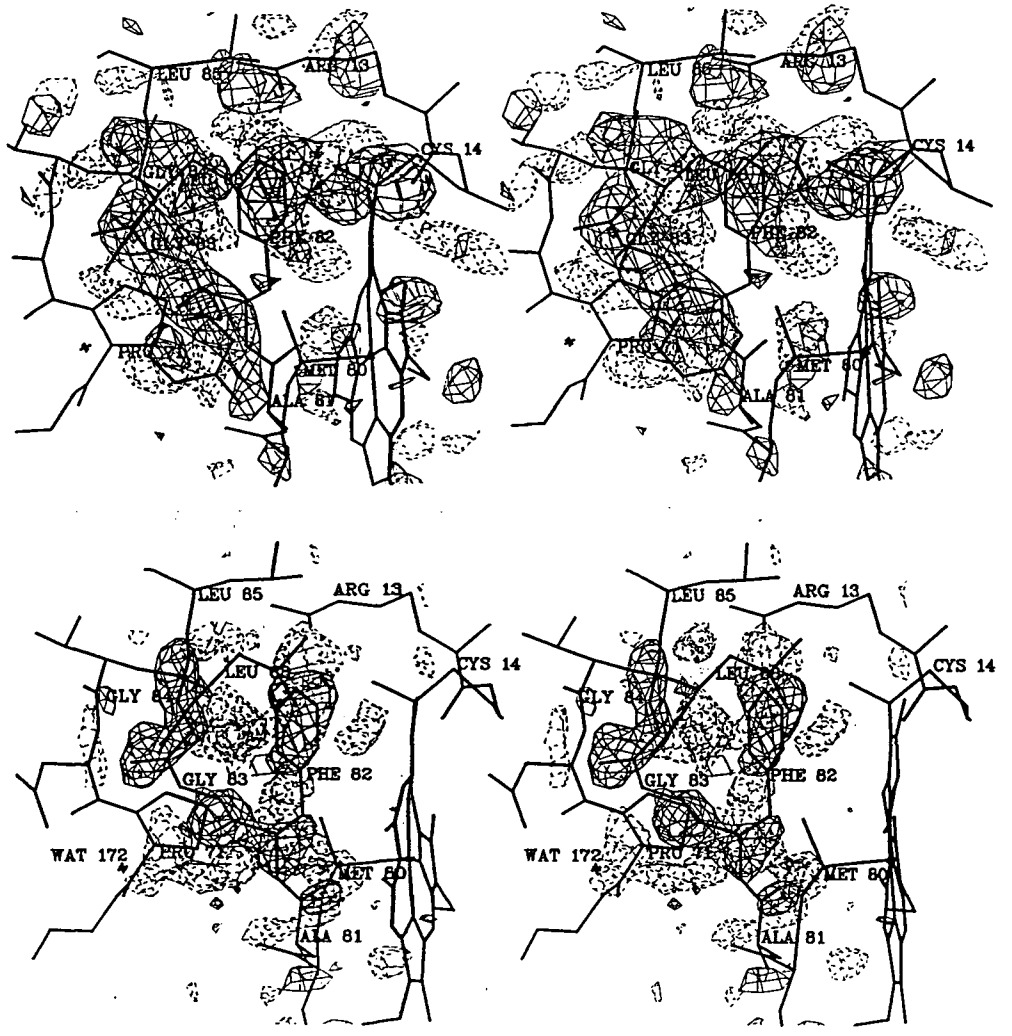

Figure VI.1. Stereo drawings of the F_o(wild-type) - F_o(Gly82) difference maps in the vicinity of residue 82 for the Gly82(Cys102) (*top*) and Gly82(Thr102) (*bottom*) variants of iso-1-cytochrome *c*. The atomic skeleton of wild-type iso-1-cytochrome *c* determined at 1.7 Å resolution is superimposed on the difference electron density envelopes. Positive density has been contoured with solid lines, and negative density with dashed lines.

peaks located around residues 82 through 84, in the direct vicinity of the substitution site.

a. Gly82(Cys102) variant

In the difference map for the Gly82(Cys102) variant, most of the phenyl ring of Phe82 was enveloped in positive density, as expected. However, positive density also occurred at the main chain of residues 82, 83 and 84, and there was significant negative density located between the phenyl ring of Phe82 and the backbone of residues 83 and 84. In addition, the Phe82 atoms CB and CG, as well as CD1 and CE1 of the inner edge of the phenyl ring, were associated with neither negative nor positive density. The difference map thus indicated that a conformational rearrangement of residues 82 to 84 occurs in the Gly82(Cys102) variant, but was not sufficiently clear to allow a refitting of this segment of polypeptide chain.

b. Gly82(Thr102) variant

The difference map for the Gly82(Thr102) variant very clearly indicated a perturbation in the conformation of residues 82 to 84. Enveloped in positive density were the upper portion of the phenyl ring of Phe82, as well as almost the entire stretch of polypeptide backbone from Ala81 O through Gly84 O. Note that a break at the CA atom of Gly83 is consistent with the later determination of a water molecule occurring at this site in the Gly82(Thr102) variant (see below). Strong peaks of negative density occurred adjacent to the carbonyl group of Ala81, directly above CA and CB of Phe82, and alongside the main chain of glycines 83 and 84. Consistent with the replacement of Cys102 by Thr102, the difference map also showed a strong positive peak at the SG atom of Cys102, and an adjacent negative peak.

The high quality of the difference map suggested that a $2F_o$(Gly82) - F_c' map may be sufficiently clear to define the altered conformation of residues 82 to 84 in the Gly82(Thr102) variant. The calculated structure factor magnitudes and phases for this map excluded the contribution of residues 81 through 85, and of several water molecules located in the vicinity of the substitution site. The resultant map was found to be very similar in appearance to both $2F_o$ - F_c and F_o - F_c' maps calculated later during the course of refinement of the

Gly82(Thr102) variant, and allowed the unambiguous refitting of the polypeptide chain from residues 81 through 84. In the refitting process, the main chain torsion angles of Met80 through Leu85 were adjusted. A $2F_o - F_c'$ map was also used to position the side chain of Thr102.

3. Refinement of the structures of the Gly82 variants

a. Gly82(Cys102) variant

The initial model used in the refinement of the Gly82(Cys102) variant was derived from a 1.7 Å resolution wild-type iso-1-cytochrome *c* structure, which included 98 water molecules and 1 sulfate ion. The atoms of the phenylalanine 82 side chain were deleted, but no other adjustments were made. The initial model was refined against diffraction data for the Gly82(Cys102) variant consisting of the 1581 structure factors with $F \geq 2\sigma_F$ in the resolution range 6.0 to 2.6 Å (corresponding to 55% of the available reflections). Nine cycles of refinement reduced the R-factor from a starting value of 0.288 to 0.206. At this point, the mean shift of all protein atoms from their positions in the initial model was 0.23 Å. Although structural refinement did not significantly alter the positions of residues 82 to 84, a $2F_o - F_c$ map calculated at this stage showed that these residues did not reside in electron density. The presence of significant positive electron density located between the heme group and the position then occupied by residues 82 to 84 indicated a possible alternative path for this segment of the polypeptide chain. A residue-deleted $F_o - F_c'$ map, where F_c' were calculated with residues 81 through 86 removed from the model, was consulted in making adjustments to the conformations of residues 81 through 85. Some minor manual adjustments were also made, through inspection of $2F_o - F_c$ maps, to the portion of the atomic model outside of the substitution site. Twelve cycles of refinement with the revised Gly82(Cys102) variant model then reduced the R-factor to 0.184. At this point, the fit of the refolded segment of polypeptide chain and the solvent structure in the vicinity of the substitution site were further evaluated using another $2F_o - F_c$ map. This led to the deletion of two water molecules hydrogen bonded to main chain atoms of residues 81 to 84 in their original conformations, and the addition of two water molecules which formed hydrogen bonds to these residues in their altered conformation. A final 18 cycles of refinement resulted in a further

reduction in R-factor to 0.158. As shown in Table VI.1, the final atomic model has excellent agreement with ideal stereochemistry.

b. Gly82(Thr102) variant

The initial model used in the refinement of the Gly82(Thr102) variant was also derived from a 1.7 Å resolution wild-type iso-1-cytochrome *c* structure. In this case, the preliminary model incorporated a number of adjustments made through inspection of $2F_o - F_c'$ maps as described above. These included a refitting of the segment of polypeptide chain Ala81-Gly82-Gly83-Gly84-Leu85, the placement of a threonine side chain at residue 102, and the deletion of several water molecules which occurred in the vicinity of the substitution site and which lacked

Table VI.1. Agreement with ideal stereochemistry in the final refined models of the Gly82 variants of yeast iso-1-cytochrome *c*

Class of restraint	R.m.s. deviation from ideality	
	Gly82(Cys102)	Gly82(Thr102)
1-2 bond distance	0.011 Å	0.024 Å
1-3 angle distance	0.028 Å	0.044 Å
1-4 planar distance	0.033 Å	0.060 Å
Planar	0.013 Å	0.020 Å
Chiral centre	0.155 Å^3	0.198 Å^3
Non-bonded contact[a]		
single torsion	0.206 Å	0.217 Å
multiple torsion	0.207 Å	0.221 Å
possible hydrogen bond	0.239 Å	0.244 Å
Staggered (±60°, 180°) torsion angle	23.2°	20.8°
Planar torsion angle (0 or 180°)	2.4°	3.4°
Temperature factor		
1-2 bond (main chain)	0.5 Å^2	1.5 Å^2
1-3 angle (main chain)	0.8 Å^2	2.1 Å^2
1-2 bond (side chain)	0.3 Å^2	1.8 Å^2
1-3 angle (side chain)	0.6 Å^2	2.7 Å^2

[a]The r.m.s. deviations from ideality for this class of restraint incorporate a reduction of 0.1 Å from the radius of each atom involved in a contact.

sufficient electron density. The course of refinement of the structure of the Gly82(Thr102) variant is summarized in Table VI.2. In general, the manual adjustments required were small, and were restricted to the side chains of amino acid residues. Further substantial manual adjustment to the conformation of the polypeptide chain in the vicinity of the substitution site was not required. The locations of four new water molecules occurring in this region of the molecule were determined. For the final refined structure of the Gly82(Thr102) variant of iso-1-cytochrome *c*, the R-factor is 0.199 for diffraction data in the resolution range 6.0–1.76 Å, and there is good agreement with ideal stereochemistry (see Table VI.1).

B. Description of structural changes

The replacement of phenylalanine at residue 82 with a glycine residue causes a substantial, but localized, refolding of the polypeptide chain. In both the Gly82(Cys102) and Gly82(Thr102)

Table VI.2. Course of refinement of the structure of the Gly82(Thr102) variant of yeast iso-1-cytochrome *c*

	Resolution range (Å)	Sigma cutoff	Number of reflections	Refinement cycles	R-factor Initial	R-factor Final	Inspection of atomic model
1	6.0 – 2.0	3.0	5199	13	0.262	0.214	Against $2F_o-F_c$ map and F_o-F_c' map for region near Gly82
2	6.0 – 2.0	3.0	5199	6	0.217	0.205	--
3	6.0 – 1.76	3.0	6176	11	0.221	0.209	Against $2F_o-F_c$ map
4	6.0 – 1.76	3.0	6176	11	0.216	0.203	--
5	6.0 – 2.0	2.0	6618	8	0.215	0.210	Against $2F_o-F_c$ map
	6.0 – 1.76	3.0					
6	6.0 – 2.0	2.0	6618	10	0.214	0.204	Against $2F_o-F_c$ map and
	6.0 – 1.76	3.0					F_o-F_c' map for region near Gly82
7	6.0 – 2.0	2.0	6618[a]	8	0.206	0.199	--
	6.0 – 1.76	3.0					

[a]The 6618 reflections used in the final refinement represents 69% of the available data in the resolution range 6.0–1.76 Å.

variants, the segment of polypeptide chain whose conformation is most greatly perturbed spans residues 81 through 85. However, the conformations of these residues in the two Gly82 variants differ largely not only from that in wild-type iso-1-cytochrome *c*, but also from each other.

The highly localized nature of the conformational rearrangement that occurs in the Gly82 variants is evident in Figure VI.2, which shows a comparison of the main chain conformations of the two Gly82 variants with that of wild-type iso-1-cytochrome *c*. Outside of the immediate

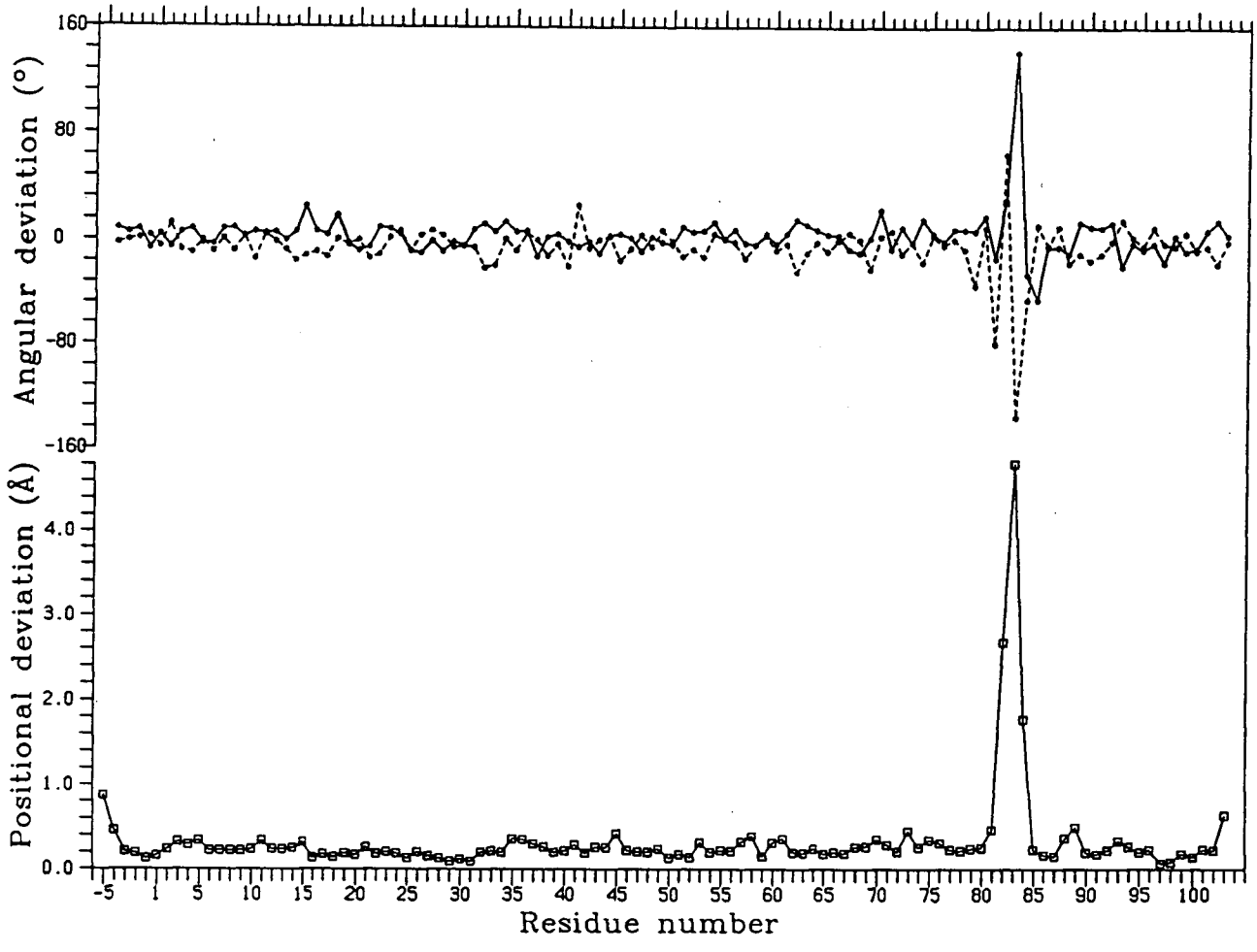

Figure VI.2. Comparison of the polypeptide backbone conformations of the Gly82 variant and wild-type iso-1-cytochromes *c*. Wild-type iso-1-cytochrome *c* is compared with the Gly82(Cys102) variant in the panel above, and with the Gly82(Thr102) variant on the panel on the following page. In each panel, the lower plot indicates the average deviation observed between the positions of main chain atoms in corresponding residues of the Gly82 variant and wild-type iso-1-cytochromes *c*. And the upper plot indicates the deviations in the main chain conformational torsion angles ϕ (—) and ψ (---) of corresponding residues of the two proteins.

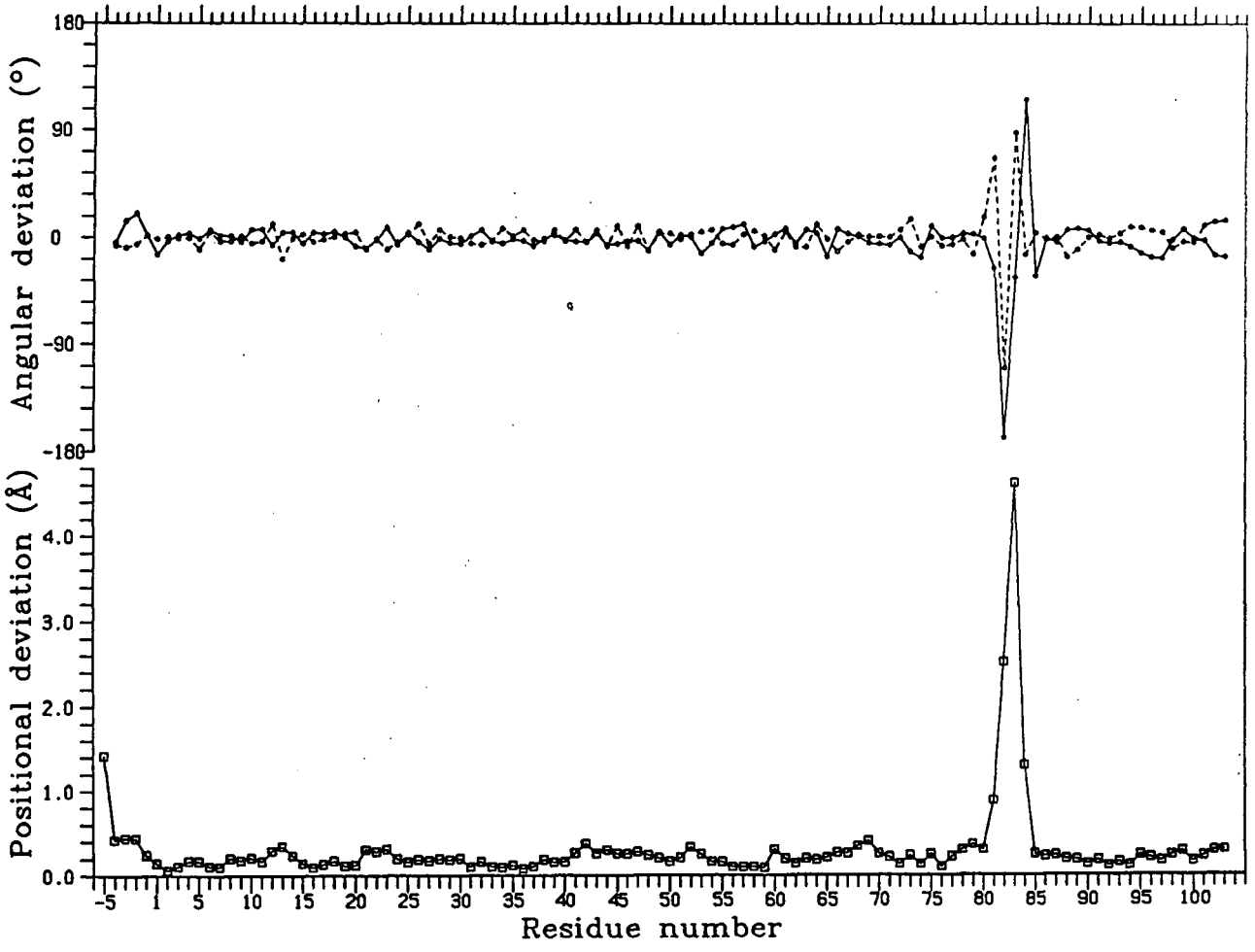

Figure VI.2 (continued).

vicinity of the substitution site, the main chain conformations of both Gly82 variants do not significantly differ from that of the wild-type protein. For the comparison of residues -5 to 80 and 86 to 103 of the wild-type protein with those of the Gly82(Cys102) variant, the overall r.m.s. positional difference for main chain atoms is 0.29 Å and the overall r.m.s. differences for phi and psi angles are 8.6° and 9.8°, respectively. The corresponding values for the comparison between the wild-type and Gly82(Thr102) variant iso-1-cytochromes *c* are 0.28 Å, 6.9°, and 6.9°, respectively.

1. Polypeptide chain refolding in the vicinity of Gly82

a. Gly82(Cys102) variant

In the refolding of the polypeptide chain in the Gly82(Cys102) variant (see Figure VI.3), Gly82, Gly83 and Gly84 are the residues most affected, with the main chain atoms of these residues shifting on average 2.7, 4.8 and 1.8 Å, respectively, with respect to their positions in the wild-type iso-1-cytochrome *c* molecule. On either side of the affected region, the large inward movement of residues 82 to 84 is accommodated by adjustments in only a single main chain torsion angle: ψ of Ala81 by -78°, and ϕ of Leu85 by -45°. It is notable that residues 82 to 84 do not undergo just a simple translational displacement, as the main chain torsion angles of these residues differ substantially from those in the wild-type protein (see Table VI.3).

The effect of the refolding is to pack the apolar portion of the main chain of residues 82 to 83 against the upper left face of the heme group. Van der Waals contacts made between this

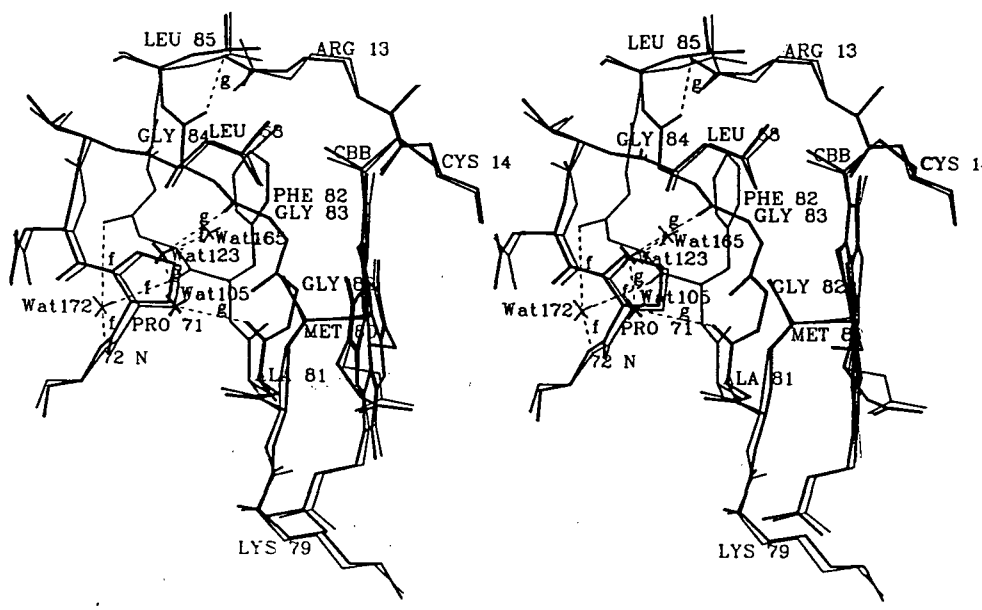

Figure VI.3. Stereo drawing of the region around residue 82 in the Gly82(Cys102) variant (thick lines) and wild-type (thin lines) iso-1-cytochromes *c*. The heme groups, the polypeptide chain of residues 13 through 14, 68 through 72, and 79 through 85, and neighboring water molecules of the two proteins are shown superimposed. Hydrogen bonds are represented as dashed lines, and have been labelled according to their occurrence in the wild-type ('f') or Gly82(Cys102) variant ('g') proteins.

Table VI.3. Comparison of main chain torsion angles[a] in the vicinity of residue 82 in wild-type and Gly82 variant iso-1-cytochromes *c*

Residue and main chain torsion angle		Wild-type	Gly82(Cys102)		Gly82(Thr102)	
			Value	Difference	Value	Difference
Met80	ϕ	-75	-56	+19	-75	0
	ψ	106	116	+10	125	+19
Ala81	ϕ	-82	-96	-14	-107	-25
	ψ	69	-9	-78	137	+68
Phe82	ϕ	-136	-105	+31	57	-167
(Gly82)	ψ	125	-170	+65	16	-109
Gly83	ϕ	-61	81	+142	-93	-32
	ψ	-45	179	136	43	+88
Gly84	ϕ	136	110	-26	-107	+117
	ψ	166	120	-46	153	-13
Leu85	ϕ	-113	-158	-45	-144	-31
	ψ	99	111	+12	105	+6

[a]All torsion angles have units of degrees.

segment of polypeptide chain and the carbon atoms of the heme group include Gly82 CA to Hem C2C and CMC (3.2 Å) and Gly83 CA to Hem CBB (3.5 Å). Interestingly, the carbonyl groups of residues 82, 83 and 84, the most polar groups (Burley and Petsko, 1986) in this segment of polypeptide chain, retain orientations similar to those occurring in the wild-type molecule, as they are all directed away from the heme group. It is notable that atoms of residues 82 through 84 have shifted from being over 6 to 8 Å removed from atoms of the heme group, to their positioning adjacent to the heme in the Gly82(Cys102) variant. In wild-type iso-1-cytochrome *c*, the phenyl ring of Phe82 has a similar placement with respect to the heme group. In the Gly82(Cys102) variant, the position of the polypeptide backbone at about Gly83 C coincides with that of the lower portion (atoms CD1 and CE1) of the Phe82 phenyl ring in the wild-type protein. It should also be noted that the large shift of the polypeptide chain in the Gly82(Cys102) variant at residues 82 to 84, and the positioning of these residues with respect to the position of the phenyl ring of Phe82 in the wild-type molecule are in good agreement with

the the initial F_o(wild-type) - F_o(Gly82) difference map.

b. Gly82(Thr102) variant

In the Gly82(Thr102) variant, refolding of the polypeptide chain most greatly affects the positions of Ala81 through Gly84 (see Figure VI.4). For residues 81, 82, 83 and 84, the average positional differences between corresponding main chain atoms of the Gly82(Thr102) variant and wild-type iso-1-cytochromes *c* are 0.9, 2.5, 4.7 and 1.3 Å, respectively. As in the Gly82(Cys102) variant, the large conformational rearrangement of residues in the vicinity of the substitution site perturbs on either side of the affected region only a single main chain torsion angle: ψ of Met80 by $+19^\circ$ and ϕ of Leu85 by -31° (see Table VI.3).

Similar to the situation in the Gly82(Cys102) variant, refolding in the Gly82(Thr102) variant positions the main chain of residues 82 to 83 in the volume of space occupied by the phenyl ring of Phe82 in wild-type iso-1-cytochrome *c*. The positions of the Gly83 N, CA and C atoms in the Gly82(Thr102) variant coincide closely with those of Phe82 CG, CD1 and CE1 in the wild-type molecule. In its position adjacent to the heme group, the polypeptide chain at residues

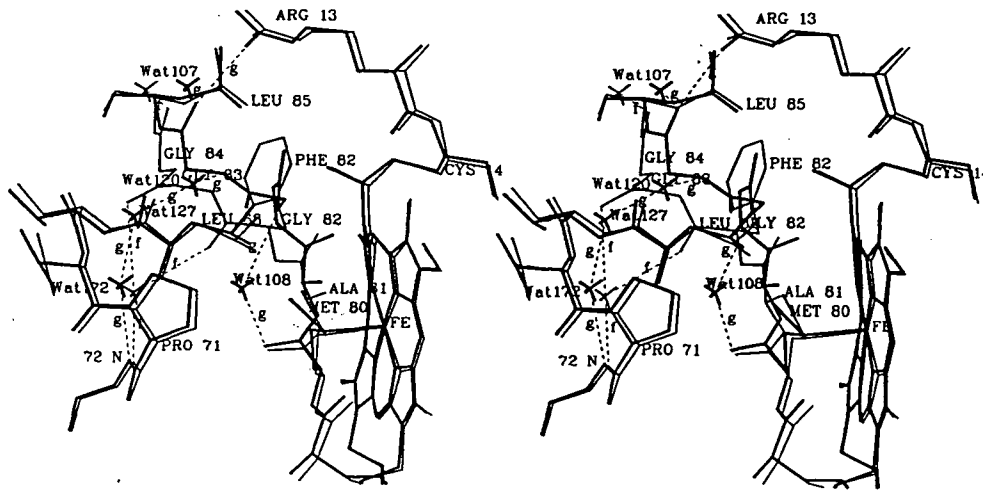

Figure VI.4. Stereo drawing of the region around residue 82 in the Gly82(Thr102) variant (thick lines) and wild-type (thin lines) iso-1-cytochromes *c*. The heme groups, the polypeptide chain of residues 13 through 14, 68 through 72, and 79 through 85, and neighboring water molecules in the two proteins are shown superimposed. Hydrogen bonds are represented as dashed lines, and have been labelled according to their occurrence in the wild-type ('f') or Gly82(Thr102) variant ('g') proteins.

81 through 83 makes several van der Waals contacts with atoms at the upper edge of this group. These contacts include Ala81 O to HemCBC (3.4 Å), Gly83 CA to HemCHC (3.5 Å), and Gly83 CA to HemCBB (3.2 Å).

2. Solvent structure in the vicinity of the substitution site

In the direct vicinity of the substitution site of both Gly82 variants, changes in the positioning of polar atoms of the polypeptide backbone cause the nearby solvent structure to be substantially altered from that in the wild-type protein.

a. Gly82(Cys102) variant

A water molecule (Wat172) which in the wild-type structure is hydrogen bonded to the carbonyl oxygens of both residues 82 and 83, as well as to the backbone amide nitrogen of residue 72, is absent in the Gly82(Cys102) variant. Refolding of the polypeptide chain at the substitution site places both carbonyl groups too distant from 72 N to participate with this atom in coordinating a water molecule. A water molecule (Wat165) hydrogen bonded to the amide nitrogen of Gly83 in the wild-type molecule is also absent, since Gly83 N in the Gly82(Cys102) variant is directed inwards into the heme pocket, rather than upward toward the external solvent. The carbonyl groups of residues 81 and 83 at their altered positions in the Gly82(Cys102) variant hydrogen bond two water molecules (Wat105 and Wat123) which have no counterparts in wild-type iso-1-cytochrome *c*.

b. Gly82(Thr102) variant

A number of water molecules bound to residues 81 to 84 in their refolded conformation in the Gly82(Thr102) variant are unique to this protein. Wat108 forms hydrogen bonds with both the carbonyl group of Met80 and the amide nitrogen of Gly83, and is notable in that it is positioned within 3.5 Å of the Met80 side chain, which ligands the heme iron. Wat120 is hydrogen bonded to the amide nitrogen of Gly84. A further unique water molecule (Wat127) is hydrogen bonded to Wat120. In addition, the inward shift of the carbonyl group of Gly84 in the Gly82(Thr102) variant causes a corresponding positional shift in Wat107, which is hydrogen bonded

to this carbonyl group. Interestingly, Wat172 which is absent in the Gly82(Cys102) variant does occur in the Gly82(Thr102) variant. In the latter protein, Wat120 participates with the amide nitrogen of residue 72 in coordinating Wat172.

3. Other structural differences

a. Gly82(Cys102) variant

Occurring nearby to the site of the amino acid substitution in the Gly82(Cys102) variant are slight alterations in the conformations of the side chain of arginine 13 and of the thioether bridge between Cys14 and the heme group. A reorientation of the guanidinium group of Arg13, coupled with the more inward positioning of the carbonyl group of Gly84, results in the formation of a hydrogen bond between these two groups. In the wild-type molecule, van der Waals contacts between the phenyl ring of Phe82 and the aliphatic portion of the side chain of Arg13 prevent the guanidinium group from approaching 84 O closely enough to form a hydrogen bond. The CBB methyl group, which is part of the heme vinyl side chain forming a thioether bond to Cys14, is packed against the phenyl ring of Phe82 in wild-type iso-1-cytochrome *c*. In the Gly82(Cys102) variant, movement of this group by 0.8 Å further into the interior of the heme pocket occurs to accommodate the positioning of Gly83 adjacent to the heme group (the CBB - Gly83 CA distance is 3.5 Å).

Overall, the heme conformations in the Gly82(Cys102) variant and wild-type iso-1-cytochromes *c* are similar, with the mean difference in the positions of all heme atoms being 0.26 Å. Aside from the CBB atom, the only other significant positional difference occurs at the aliphatic portion of the propionic acid group of pyrrole ring A. The largest positional shift of 1.1 Å occurs at the CBA carbon atom, and is associated with a smaller shift (average 0.4 Å) in the carboxyl group of this propionic acid.

Finally, a positional shift is observed in the Gly82(Cys102) variant at an internal water molecule (Wat166) located within the lower left of the heme pocket (Figure VI.5). In wild-type iso-1-cytochrome *c*, Wat166 forms hydrogen bonds to the side chains of Asn52, Tyr67 and Thr78. The position of this water molecule in the Gly82(Cys102) variant differs by 1.2 Å from that in

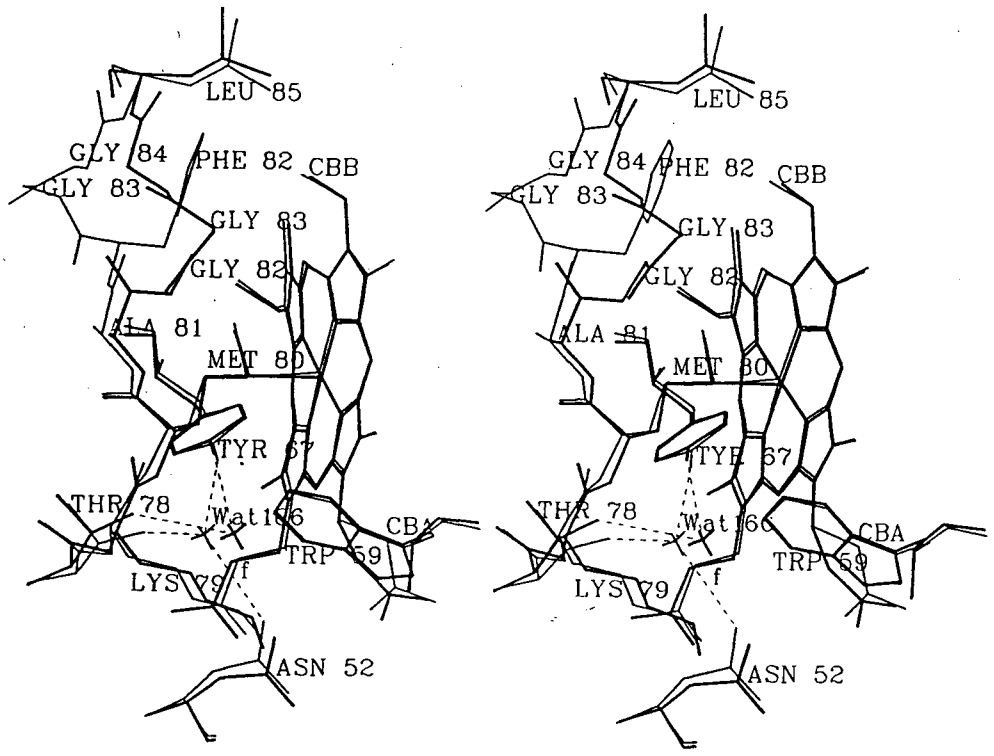

Figure VI.5. Stereo drawing showing structural differences between the Gly82(Cys102) variant (thick lines) and wild-type (thin lines) iso-1-cytochromes *c* in the region around the internal water molecule Wat166. The heme groups, the internal water molecule Wat166, and residues 52, 67, 78 and 79 through 85 of the two proteins are shown superimposed. Hydrogen bonds involving Wat166 are shown as dashed lines, with the Wat166–Asn52 hydrogen bond present only in the wild-type molecule labelled 'f'.

the wild-type protein. In the Gly82(Cys102) variant, Wat166 is positioned 0.8 Å closer to the heme iron (5.6 Å vs. 6.4 Å in the wild-type structure) and no longer forms a hydrogen bond to Asn52.

b. Gly82(Thr102) variant

Apart from the refolding of the polypeptide chain about Gly82, remarkably little dissimilarity exists between the structures of the Gly82(Thr102) variant and wild-type iso-1-cytochromes *c*. The heme groups in these two proteins have essentially identical conformations and positioning, with the mean deviation in the positions of all heme atoms being only 0.15 Å.

i. Vicinity of the substitution site

As in the Gly82(Cys102) variant, a slight adjustment in the conformation of the side chain of Arg13 occurs in the Gly82(Thr102) variant, allowing the guanidinium group of Arg13 to form a hydrogen bond to the carbonyl group of Gly84.

ii. Oxidation-state dependent structural changes

Structural differences between the wild-type and Gly82(Thr102) variant iso-1-cytochromes *c* may result from two factors: the phenylalanine to glycine amino acid substitution, and the difference in oxidation states of the two proteins. Due to their strong similarity to differences that occur between the two redox forms of cytochrome *c* (see Section IV.C), the differences described in this section have been considered to be dependent on the oxidation state of the iso-1-cytochrome *c* molecule. However, it has already been demonstrated that conformational shifts that have occurred remote from the substitution site in the reduced form of the Ser82 variant resemble those expected as iso-1-cytochrome *c* is oxidized during an electron transfer event (see Section V.D). Thus, in interpreting structural differences between the wild-type and Gly82(Thr102) variant iso-1-cytochromes *c*, the two above-mentioned effects are particularly difficult to differentiate. In the present case, it would seem most likely that both factors contribute to the observed differences.

The most substantial positional shifts occur at Wat166, and at the side chain of Asn52 (see Figure VI.6). In the Gly82(Thr102) variant, Wat166 is displaced by 0.6 Å with respect to its position in wild-type iso-1-cytochrome *c*. This displacement decreases the distance between Wat166 and the heme iron from 6.4 Å to 5.9 Å, and also appears to induce the upward shift (average deviation 0.8 Å) of the side chain of Asn52, to which Wat166 is hydrogen bonded. The Wat166-Asn52 ND2 hydrogen bond distances in the two structures are similar [2.9 Å in the Gly82(Thr102) variant, and 3.1 Å in wild-type iso-1-cytochrome *c*]. Significant conformational differences between the Gly82(Thr102) variant and wild-type iso-1-cytochromes *c* are also observed at two other side chains. In the Gly82(Thr102) variant, the aromatic rings of Trp59 and Tyr67 have shifted approximately 0.5 Å and 0.3 Å, respectively, in a direction toward the rear of

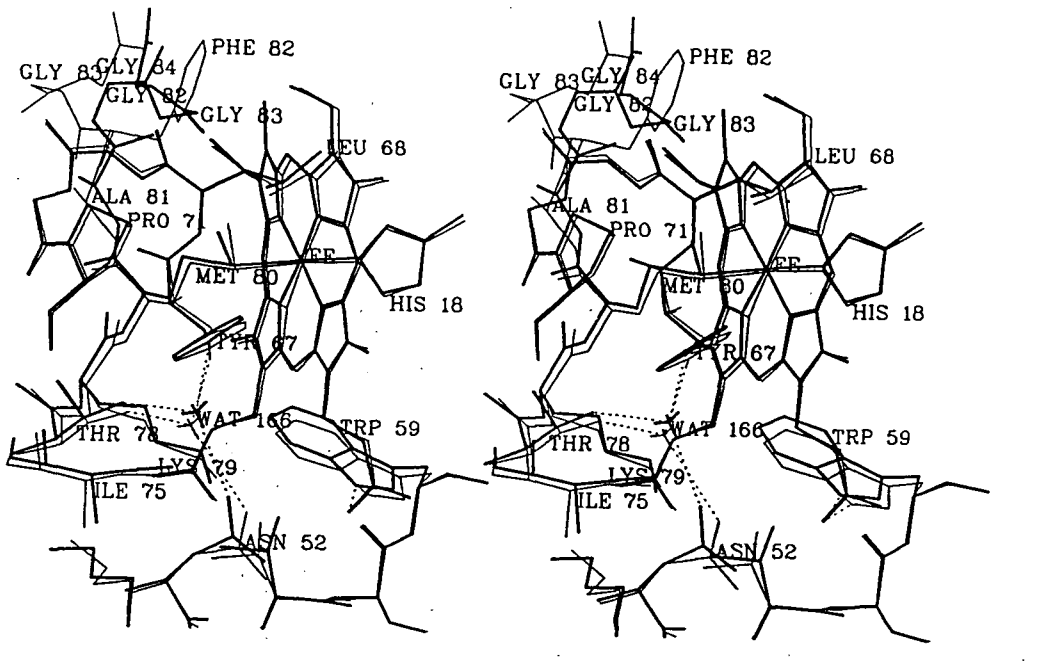

Figure VI.6. Oxidation-state dependent structural differences between the oxidized Gly82(Thr102) variant (thick lines) and reduced wild-type (thin lines) iso-1-cytochromes *c*. The heme groups, the internal water molecule Wat166, and residues 52 through 59, 67 through 71, and 75 through 84 of the two proteins are shown superimposed. Oxidation-state dependent changes are restricted to the side chains of Trp59, Asn52 and Tyr67 and the water molecule Wat166. Note the strong resemblance to conformational differences observed between reduced yeast and tuna, and oxidized tuna and rice cytochromes *c* (see Figure IV.9). Hydrogen bonds involving groups whose positions are sensitive to the oxidation state have been drawn as dashed lines. The large rearrangement occurring in the vicinity of the Gly82 substitution site can also be seen in this figure.

the iso-1-cytochrome *c* molecule.

iii. Threonine 102

In the Gly82(Thr102) variant, the replacement of the naturally occurring cysteine at residue 102 with a threonine causes no significant structural perturbation (see Figure VI.2b). The side chain of this threonine residue adopts a g^- conformation ($\chi 1 = 71^\circ$), which is the same as that adopted by Thr102 in both tuna and rice cytochromes *c*. In iso-1-cytochrome *c*, the Thr102 OG1 atom forms a hydrogen bond to the carbonyl group of Leu98. A homologous hydrogen bond (Cys102 SG - Leu98 O) occurs in the wild-type protein, but it is notable that the Cys102 SG and Thr102 OG1 atoms do not occupy the same position. Instead, the position of the CG2 atom of the Thr102 side chain in the Gly82(Thr102) variant coincides with that of the SG atom of

Cys102 in the wild-type protein, which is consistent with the non-polarity of the pocket occupied by these two atoms (see Section III.G.4). It is also of note that the close proximity (~4.3 Å) of the methyl groups of Thr102 CG2 and Val20 CG1 is consistent with NMR studies (Pielak *et al.*, 1988b), which show a strong NOE between protons of these two groups.

4. Comparison of the refolded conformations in the Gly82(Cys102) and Gly82(Thr102) variants

The refolding occurring in both the Gly82(Cys102) and Gly82(Thr102) variants results in a final conformation of the polypeptide chain which differs between the two proteins (see Figure VI.7). The most notable conformational differences are the opposite orientations of the peptide bonds between residues 81 and 82, and residues 83 and 84. The largest positional discrepancy between the two Gly82 variants occurs at Gly82. In the Gly82(Cys102) variant, Gly82 is positioned deeper in the heme pocket, and is packed closer to the heme plane; in the Gly82(Thr102) variant Gly82 has a more external positioning.

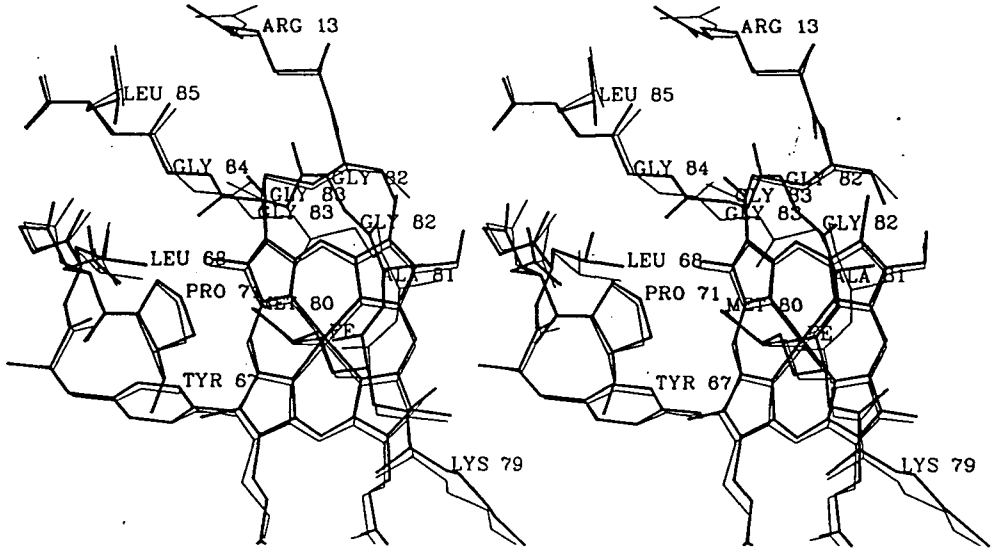

Figure VI.7. Comparison of the conformations of the polypeptide chain around residue 82 in the Gly82(Cys102) (thin lines) and Gly82(Thr102) (thick lines) variant iso-1-cytochromes *c*. The heme groups and the polypeptide chain of residues 13 through 14, 67 through 71, and 79 through 85 of the two proteins are shown superimposed.

Unclear are the causes of the differing conformation of the polypeptide chain at residues 81 to 84 in the two Gly82 variants. Two obvious differences between these proteins are their oxidation states and the identity of residue 102. However, it would seem unlikely that these factors are of primary importance. In general, redox-state dependent conformational changes in cytochrome *c* are observed to be small, and appear to be localized to the lower left of the heme pocket (see Sections I.B.4 and IV.C). In addition, a preliminary difference map for the reduced Gly82(Thr102) variant indicates that the conformation of residues 81 to 84 in this protein resembles that in the oxidized Gly82(Thr102) variant. The cysteine to threonine replacement at position 102 has not significantly affected the conformation of the polypeptide chain local to this substitution site, and residue 102 is 18 Å removed from the region of the polypeptide chain around residue 82. It may be that the observed conformations for residues 81 to 84 simply represent two relatively stable members of a population of existent conformations.

It should be emphasized that despite the differences in the conformation of residues 81 through 84, the two Gly82 variants are similar in that this segment of polypeptide chain has undergone a large inward shift with respect to its position in the wild-type protein. The result is that the degree of exclusion of solvent from the heme pocket is about the same in the two proteins. Also similar is the number of polar groups from the polypeptide chain which are introduced into the heme pocket. Therefore, in much of the subsequent discussions, interpretations of altered functional properties of the Gly82 variant in terms of structural changes apply equally to both the Gly82(Cys102) and Gly82(Thr102) variants.

C. Correlation of structural changes with differences in functional properties

1. Accessibility of the heme pocket to solvent

The important role played by the phenyl ring of Phe82 in forming the upper left wall of the heme pocket and in excluding solvent from the heme pocket has been discussed earlier (see Sections III.H.1 and V.C.1). The modelled structure of iso-1-cytochrome *c* in which the side chain of residue 82 is removed, but the native conformation of the polypeptide backbone is retained, shows that the absence of the bulky phenyl ring causes a large part of the upper left

face of the heme group as well as several hydrophobic groups forming the heme pocket (including the Met80 ligand) to become exposed to solvent.

However, in the Gly82 variants, the repositioning of the main chain of residues 82 to 83 into the space normally occupied by the phenyl ring of Phe82 results in the heme groups having a degree of solvent accessibility approximately the same as that of the heme group in wild-type iso-1-cytochrome *c* (see Figure VI.8). The solvent accessibilities of the heme group in both the Gly82(Cys102) and Gly82(Thr102) variants (44.4 $Å^2$ and 44.5 $Å^2$, respectively) are actually slightly lower than that in wild-type iso-1-cytochrome *c* (48.8 $Å^2$). That the Gly82 variant maintains a high degree of exclusion of solvent from the heme crevice is in contrast to the situation observed in the Ser82 variant. In that variant, no refolding of the polypeptide chain occurs, and the presence of the smaller serine side chain results in the creation of a channel extending down into the heme pocket (see Section V.C.1) and consequently in a greatly increased solvent accessibility to the heme group (70.0 $Å^2$).

The Gly82 variant reacts with the reductant $Fe(EDTA)^{2-}$ at twice the rate as does wild-type iso-1-cytochrome *c* (13.7×10^4 vs. 7.2×10^4 $M^{-1}s^{-1}$). The Ser82 variant shares a similar two-fold increase in reactivity (14.8×10^4 $M^{-1}s^{-1}$), which can be correlated with the substantially increased exposure to the solvent medium of the heme group in this protein. A simple structural explanation is not apparent for the Gly82 variant, in which heme solvent accessibility is about the same as that in the wild-type protein. It may be that the inward folding of residues 81 to 84 permits the closer approach of the $Fe(EDTA)^{2-}$ reactant to the heme group of cytochrome *c*, or that the flexibility of this segment of polypeptide chain (see below) permits transient unfolding of this region and thus transiently increased heme exposure.

2. Reduction potential and the polarity of the heme environment

The reduction potential of the Gly82 variant is measured to be 247 mV, whereas that of the wild-type protein is 290 mV (Rafferty *et al.*, 1990). The ~50 mV decrease in reduction potential observed for this variant is similar to that observed for the Ser82 variant. For the Ser82 variant, the lowering in reduction potential can be largely attributed to the increased polarity of

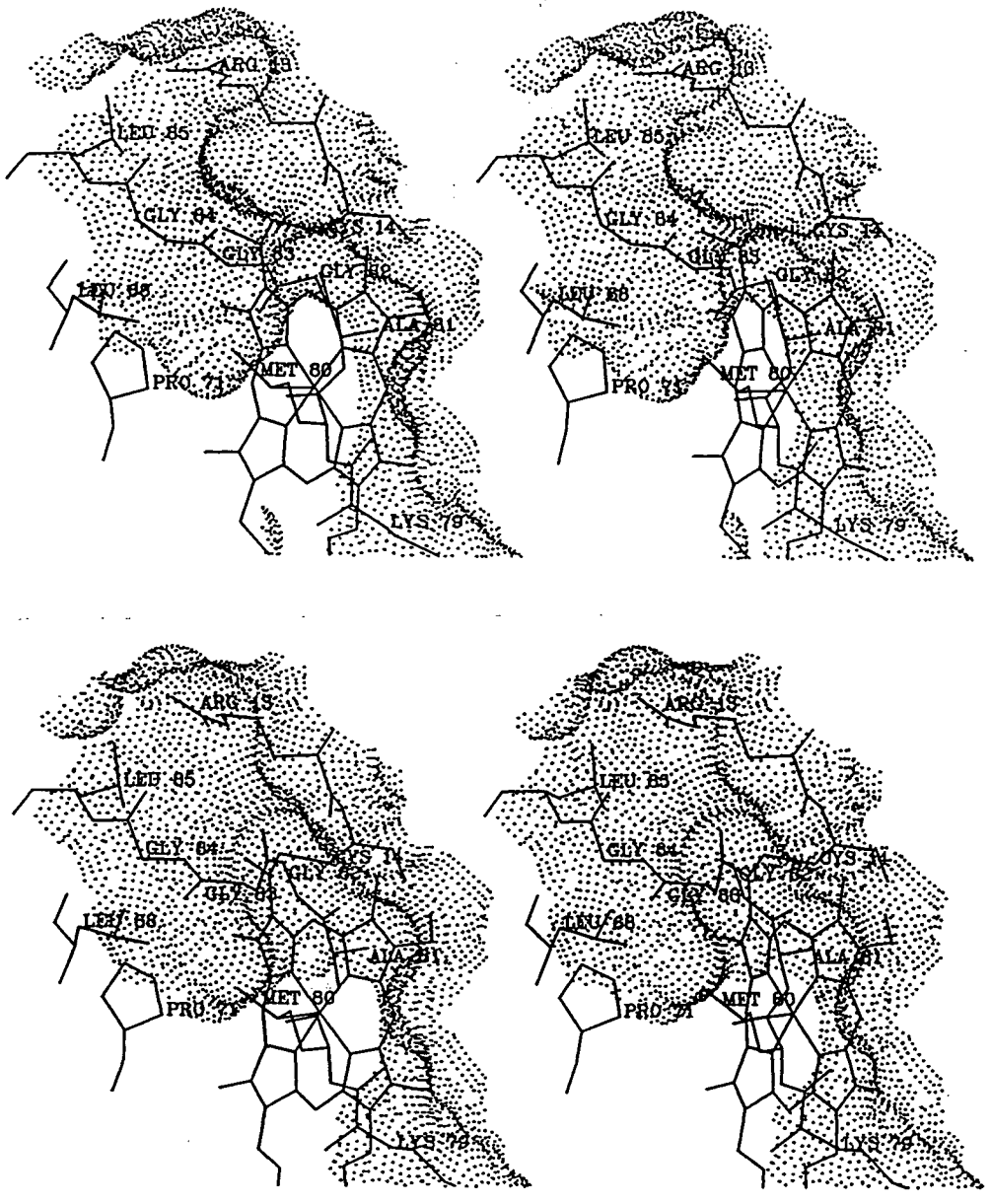

Figure VI.8. Molecular surface in the region of residue 82 in the Gly82(Cys102) (*top*) and Gly82(Thr102) (*bottom*) variant iso-1-cytochromes *c*. The dot surfaces represent the regions of the two protein molecules which are accessible to solvent, and were calculated (Connolly, 1983) using a probe sphere of radius 1.4 Å. The directions of view are identical to that in Figure V.4, and thus the illustrated regions of the surfaces of the two Gly82 variants can be directly compared with that of the wild-type iso-1-cytochrome *c* molecule presented in the earlier figure.

the heme pocket arising from the increased accessibility of solvent (see Section V.C.1.a). For the Gly82 variant, the exposure of the heme group does not differ greatly from that in the wild-type protein, but a different mechanism may effect a similar increase in the polarity of the heme

environment: refolding of residues 81 through 84 places a number of polar groups from the polypeptide chain adjacent to the heme group. In the Gly82(Cys102) variant, the amide nitrogens of residues 81 through 83 and the carbonyl group of residue 82 are directed into the heme pocket. In the Gly82(Thr102) variant, the carbonyl groups of residues 81 and 83 and the amide nitrogen of residue 83 are directed into the heme pocket. It is notable that in the wild-type iso-1-cytochrome *c* molecule, these groups are well removed from the heme and interact with solvent molecules, and that the corresponding region of the heme pocket is occupied by the non-polar phenyl group of Phe82. An additional factor is that despite the similarity in the degree of exposure to solvent of the heme groups of the Gly82 variant and wild-type iso-1-cytochromes *c*, the thickness of the protein layer separating the heme from bulk solvent is decreased in the Gly82 variant. The inward folding of the main chain of residues 82 through 84 has removed a outward bulge at the "shoulder" of the molecule, and has instead created a depression in the molecular surface to the immediate left of the Met80 ligand (see Figure VI.8). The result is that components of the hydrophobic heme pocket normally buried 6 Å below the surface of the molecule have become exposed to the external solvent medium. In particular, Wat108 in the Gly82(Thr102) variant, hydrogen bonded to Met80 O and Gly83 N, is positioned within van der Waals contact of the Met80 ligand (see Figure VI.4). Thus, for the Gly82 variant, the decreased reduction potential would appear to arise from the increased polarity of the heme environment due to the presence of polar protein atoms in the immediate vicinity of the heme group, and from the closer proximity of the external solvent medium.

In comparison to its position in wild-type iso-1-cytochrome *c*, Wat166 in both the Ser82 and Gly82 variants has a position ~0.6 Å closer to the heme iron. As discussed in Section V.D for the Ser82 variant, the shift of Wat166 apparently occurs in response to an increased electropositivity of the heme iron, which is in turn associated with the lower reduction potential of this protein. A similar effect may be responsible for the movement of Wat166 in the Gly82 variant, which also shows an ~50 mV decrease in reduction potential. In fact, that the same shift of the internal water molecule occurs in two proteins in which structural differences are otherwise

disparate supports the proposal that the electropositivity at the heme iron affects the positioning of Wat166.

3. Stability of the heme crevice

The pK_a for the alkaline isomerization of the Gly82 variant is observed to be 7.7, 0.8 pH units below that of wild-type iso-1-cytochrome *c*. As discussed for the Ser82 variant, the decreased pK_a indicates that the stability of the heme crevice is reduced in the Gly82 variant. Several structural properties of the Gly82 variant may contribute to this lowering in stability. The presence in the Gly82 variants of a number of polar groups (amide nitrogens and carbonyl groups) which are directed into the non-polar heme pocket gives rise to a destabilizing effect in itself, and in addition generates a number of unsatisfied hydrogen bonding groups. In addition, the refolding of the polypeptide chain in the vicinity of the substitution site disrupts a network of hydrogen bonds which involves the water molecule Wat172 and which may have a role in stabilizing the native conformation of this region of the iso-1-cytochrome *c* molecule. It is notable that Wat172 is conserved in four cytochrome *c* structures (yeast, rice and both reduced and oxidized tuna), and can thus be considered an integral part of the cytochrome *c* molecule (see Section IV.G.1). Another factor of likely importance is the particularly high conformational entropy (Matthews *et al.*, 1987) of the unfolded state of the Gly82 variant. The occurrence of three consecutive glycine residues at positions 82 through 84 provide this region of the polypeptide chain considerable flexibility, which would be lost if a single fixed conformation were to be adopted. In fact, that two distinct conformations are observed for this segment of polypeptide chain can be taken as a direct indication of the flexibility of this region of the Gly82 variant molecule.

4. Electron transfer activity

In comparison to wild-type iso-1-cytochrome *c*, the Gly82 variant has a considerably reduced ability to carry out electron transfer with cytochrome *c* peroxidase. For the Gly82 variant, the steady state activity for oxidation by cytochrome *c* peroxidase is only 20% that for the wild-type protein (Pielak *et al.*, 1985). Furthermore, in the complex between iso-1-cytochrome *c*

and zinc-substituted cytochrome *c* peroxidase, the rate of reduction of Zn^{+}-cytochrome *c* peroxidase by cytochrome *c* is 10^{4} times slower for the Gly82 variant than for the wild-type protein. A similar decrease in activity with Zn^{+}-cytochrome *c* peroxidase is observed for the Ser82 variant. For this latter variant, because no substantial conformational rearrangements occur in the vicinity of the substitution site, the cause of the lower reaction rate appears to be the altered nature of the medium encountered by the electron being transferred, and in particular the lack of a π-electron system coupled to that of the heme (see Section V.C.3). For the Gly82 variant, a similar explanation also seems appropriate. In this variant, the volume of space corresponding to that occupied by the aromatic side chain of Phe82 in the wild-type molecule is filled by the more polar atoms of the main chain of residues 82 and 83. The polypeptide chain backbone would lack the ability to interact directly with the π-electron system of the heme group, and may also result in the inclusion of bound water molecules at the interface between the cytochrome *c* and cytochrome *c* peroxidase molecules. The altered character of the polypeptide chain groups in contact with the heme group would also be expected to affect the distribution of electrons within the π-orbitals of the heme (Shellnutt *et al.*, 1981, 1979).

In the series of position 82 variants, the Gly82 variant is anomalous in having a very low reduction rate in the photo-initiated electron transfer from ^{3}Zn-cytochrome *c* peroxidase to iso-1-cytochrome *c* (13 s^{-1} as compared to ~200 s^{-1} for the wild-type protein) (Liang *et al.*, 1987, 1988). Since this rate is dependent on the distance separating the redox centres, Liang *et al.* have inferred that the Gly82 variant forms a perturbed complex with zinc-substituted cytochrome *c* peroxidase. That a large conformational rearrangement in the vicinity of the substitution site occurs in this variant provides support for this interpretation. Modelling of the complex between cytochrome *c* and cytochrome *c* peroxidase (Poulos and Kraut, 1981; Lum *et al.*, 1987) indicates that the region of the iso-1-cytochrome *c* molecule about Phe82 participates in forming the interprotein contact interface. Comparison of the structures of the Gly82 variant and wild-type iso-1-cytochromes *c* in the vicinity of residue 82 indicates at least two ways by which polypeptide chain refolding in this region of the protein could affect complexation interactions with

cytochrome *c* peroxidase. First, there is an altered disposition of non-polar and polar groups, with the non-polar phenyl ring of Phe82 being replaced by more polar atoms of the polypeptide chain backbone (see Figures VI.3 and VI.4). Secondly, the inward movement of residues 82 through 84 changes the topography of the surface of the cytochrome *c* molecule, and in particular reduces the surface area available to interface the cytochrome *c* peroxidase molecule (Figures VI.8 and V.5). The end result of these changes is an interaction surface that is likely to be less compatible with the complementary surface of cytochrome *c* peroxidase.

The proposal that for the Gly82 variant, these conformational rearrangements are responsible for the perturbed complex formed with cytochrome *c* peroxidase is supported by the structural analysis of the Ser82 variant. This protein also has a small side chain at position 82, but undergoes no substantial polypeptide chain rearrangements. For the Ser82 variant, reduction by ^{3}Zn-cytochrome *c* peroxidase occurs at a near normal rate. The extreme sensitivity of the rate of reaction between ^{3}Zn-cytochrome *c* peroxidase and ferricytochrome *c* to optimal complex formation is further indicated by studies using horse and tuna cytochrome *c*. These proteins are highly homologous to yeast iso-1-cytochrome *c*, but have a lower affinity for yeast cytochrome *c* peroxidase (Kang *et al.*, 1977). Compared to yeast iso-1-cytochrome *c*, both horse and tuna cytochromes *c* are reduced by ^{3}Zn-cytochrome *c* peroxidase at a decreased rate (Ho *et al.*, 1985; Liang *et al.*, 1986) which is similar to that observed for the reduction of the Gly82 variant. Thus it would appear that for the Gly82 variant, the abberant electron transfer properties observed can be explained by structural changes which occur about the substitution site and which in turn lead to non-optimal complexation with cytochrome *c* peroxidase.

D. Discussion of the polypeptide chain refolding

It is well established that due to hydrophobic interactions, globular proteins have a native fold in which non-polar groups are positioned in the interior of the protein molecule, sequestered from contact with the solvent medium (Schulz and Schirmer, 1979). As described earlier, if removal of the phenyl ring of Phe82 were to occur unaccompanied by adjustment of the conformation of the main chain in this region of the molecule, several hydrophobic groups internal

to the heme pocket would become exposed to the solvent. Exposure of these non-polar groups would be energetically unfavorable. It is notable that in both Gly82 variants, a high degree of exclusion of solvent from the heme crevice is maintained. It would thus seem likely that in the Gly82 variant, polypeptide chain refolding occurs in order to bury that portion of the hydrophobic heme pocket which would otherwise be rendered accessible to solvent.

In comparison to amino acids having a β-carbon, glycine is considerably less restricted in the backbone conformations which it can adopt (Richardson, 1981). In the Gly82 variant, the introduction of Gly82 creates a sequence which contains three consecutive glycine residues. Thus, clearly a factor in the substantial refolding that occurs about residues 82 to 84 is the high flexibility of this region of the polypeptide chain. In the Gly82(Cys102) variant, both Gly83 ($\phi = 81^\circ$, $\psi = 179^\circ$) and Gly84 ($\phi = 110^\circ$, $\psi = 120^\circ$) adopt main chain conformational torsion angles that are disallowed for residues bearing a side chain. The same is true for Gly82 ($\phi = 57^\circ$, $\psi = 16^\circ$) in the Gly82(Thr102) variant. The occurrence of three consecutive glycine residues facilitates in an additional manner the polypeptide chain refolding which packs residues 82 to 83 against the heme group. For the polypeptide chain conformation observed in both Gly82 variants, a CB carbon atom on either residue 82 or 83 would project directly into the heme pocket and form prohibitive steric conflicts with atoms of the heme and of the Met80 ligand. These observations suggest that replacements at residue 82 with amino acids other than glycine would result in overall retention of the native conformation of the polypeptide chain about this site. As discussed above and also later, this is supported by the observed structures of the Ser82, Tyr82, and Ile82 variant proteins.

The Gly82 variant of iso-1-cytochrome *c* demonstrates that a single amino acid replacement can lead to a substantial refolding of the polypeptide chain. Structural studies of other mutated proteins have shown that in general single-site amino acid replacements are accommodated without gross rearrangements from the native conformation (Alber *et al.*, 1987, 1988; Howell *et al.*, 1986). Clearly replacement of an amino acid residue with a glycine residue effects fairly drastic steric and entropic changes, as a void is created at the site of the original side chain,

and an increase in flexibility of the polypeptide backbone is induced. Nonetheless, with two glycine substitutions in T4 lysozyme, only slight changes in polypeptide chain conformation have been observed to occur at the substitution site. In a proline to glycine substitution occurring within an α-helix, the space originally occupied by the proline side chain is filled by adjacent groups of the protein molecule which have undergone slight positional shifts (Alber *et al.*, 1988). In a threonine to glycine substitution within a surface loop of the protein, the vacated space is filled by a water molecule (Alber *et al.*, 1987).

In iso-1-cytochrome *c*, the environment of residue 82 which allows the substantial refolding to occur is unique in several respects. Because of the large volume of the phenylalanine side chain, a glycine substitution creates a particularly sizeable void. Also, since the phenyl ring of Phe82 occupies a non-polar pocket in the surface of the molecule, adjacent to the rigid planar heme group (see Section III.H.1), two means used in T4 lysozyme to fill vacated space are unavailable. The apolarity of the surrounding protein atoms is unfavorable for binding solvent molecules, and there are no other protein groups which can adjust their conformations to fill the vacated space. In addition, two factors provide the polypeptide chain around the substitution site considerable flexibility. The first is, as described above, the occurrence of three consecutive glycine residues. Second, this strand of polypeptide chain has an extended conformation, and thus there are no intrastrand hydrogen bonds to confer rigidity. In summary, the ability of the segment of polypeptide chain spanning residues 81 to 84 to access a large number of conformations, coupled with the inavailability of alternative means of filling the space vacated by the Phe82 phenyl ring, likely accounts for the substantial refolding that is observed in the Gly82 variant.

The occurrence of substantial refolding of the polypeptide chain of the Gly82 variant suggests an interesting speculation regarding the protein-folding pathway of cytochrome *c*. If folding occurs via a unique pathway (see Kim and Baldwin, 1982), then either Phe82 is not important as a determinant of the overall conformation, or the conformation of Phe82 is fixed late in the folding process. The latter supposition is supported by experimental data from two sources. First, ligation of the nearby Met80 to the heme iron has been shown to occur during the final

stage of protein folding (Ridge *et al.*, 1981). Second, a qualitative description of the folding pathway of cytochrome *c* has been obtained using hydrogen exchange labelling in combination with NMR spectroscopy (Roder *et al.*, 1988). The indicated pathway shows that the N- and C-terminal α-helices form early during protein folding, whereas the 60's and 70's helices at the Met80 side of the molecule are formed later.

VII. TYROSINE 82 VARIANT OF YEAST ISO-1-CYTOCHROME *c*

A. Details of the structure determination

1. X-ray diffraction data

Diffraction data to 1.97 Å, 7319 reflections in total, for the Tyr82 variant of iso-1-cytochrome *c* were collected from a single crystal. Background averaging was used in the processing of diffraction intensities. For the resultant set of structure factors, a scale factor of 23.2 was indicated by a linear rescale against common reflections in the resolution range ∞ to 2.2 Å from the data set for wild-type iso-1-cytochrome *c*. The R-factor between the two data sets, considering all common structure factors, was 0.25.

2. Difference maps

A difference map (Figure VII.1) was calculated using the coefficients F_o(Tyr82) - F_o(wild-type), and phases derived from a wild-type iso-1-cytochrome *c* atomic model which had been refined at 1.7 Å resolution to an R-factor of 0.197. Consistent with the presence of an

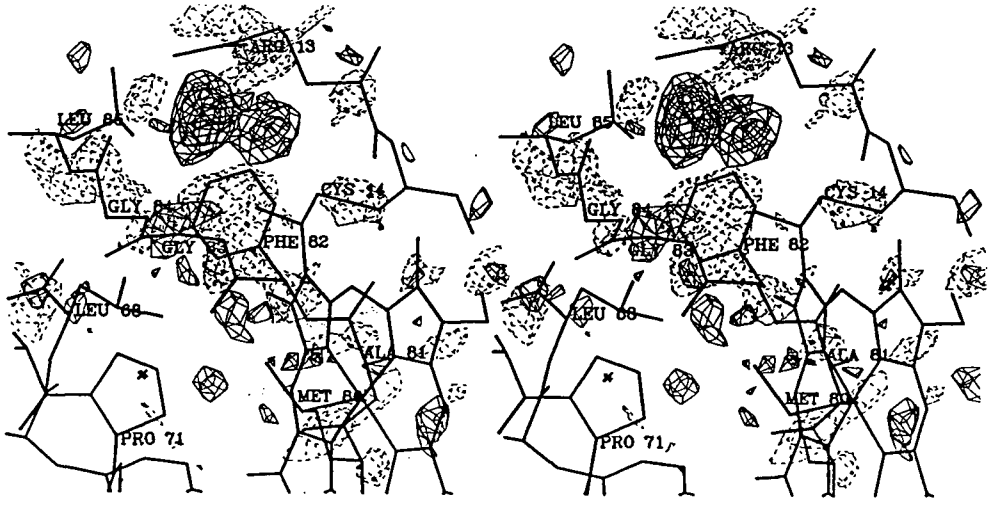

Figure VII.1. Stereo drawing of the F_o(Tyr82) - F_o(wild-type) difference map in the vicinity of residue 82. The atomic skeleton of wild-type iso-1-cytochrome *c* determined at 1.7 Å resolution is superimposed on the difference electron density envelopes. Positive density (solid lines) has been contoured at two levels, and negative density (dashed lines) at a single level.

extra hydroxyl group on the phenolic ring of residue 82 in the Tyr82 variant, the difference map showed a significant peak of positive electron density adjacent to the position occupied by the CZ atom of Phe82 in the wild-type molecule. However, negative density enveloping the lower portion of the phenyl ring, and additional positive density adjacent to the upper edge indicated that this ring occupies a more external positioning in the Tyr82 variant than in the wild-type protein. In addition, a series of paired positive and negative peaks occurring adjacent to the main chain of residues 82 to 85 indicated that these residues in the Tyr82 variant have undergone a slight inward and upward shift in position. Also, the replacement of Cys102 by Thr102 was indicated by a strong negative peak at the SG atom of Cys102, and an adjacent positive peak. A $2F_o$(Tyr82) - F_c' map, in which the calculated structure factor magnitudes and phases excluded the contribution of residues 81 through 85, was consulted in positioning the side chain of Tyr82, and in adjusting slightly the polypeptide backbone conformation of residues 81 through 85. A $2F_o - F_c$' map was also used to position the side chain of Thr102.

3. Refinement of the structure of the Tyr82 variant

The initial model used in the refinement of the Tyr82 variant was derived from a 1.7 Å resolution wild-type iso-1-cytochrome *c* structure, which included 98 water molecules and 1 sulfate ion. The preliminary model incorporated adjustments made through inspection of $2F_o - F_c$' maps, as described above. The amount of diffraction data included in the refinement was increased in three stages. Initially, 14 cycles of refinement against data with $F \geq 3\sigma_F$ in the resolution range 6.0 - 2.1 Å (3466 reflections) yielded a reduction in R-factor from 0.230 to 0.187. Next, 4057 reflections with $F \geq 3\sigma_F$ in the resolution range 6.0 - 1.97 Å were included, and 4 cycles of refinement resulted in an atomic model with an R-factor of 0.176. In the final phase, 15 refinement cycles were carried out, using data with $F \geq 2\sigma_F$ in the resolution range 6.0 - 1.97 Å (4546 reflections, corresponding to 70% of the available reflections). At various intervals, the atomic model was adjusted against $2F_o - F_c$ or $F_o - F_c$' electron density maps. The manual adjustments required were small, and were restricted to the side chains of amino acid residues. An additional 21 water molecules were incorporated into the atomic model. For the final refined

Table VII.1. Agreement with ideal stereochemistry in the final refined model of the Tyr82 variant of yeast iso-1-cytochrome *c* at 1.97 Å resolution

Class of restraint	R.m.s. deviation from ideality
1-2 bond distance	0.018 Å
1-3 angle distance	0.037 Å
1-4 planar distance	0.047 Å
Planar	0.016 Å
Chiral centre	0.184 Å^3
Non-bonded contact[a]	
single torsion	0.209 Å
multiple torsion	0.193 Å
possible hydrogen bond	0.233 Å
Staggered (±60°, 180°) torsion angle	21.4°
Planar torsion angle (0 or 180°)	2.9°
Temperature factor	
1-2 bond (main chain)	1.0 Å^2
1-3 angle (main chain)	1.6 Å^2
1-2 bond (side chain)	1.3 Å^2
1-3 angle (side chain)	2.1 Å^2

[a]The r.m.s. deviations from ideality for this class of restraint incorporate a reduction of 0.1 Å from the radius of each atom involved in a contact.

structure of the Tyr82 variant of iso-1-cytochrome *c*, the R-factor is 0.183 and the agreement with ideal stereochemistry is very good (see Table VII.1).

B. Structural differences between the Tyr82 variant and wild-type iso-1-cytochrome *c* molecules

The replacement of the phenylalanine at position 82 with a tyrosine has little effect on the conformation of the iso-1-cytochrome *c* molecule. Figure VII.2 shows a comparison of the Tyr82 variant and the wild-type protein in terms of the average positional difference between main chain atoms, and the differences in conformational torsion angles of corresponding residues. The overall r.m.s. difference between positions of corresponding main chain atoms is 0.19 Å, and the overall r.m.s. differences between phi and psi angles are 6.2° and 6.3°, respectively. The heme groups of

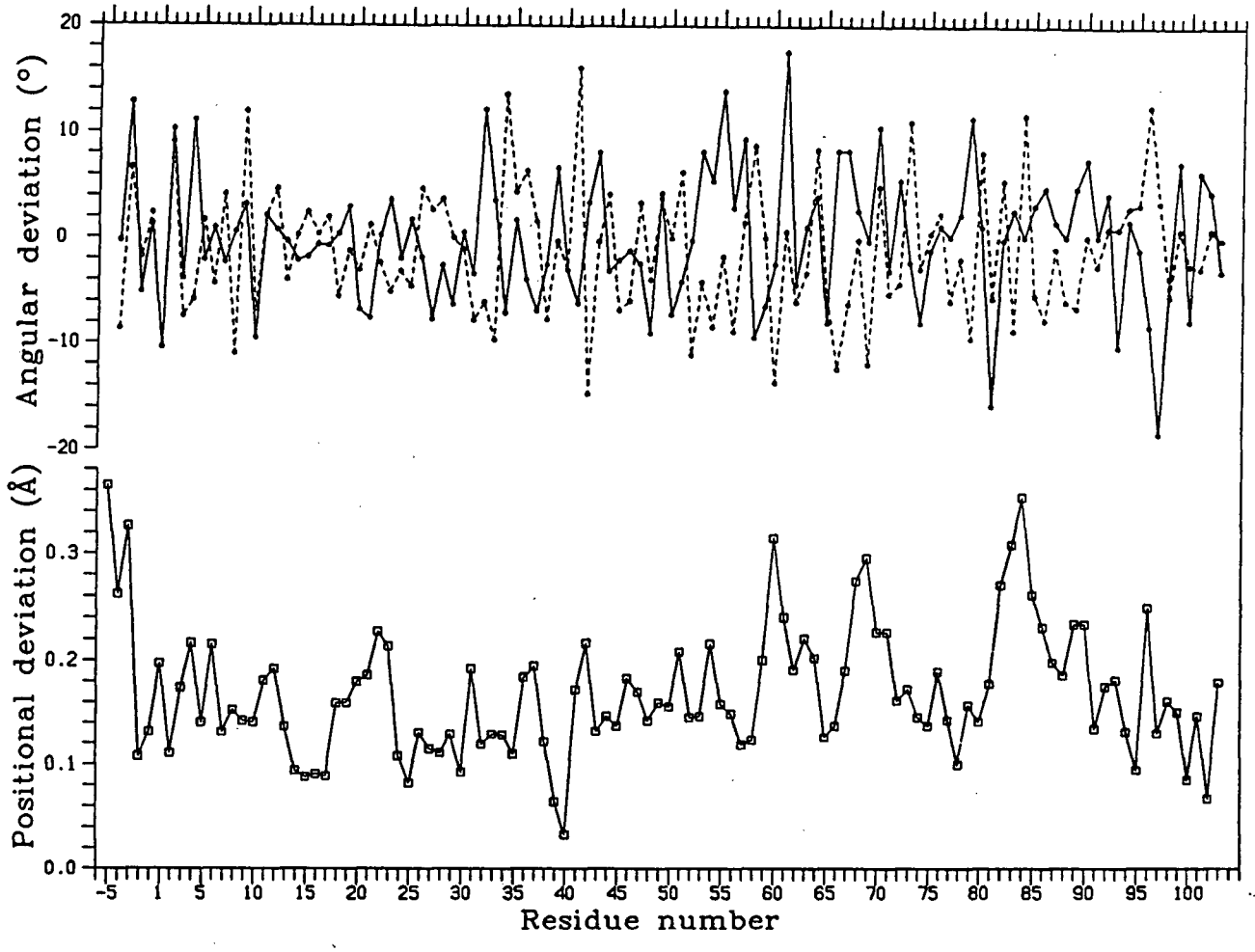

Figure VII.2. Comparison of the polypeptide backbone conformations of the Tyr82 variant and wild-type iso-1-cytochromes *c*. The lower plot indicates the average deviations observed between the positions of main chain atoms in corresponding residues of the two proteins. The upper plot indicates the deviations in the main chain conformational torsion angles ϕ (—) and ψ (---) of corresponding residues.

the Tyr82 variant and wild-type iso-1-cytochromes *c* do not differ significantly in either overall positioning or conformation, as the mean difference in atomic positions of the 43 heme atoms is only 0.13 Å.

The side chain of Tyr82 adopts a conformation ($\chi1=164°$, $\chi2=-107°$) very similar to that of Phe82 in the wild-type iso-1-cytochrome *c* molecule ($\chi1=174°$, $\chi2=-110°$; see Section III.H.1). However, the phenolic ring of the Tyr82 side chain is positioned ~0.7 Å further out of the heme pocket than the phenyl ring of Phe82 (see Figure VII.3). The shift of the phenyl ring

toward the surface of the molecule in the Tyr82 variant is likely the result of two factors. First, the additional hydroxyl group cannot be accommodated with the positioning of the phenyl ring present in the wild-type protein, as it would sterically conflict with the side chain of Leu85. Secondly, the positioning observed for the side chain of Tyr82 directs the hydroxyl group toward the external solvent medium, and allows this group to form a hydrogen bond with the nearby guanidinium group of Arg13. It is notable that in adopting a more external positioning, the phenolic ring of Tyr82 is also able to pack against the heme group with a greater degree of parallelism (the interplanar angle is 8°, as compared to 23° for Phe82 in wild-type iso-1-cytochrome *c*).

Significant differences between the Tyr82 variant and wild-type iso-1-cytochromes *c* do occur in the conformation of the polypeptide backbone of residues 82 through 85. The set of shifts observed apparently serves to accommodate the more external positioning of the side chain of Tyr82. The atoms of the main chain of residue 82 have shifted on average 0.27 Å in an upward direction (0.34 Å at the α-carbon). However, due to a 16° rotation in the ϕ main chain torsion angle of the preceding residue (Ala81), the upward shift in the position of the side chain

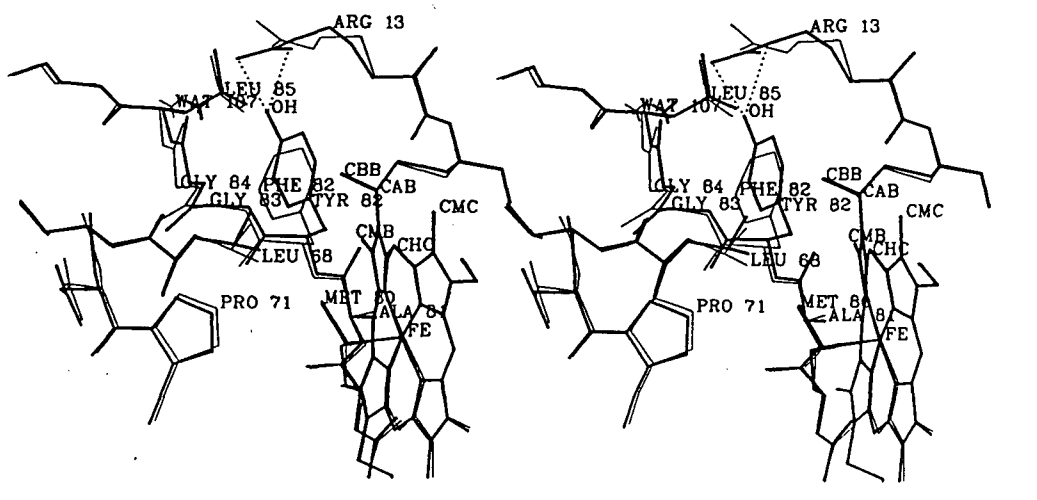

Figure VII.3. Stereo drawing of the region around residue 82 in the Tyr82 variant (thick lines) and wild-type (thin lines) iso-1-cytochromes *c*. The heme groups and the polypeptide chain of residues 13 through 14, 68 through 71, and 79 through 85 of the two proteins are shown superimposed. Hydrogen bonds between the side chain of Arg13 and the hydroxyl group of the Tyr82 tyrosyl ring are shown as thin dashed lines.

CB atom of Tyr82 is 0.68 Å, about the same as that of the phenolic ring overall. Also observed are the main chain movements, primarily in an inward direction, of Gly83, Gly84, and Leu85 by 0.31, 0.36, and 0.26 Å (on average), respectively. The positional shifts of residues 81 through 84 are accommodated by small adjustments in the main chain torsion angles within this region, most notably ϕ of Ala81 by 16°, and ψ of Gly84 by −12°. The positioning of the side chain of Arg13 and of Wat107, which are both located in the vicinity of the substitution site, are also perturbed in the Tyr82 variant. The more external positioning of the side chain of Tyr82 causes an upward shift of the alkyl portion of the Arg13 side chain. In addition, an adjustment in the orientation of the guanidinium group of Arg13 allows the NH1 and NE atoms to form hydrogen bond interactions with the hydroxyl group of Tyr82. The inward shift of the main chain atoms of Gly84 appears to induce a similar shift in Wat107, which is hydrogen bonded to the carbonyl group of Gly84. No other significant differences are observed between the Tyr82 variant and wild-type iso-1-cytochromes *c* in the positions of either protein groups or bound solvent molecules.

Pielak *et al.* (1988a) have studied using NMR spectroscopy the structural differences between the Tyr82 variant and wild-type iso-1-cytochromes *c*. They have shown that in the Tyr82 variant, changes in chemical shifts occur only at groups located within ~5 Å of the site of the amino acid substitution. This result is consistent with the absence of significant differences between the crystal structures of the two proteins, aside from the more external positioning of the Tyr82 side chain. However, the NMR and crystallographic studies do not agree on the details of the minor conformational rearrangements occurring in the vicinity of Tyr82. Both the magnitude of the heme-induced diamagnetic shifts experienced by the protons of the aromatic side chain of residue 82, and the strength of the NOE's between these protons and protons of adjacent internal groups (Leu68 CD1 and CD2; Met80 CE; and Hem CMC and CBB) are similar in the two proteins. Thus, Pielak *et al.* infer that the side chains of Tyr82 and Phe82 occupy identical positions in their respective proteins, and attribute the observed changes in the chemical shifts of proton resonances of groups in the vicinity of Tyr82 to the decreased ring current shift factor of

the phenolic side chain. Between the Tyr82 variant and wild-type iso-1-cytochromes *c*, the most significant differences in the NOE's involving residue 82 occur for the side chain of Leu85, which is thus inferred to adopt a substantially different conformation in the Tyr82 variant. These interpretations are not supported by the crystallographic results. Comparison of the crystal structures of the Tyr82 variant and wild-type proteins shows that the positioning of the Tyr82 and Phe82 aromatic rings differ significantly, whereas the conformation of the Leu85 side chain is very similar in the two proteins [(χ1, χ2) are (-98°, -70°) and (-99°, -81°), respectively]. The origin of the discrepancy between the results from the two methods is unclear. It should be noted, however, that the crystallographically observed structural changes give rise to only very small differences in interatomic distances between Tyr82 CD1,CE1 and adjacent internal methyl groups (of Met80 CE, Hem CBB and CMC, and Leu 68 CD1), for which the corresponding NOE's have been observed to be little affected. In addition, that the conformation of the Leu85 side chain differs in the two proteins is inconsistent with the absence of significant perturbations both in the chemical shifts of the protons of this group, and in the NOE's between this group and adjacent groups in the heme pocket other than the side chain of residue 82 (Pielak *et al.*, 1988a).

C. Correlation of structural changes with differences in functional properties

1. Accessibility of the heme crevice and polarity of the heme environment

Despite the more external positioning and slightly altered orientation of the Tyr82 phenolic ring with respect to the plane of the heme group, the degree of exclusion of solvent from the upper left portion of the heme crevice in the Tyr82 variant differs little from that in wild-type iso-1-cytochrome *c* (see Figure VII.4). The solvent accessibility of the heme group in the Tyr82 variant (46.8 $Å^2$) is actually decreased slightly from that in the wild-type protein (48.8 $Å^2$). These results are consistent with the slightly reduced reactivity of the Tyr82 variant with $Fe(EDTA)^{2-}$ (6.2 X 10^4 $M^{-1}s^{-1}$ as compared to 7.2 X 10^4 $M^{-1}s^{-1}$ for wild-type iso-1-cytochrome *c*).

A comparison of the heme pockets of the Tyr82 variant and wild-type iso-1-cytochromes *c* shows that the most notable difference is the presence of a polar hydroxyl group on the side chain of Tyr82. The reduction potential of the Tyr82 variant is observed to be

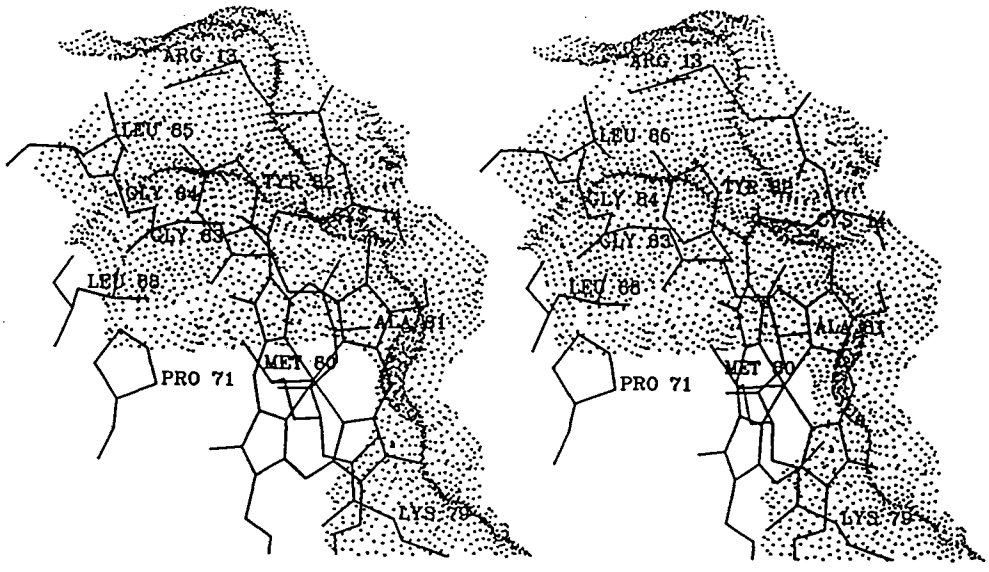

Figure VII.4. Molecular surface in the region of residue 82 in the Tyr82 variant iso-1-cytochrome *c*. The dot surface represents the region of the protein molecule which is accessible to solvent, and was calculated (Connolly, 1983) using a probe sphere of radius 1.4 Å. The direction of view is identical to that in Figure V.4, and thus the illustrated region of the surface of Tyr82 variant can be directly compared with that of the wild-type iso-1-cytochrome *c* molecule presented in the earlier figure.

slightly lower than that of wild-type iso-1-cytochrome *c* (280 mV vs. 290 mV). The 10 mV drop is the smallest perturbation in reduction potential observed for any of the position 82 variants studied thus far, and is in agreement with the similarity in the heme environments of the Tyr82 variant and wild-type iso-1-cytochromes *c*. The small drop in reduction potential may arise from the slight increase in the polarity of the heme pocket effected by the Tyr82 hydroxyl group. Notably, Wat166, whose positioning appears to be affected by the reduction potential of the iso-1-cytochrome *c* protein, occupies a position in the Tyr82 variant not significantly different from that of Wat166 in the wild-type molecule. In contrast, in the Ser82 and Gly82 variants, which have reduction potentials ~50mV below that of wild-type iso-1-cytochrome *c*, Wat166 has undergone a significant shift toward the heme iron (see Sections V.B.3 and VI.B.3).

2. Stability of the heme crevice

The pK_a for the alkaline isomerization of the Tyr82 variant is observed to be 8.4, only 0.1 pH units below that of wild-type iso-1-cytochrome *c*. Thus, it can be inferred that the stability

of the heme crevice of the Tyr82 variant is not significantly different from that of the wild-type protein. Nonetheless, comparison of the structures of the Tyr82 variant and wild-type iso-1-cytochromes *c* in the vicinity of residue 82 does suggest at least two structural features which would be expected to destabilize the folded state of the Tyr82 variant. First, the more external positioning of the phenolic ring of Tyr82 would be expected to perturb the packing interactions formed between the side chain of residue 82 and the surrounding hydrophobic groups of the heme crevice (see Section III.H.1) and to lower the packing efficiency in this region of the iso-1-cytochrome *c* molecule. Secondly, in the conformation adopted by the side chain of Tyr82, the polar hydroxyl group makes a repulsive van der Waals contact with the side chain of Leu85 (the Tyr82 OH - Leu85 CG distance is 3.01 Å). A comparison of atomic temperature factors does give a direct indication of greater mobility (and thus by inference, a less stable conformation) of this region of the Tyr82 variant. The average B-factor for the phenyl ring of Phe82 in the wild-type molecule is 15.1 $Å^2$, whereas that for the phenolic ring of Tyr82 is 18.5 $Å^2$.

3. Electron transfer activity

Of the position 82 variants studied thus far, the Tyr82 variant is the only one that is oxidized by Zn^+-cytochrome *c* peroxidase in the preformed complex between these two proteins at a rate comparable to that observed for wild-type iso-1-cytochrome *c* (1.5×10^4 s^{-1} vs. 1.9×10^4 s^{-1}, respectively) (Liang *et al.*, 1987, 1988). The interpretation of these results is that within the cytochrome *c* molecule, coupling of the π-electron systems of an aromatic group at position 82 and the heme group functions to facilitate electron transfer between Zn^+-cytochrome *c* peroxidase and iso-1-cytochrome *c* (Liang *et al.*, 1987, 1988). This proposal is supported by structural comparison of the Tyr82 variant and wild-type iso-1-cytochromes *c*, which shows that the phenolic ring of Tyr82 and the phenyl ring of Phe82 have an analogous positioning parallel to the heme group of the iso-1-cytochrome *c* molecule.

Although the Tyr82 variant has a normal reactivity with both ^3Zn- and Zn^+-cytochrome *c* peroxidase, its steady state activity with native cytochrome *c* peroxidase is decreased to 30% of that observed for wild-type iso-1-cytochrome *c*. Because the capacity of the Tyr82 variant to transfer

an electron to cytochrome *c* peroxidase is apparently normal, the decreased steady state activity would appear to result from an alteration in some other aspect of the bimolecular reaction.

One possibility is perturbed complexation. In the modelled complex between cytochrome *c* and cytochrome *c* peroxidase, the side chain of Phe82 forms important hydrophobic interactions at the interface between the two protein molecules, and the side chain of Arg13 is involved in an intermolecular salt bridge with an aspartate side chain (Lum *et al.*, 1987; Poulos and Kraut, 1981). Consideration of the proposed model suggests two factors which may inhibit the optimal association of the Tyr82 variant with cytochrome *c* peroxidase. First, the more external positioning of the side chain of Tyr82 may interfere sterically with complex formation. Secondly, because the side chain of Arg13 is involved in intramolecular hydrogen bonding with the side chain of Tyr82, it would be less available for participation in an intermolecular salt bridge. It should be noted that the rate of electron transfer from ^{3}Zn-cytochrome *c* peroxidase to iso-1-cytochrome *c*, which is affected by the distance separating the redox centres, is normal for the Tyr82 variant. This suggests that the iron-zinc distance in the preformed complex between the Tyr82 variant and Zn-cytochrome *c* peroxidase is not greatly perturbed. Therefore, if the structural differences between the Tyr82 variant and wild-type iso-1-cytochrome *c* are responsible for the observed decrease in steady state activity of the Tyr82 variant in reducing cytochrome *c* peroxidase, their effects are in slowing the rate of complex formation, or alternatively in inhibiting conformational rearrangements that occur in the cytochrome *c* molecule after the electron transfer has taken place (Liang *et al.*, 1987, 1988; Hoffman and Ratner, 1987).

VIII. ISOLEUCINE 82 VARIANT OF YEAST ISO-1-CYTOCHROME *c*

A. Details of the structure determination

1. X-ray diffraction data

The set of diffraction intensities to 2.3 Å resolution, 4375 reflections in total, for the Ile82 variant of iso-1-cytochrome *c* was measured from a single crystal, and was obtained from Dr. W. Hutcheon. Background averaging was used in the processing of the diffraction data. For the resultant set of structure factors, a scale factor of 38.2 was indicated by a linear rescale against common structure factors in the resolution range ∞ to 2.6 Å from the data set for wild-type iso-1-cytochrome *c*. The R-factor between the two data sets was 0.256 (considering only the reflections included in the linear rescaling).

2. Difference maps

A difference map (Figure VIII.1) was calculated using the coefficients F_o(wild-type) - F_o(Ile82), and phases derived from a wild-type iso-1-cytochrome *c* atomic model which had been

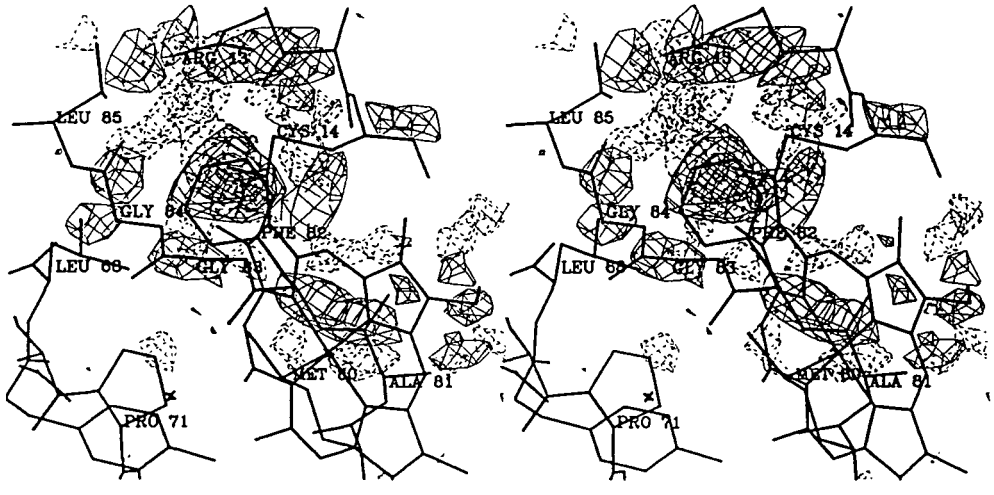

Figure VIII.1. Stereo drawing of the F_o(wild-type) - F_o(Ile82) difference map in the vicinity of residue 82. The atomic skeleton of iso-1-cytochrome *c* determined at 1.23 Å resolution is superimposed on the difference electron density envelopes. Positive density (solid lines) has been contoured at two levels, and negative density (dashed lines) at a single level.

refined at 1.23 Å resolution to an R-factor of 0.192 (see Section III). The most significant features of this map, as interpreted with respect to the superimposed structure of the wild-type protein, occurred in the vicinity of Phe82 and Cys102. Positive density enveloped the upper edge of the phenyl ring of Phe82 and also portions of the main chain atoms of residues 82 through 84. A diffuse negative peak occurred adjacent to Phe82 CB. The difference electron density peaks near Phe82 thus confirmed the amino acid replacement at this site, but could not clearly define the positioning of the substituted isoleucine side chain. This was more clearly indicated by a $2F_o$(Ile82) - F_c' map, for which the calculated structure factors excluded the contribution of residues 81 through 85 of the wild-type protein. This map was also consulted in adjusting slightly the conformation of the polypeptide backbone at residues 81 through 85. The positive difference electron density peak centred at the position of Cys102 SG and an adjacent negative peak were identical to those observed in difference maps for the Thr102-containing Gly82 and Tyr82 variants.

3. Refinement of the structure of the Ile82 variant

The initial model used in the refinement of the Ile82 variant was derived from the 1.23 Å resolution wild-type iso-1-cytochrome *c* structure. It incorporated adjustments to residues 81 through 85 based on the difference maps discussed above, as well as the sulfate ion and 73 water molecules having temperature factors less than 40 Å^2 from the wild-type model. The amount of diffraction data included in the refinement was increased in three stages. Initially, 5 cycles of refinement against data with $F \geq 3\sigma_F$ in the resolution range 6.0 - 2.6 Å (1600 reflections) yielded a reduction in R-factor from 0.211 to 0.163. Next, 1848 reflections with $F \geq 3\sigma_F$ in the resolution range 6.0 - 2.3 Å were included, and 11 cycles of refinement resulted in an atomic model with an R-factor of 0.162. In the final stage, 4 refinement cycles were carried out, using data with $F \geq 2\sigma_F$ in the resolution range 6.0 - 2.3 Å (2460 reflections, corresponding to 58% of the available reflections). At various intervals, minor manual adjustments to the atomic model were made through inspection of $2F_o - F_c$ or $F_o - F_c$' difference electron density maps. The final refined structure of the Ile82 variant of iso-1-cytochrome *c* has an R-factor of 0.182 and has very good agreement with ideal stereochemistry (see Table VIII.1).

Table VIII.1. Agreement with ideal stereochemistry in the final refined model of the Ile82 variant of yeast iso-1-cytochrome *c* at 2.3 Å resolution

Class of restraint	R.m.s. deviation from ideality
1-2 bond distance	0.012 Å
1-3 angle distance	0.028 Å
1-4 planar distance	0.034 Å
Planar	0.012 Å
Chiral centre	0.151 Å^3
Non-bonded contact[a]	
single torsion	0.203 Å
multiple torsion	0.167 Å
possible hydrogen bond	0.211 Å
Staggered ($\pm 60°$, $180°$) torsion angle	20.6°
Planar torsion angle (0 or $180°$)	2.3°
Temperature factor	
1-2 bond (main chain)	1.0 Å^2
1-3 angle (main chain)	1.6 Å^2
1-2 bond (side chain)	1.8 Å^2
1-3 angle (side chain)	3.0 Å^2

[a]The r.m.s. deviations from ideality for this class of restraint incorporate a reduction of 0.1 Å from the radius of each atom involved in a contact.

B. Structural comparison of the Ile82 variant and wild-type iso-1-cytochromes *c*

The replacement of Phe82 by isoleucine causes a slight rearrangement in the conformation of the polypeptide backbone at residues 81 through 85. However, the overall conformation of the iso-1-cytochrome *c* molecule is little affected. A residue-by-residue comparison of the main chain atomic positions and of conformational torsion angles in the Ile82 variant and wild-type iso-1-cytochromes *c* is shown in Figure VIII.2. The overall r.m.s. difference between positions of corresponding main chain atoms is 0.20 Å, and the overall r.m.s. differences between phi and psi angles are 7.3° and 6.5°, respectively. The heme groups of the Ile82 variant and wild-type iso-1-cytochromes *c* do not differ significantly in either overall positioning or conformation, as the mean difference in atomic positions of the 43 heme atoms is only 0.15 Å. The side chain of

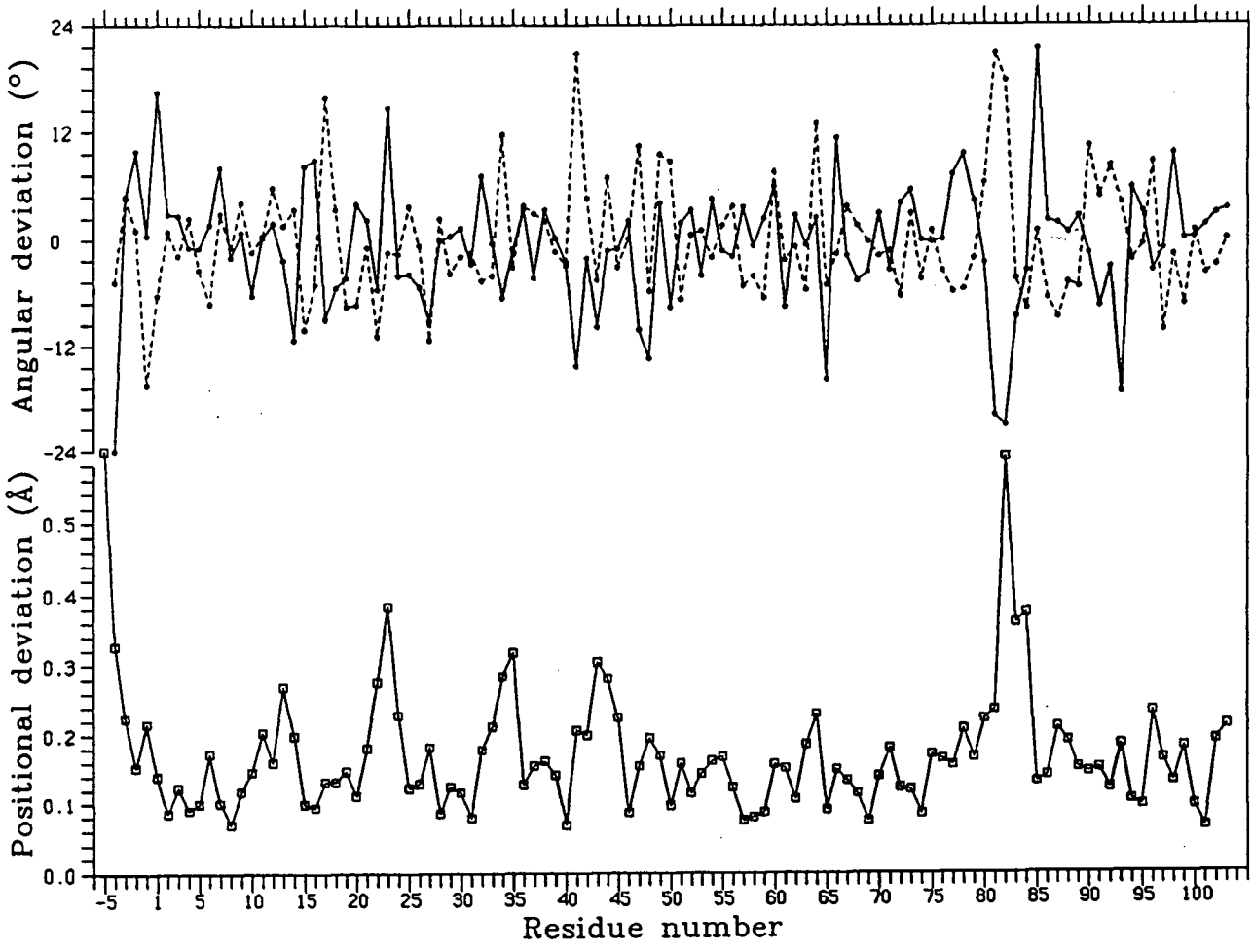

Figure VIII.2. Comparison of the polypeptide backbone conformations of the Ile82 variant and wild-type iso-1-cytochromes *c*. The lower plot indicates the average deviations observed between the positions of main chain atoms in corresponding residues of the two proteins. The upper plot indicates the deviations in the main chain conformational torsion angles ϕ (——) and ψ (---) of corresponding residues.

Ile82 adopts a tg^+ conformation ($\chi1=169°$, $\chi2=-97°$) such that the smaller of the two branches emanating from CB is directed downward toward the Met80 ligand. The CG2 atom (see Figure VIII.3) makes van der Waals contacts with both the Met80 CG (interatomic distance 3.2 Å) and CE (3.3 Å) carbon atoms. The observed adjustments in the conformation of the polypeptide backbone in the vicinity of Ile82 likely occur in order to relieve these steric conflicts. In the Ile82 variant, the main chain of residue 82 has shifted ~0.6 Å upward with respect to its position in the wild-type molecule. Accompanying shifts of Gly83 and Gly84 by 0.4 Å also occur. These

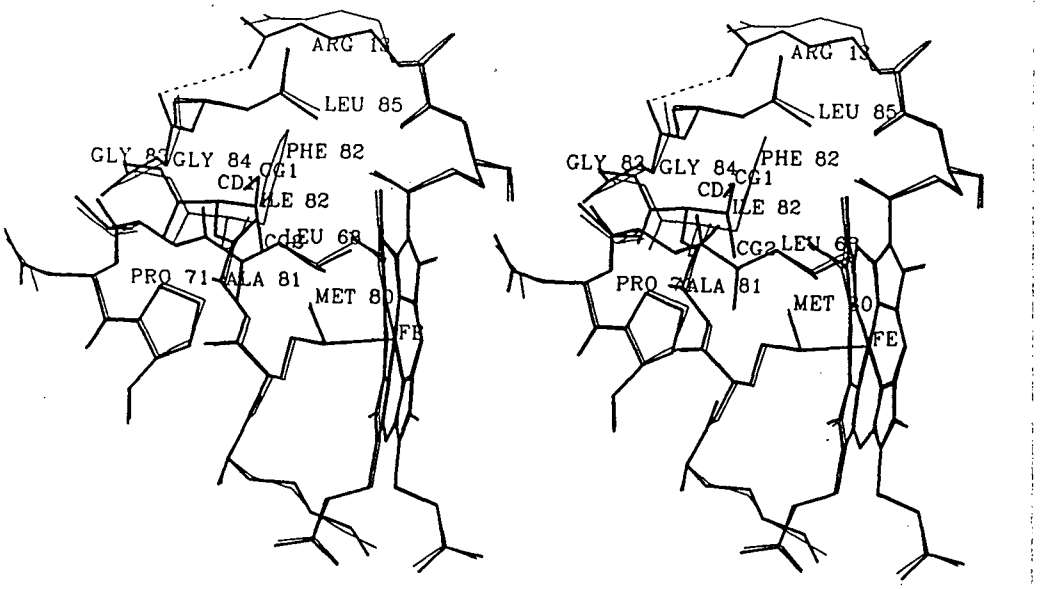

Figure VIII.3. Stereo drawing of the region around residue 82 in the Ile82 variant (thick lines) and wild-type (thin lines) iso-1-cytochromes *c*. The heme groups and the polypeptide chain of residues 13 through 14, 68 through 71, and 79 through 85 of the two proteins are shown superimposed.

movements are accommodated by small adjustments in main chain torsion angles within this region, most notably ψ of Ala81 by -21°, and ϕ of Leu85 by -21°.

With the observed positioning of the Ile82 side chain, the positions of the CG1 and CD1 atoms coincide roughly with those of the Phe82 CD2 and CE1 atoms in the wild-type molecule. The vacated space formerly occupied by CE2 and CZ of Phe82 allows the side chain of Arg13 to adopt a positioning closer to the body of the molecule and to form a hydrogen bond to the carbonyl group of Gly84.

The occurrence at residue 82 of a side chain branched at the β-carbon appears to induce considerable steric strain. As described above, the conformation adopted by the Ile82 side chain projects the CG2 atom directly at the Met80 ligand. However, a search for other possible orientations for the Ile82 side chain related by a rotation about the CA-CB bond clearly demonstrates that only the average conformation observed in the final refined atomic model is sterically permissible. Due to its branching at the CB atom, an isoleucine side chain typically strongly prefers a conformation which places its CG atoms *gauche* to both the N and C main chain atoms (McGregor *et al.*, 1987). Adoption of the g^- conformation would rotate the longer

arm of the Ile82 side chain downward into the heme pocket, and is thus prohibited by close contacts between the CD1 atom and the Met80 side chain. Rotation in the opposite direction (toward the exterior of the molecule) would generate the most commonly favored g^+ conformation for an Ile side chain, but is prevented by conflicts between the CG1 atom and the carbonyl group of Ala81. Thus the environment of Ile82 requires that this side chain adopt the least preferred t conformation [observed in less than 15% of isoleucines surveyed in various protein structures (McGregor *et al.*, 1987; Ponder and Richards, 1987)]. However, some disorder in the positioning of the Ile82 side chain is apparent from the diffuse electron density observed for this group in electron density maps.

It is notable that despite the poor accommodation of the Ile82 side chain, no substantial refolding of the polypeptide backbone occurs, thus indicating that the main chain of residues 81 through 85 may be constrained to adopt a native-like conformation. As discussed earlier, an important determinant of this conformation in the wild-type molecule appears to be favorable packing interactions formed by the phenyl ring of Phe82. As these interactions are clearly disrupted in the Ile82 variant. it seems likely that in the Ile82 variant the presence of a side chain at residue 82 limits the conformational flexibility of the adjacent region of polypeptide chain. (Recall that in the Gly82 variant, it is the adoption by Gly82 of conformational torsion angles permissible only for a residue without a side chain that allows the substantial localized refolding.) In addition, restrictions on the conformations accessible to the main chain about residue 82 may be imposed by the requirement to exclude polar protein groups and solvent from the hydrophobic pocket located below the side chain of residue 82, or by stabilizing hydrogen bond interactions formed between polar main chain groups and water molecules (in particular Wat172, see Section VI.C.3).

C. Correlation of structural changes with differences in functional properties

1. Accessibility of the heme crevice and polarity of the heme environment

The reduction potential of the Ile82 variant is ~20 mV lower than that of wild-type iso-1-cytochrome *c* (273 mV vs. 290 mV). A structural explanation for this drop in reduction

potential is not readily apparent. The side chain of Ile82 in the Ile82 variant occupies the same hydrophobic cavity in the heme pocket as the phenyl ring of Phe82 in the wild-type protein, although the smaller volume and shorter length of the isoleucine side chain means that this cavity is less completely filled in the former protein. This results in a small indentation in the surface of the Ile82 variant molecule, located between the Ile82 CD1 and Leu85 CG atoms (see Figure VIII.4). However, because this perturbation in packing occurs directly below the aliphatic portion of the side chain of Arg13, the heme group of the Ile82 variant is not rendered more accessible to solvent (the total solvent accessibility is 46.6 Å^2, as compared to 48.8 Å^2 in the wild-type protein). The polarity of the heme pocket in the Ile82 variant would also not be expected to altered greatly by the occurrence of the non-polar isoleucine side chain.

Considering the absence of obvious differences between the Ile82 variant and wild-type iso-1-cytochromes *c* in heme pocket polarity and solvent acessibility, one can speculate on the

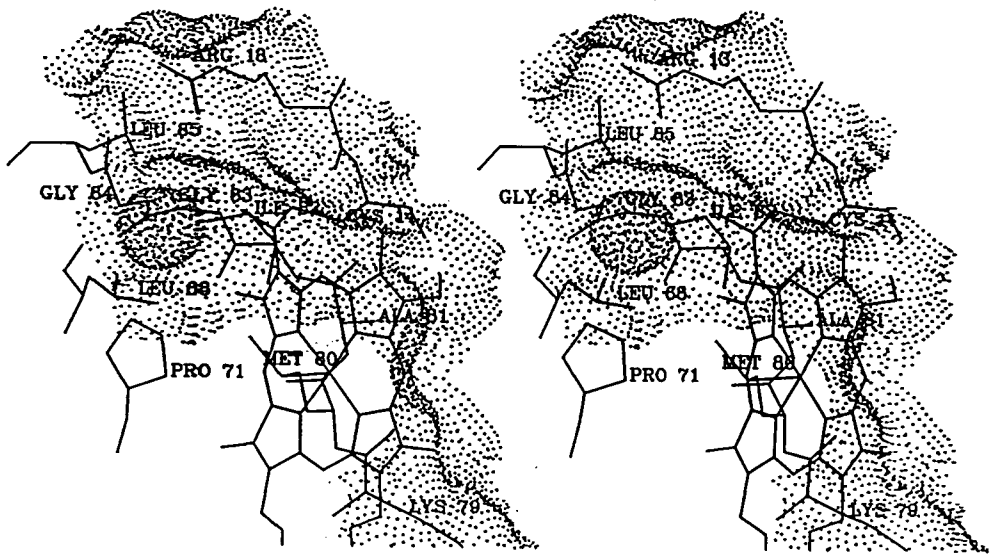

Figure VIII.4. Molecular surface in the region of residue 82 in the Ile82 variant iso-1-cytochrome *c*. The dot surface represents the region of the protein molecule which is accessible to a probe sphere of radius 1.4 Å (Connolly, 1983). Note the occurrence of a cavity beneath the external (solvent accessible) surface of the Ile82 variant molecule. This cavity is formed by the side chains of Leu68, Ile82 and Leu85, and the main chain of residues Gly83 and Gly84. The direction of view is identical to that in Figure V.4, and thus the illustrated region of the surface of Ile82 variant can be directly compared with that of the wild-type iso-1-cytochrome *c* molecule presented in the earlier figure.

cause of the decreased reduction potential of the Ile82 variant. One possible explanation is that the phenyl group of the Phe82 exerts a specific inductive effect on the π-electron system of the heme group. Wallace *et al.* (1989) have proposed that such an effect accounts for the drop in reduction potential arising from a Phe substitution at Tyr67. An alternative explanation is that steric contacts involving the Ile82 side chain may destabilize the conformation of the polypeptide backbone about residue 82, and thus cause transient unfolding of this segment of the polypeptide chain. Such transient unfolding may be reflected in the diffuse electron density apparent for the main chain of residues 82 and 83, and also in the decreased stability of the heme crevice in the Ile82 variant (see below), and would be expected to increase the effective accessibility of solvent to the heme group in this protein. This proposal may also explain the increased exposure of the heme group as indicated by the slightly elevated reactivity of the Ile82 variant with $Fe(EDTA)^{2-}$ (9.4×10^4 $M^{-1}s^{-1}$ as compared to 7.2×10^4 $M^{-1}s^{-1}$ for wild-type iso-1-cytochrome *c*).

As previously discussed, the position of the buried water molecule Wat 166 at the lower left of the heme crevice appears to be sensitive to the reduction potential of the heme iron. Interestingly, the position of Wat166 in the Ile82 variant is intermediate to those occurring in the Tyr82 variant or wild-type protein (which have normal reduction potentials) and the Ser82 or Gly82 variants (which have 50mV lower reduction potentials). Thus, in this series of position 82 variants, the magnitude of the positional shift of Wat166 appears to be correlated with the size of the drop in reduction potential.

2. Stability of the heme crevice

Of the position 82 variants studied thus far, the Ile82 variant has the most greatly perturbed pK_a for alkaline isomerization, 7.2 as compared to 8.5 for wild-type iso-1-cytochrome *c*. The decrease in pK_a indicates that the stability of the heme crevice is considerably reduced in this variant. The most important factor in the reduced stability is likely the destabilization of the native conformation of the polypeptide backbone about Ile82 induced by steric contacts between the Ile82 side chain and adjacent groups in the heme crevice. As shown in Figure VIII.4, a small void is created in the heme crevice adjacent to the Ile82 side chain. Thus, a disruption of the

packing interactions normally formed by the phenyl ring within its hydrophobic cavity may also contribute to the lowered stability of the Ile82 variant.

3. Electron transfer activity

The Ser82, Gly82 and Ile82 variants all show a decrease by a factor of ~10^4 in reaction rate with Zn^+–cytochrome *c* peroxidase. As discussed earlier, this decreased reactivity is attributed to the lack of an aromatic side chain at position 82, which apparently functions to facilitate electron transfer by interacting with the π–electron system of the heme group. Consistent with this proposal is that in the Ile82 variant protein, the volume of space filled by the phenyl ring of Phe82 in the wild–type protein is now occupied by the aliphatic side chain of Ile82 and empty space.

The Ile82 variant also has a significantly lowered reduction rate for the photo–initiated electron transfer from ^{3}Zn–cytochrome *c* peroxidase. This result indicates that the Ile82 variant may interact with Zn–cytochrome *c* peroxidase to form a perturbed complex, in which the distance separating the redox centres is longer than normal. Such a perturbation may be caused by the fairly substantial (~0.7 Å) displacement, in an upward direction, of the polypeptide backbone in the vicinity of Ile82.

IX. ALANINE 38 VARIANT OF YEAST ISO-1-CYTOCHROME *c*

A. Details of the structure determination

1. X-ray diffraction data

Diffraction data to 2.0 Å (7050 reflections in total), for the Ala38 variant of iso-1-cytochrome *c* were collected from a single crystal. Background averaging was used in the processing of diffraction intensities. For the resultant set of structure factors, a scale factor of 43.6 was indicated by a linear rescale against common structure factors in the resolution range ∞ to 2.6 Å from the data set for wild-type iso-1-cytochrome *c*. The R-factor between the two data sets was 0.27 (for the reflections included in scaling).

2. Difference map

A difference map (Figure IX.1) was calculated using the coefficients F_o(wild-type) - F_o(Ala38), and phases derived from a wild-type iso-1-cytochrome *c* atomic model which had been refined at 1.7 Å resolution to an R-factor of 0.197. This map was interpreted with respect to the

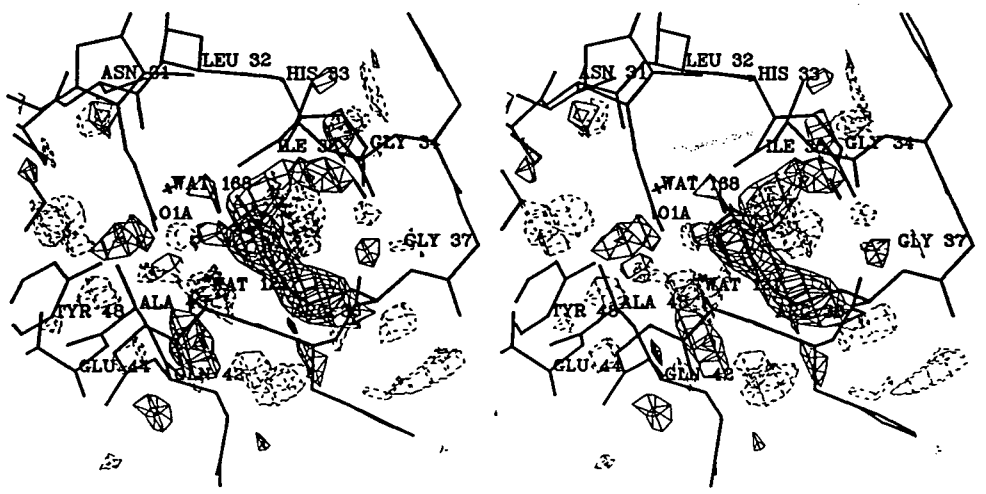

Figure IX.1. Stereo drawing of the F_o(wild-type) - F_o(Ala38) difference map in the vicinity of residue 38. The atomic skeleton of iso-1-cytochrome *c* determined at 1.7 Å resolution is superimposed on the difference electron density envelopes. Positive density (solid lines) has been contoured at two levels, and negative density (dashed lines) at a single level.

superimposed structure of the wild-type protein, and confirmed the replacement of Arg38. Positive electron density enveloped the entire Arg38 side chain, with the exception of the CB atom. No other significant positive or negative peaks were observed in the vicinity of the amino acid substitution site. The replacement of Cys102 by Thr102 in the Ala38 variant was evidenced by a strong positive peak at the SG atom of Cys102, and an adjacent negative peak.

3. Refinement of the structure of the Ala38 variant

The difference map indicated that structural changes occurring in the Ala38 variant were essentially limited to a deletion of the terminal 6 atoms (CG onwards) of the Arg38 side chain. Therefore, an initial model for the Ala38 variant was constructed from the 1.7 Å resolution wild-type iso-1-cytochrome *c* structure. For residue 38, only the CB carbon atom was retained; and for residue 102, a threonine side chain having a g^- conformation [as observed in the structures of the Gly82(Thr102) and Tyr82 variants] was appended to the α-carbon. Of the 98 water molecules and 1 sulfate ion present in the original wild-type iso-1-cytochrome *c* model, all were retained except for three water molecules (Wat108, 121 and 168) which form hydrogen bonds to the guanidinium group of Arg38.

The amount of diffraction data included in the refinement of the Ala38 variant was increased in three stages. The initial model was refined against data with $F \geq 2.5\sigma_F$ in the resolution range 6.0 - 2.5 Å (1664 reflections). Nine cycles lowered the R-factor from 0.226 to 0.167. Next, 2083 reflections with $F \geq 2.5\sigma_F$ in the resolution range 6.0 - 2.0 Å were included, and 4 cycles of refinement resulted in an atomic model with an R-factor of 0.198. In the final stage, 11 refinement cycles were carried out, using data with $F \geq 2\sigma_F$ in the resolution range 6.0 - 2.5 Å and $F \geq 2.5\sigma_F$ in the range 2.5 - 2.0 Å (2356 reflections in total, corresponding to 36% of the available reflections in the 6.0-2.0 Å range). At various intervals, the atomic model was inspected against $2F_o - F_c$ or $F_o - F_c$' electron density maps. Slight adjustments to the conformation of the polypeptide chain of residues 35 through 40 were made after the first stage of refinement. No other substantial manual adjustments to the protein molecule were required. However, a number of changes were made to the solvent structure in the vicinity of the

Table IX.1. Agreement with ideal stereochemistry in the final refined model of the Ala38 variant of yeast iso-1-cytochrome *c* at 2.0 Å resolution

Class of restraint	R.m.s. deviation from ideality
1-2 bond distance	0.012 Å
1-3 angle distance	0.028 Å
1-4 planar distance	0.033 Å
Planar	0.012 Å
Chiral centre	0.147 Å^3
Non-bonded contact[a]	
single torsion	0.201 Å
multiple torsion	0.189 Å
possible hydrogen bond	0.203 Å
Staggered (±60°, 180°) torsion angle	23.7°
Planar torsion angle (0 or 180°)	2.3°
Temperature factor	
1-2 bond (main chain)	0.6 Å^2
1-3 angle (main chain)	1.0 Å^2
1-2 bond (side chain)	0.5 Å^2
1-3 angle (side chain)	0.9 Å^2

[a]The r.m.s. deviations from ideality for this class of restraint incorporate a reduction of 0.1 Å from the radius of each atom involved in a contact.

substitution site, including the reincorporation of Wat121 and Wat168 (see below). The final refined structure of the Ala38 variant of iso-1-cytochrome *c* has an R-factor of 0.187 and good agreement with ideal stereochemistry (see Table IX.1).

B. Structural comparison of the Ala38 variant and wild-type iso-1-cytochromes *c*

1. Polypeptide chain conformation

The replacement of an arginine residue at position 38 with an alanine has very little effect on the overall conformation of the iso-1-cytochrome *c* molecule. Figure IX.2 shows a comparison of the Ala38 variant and the wild-type protein in terms of the average positional difference between main chain atoms and the differences in conformational torsion angles of corresponding

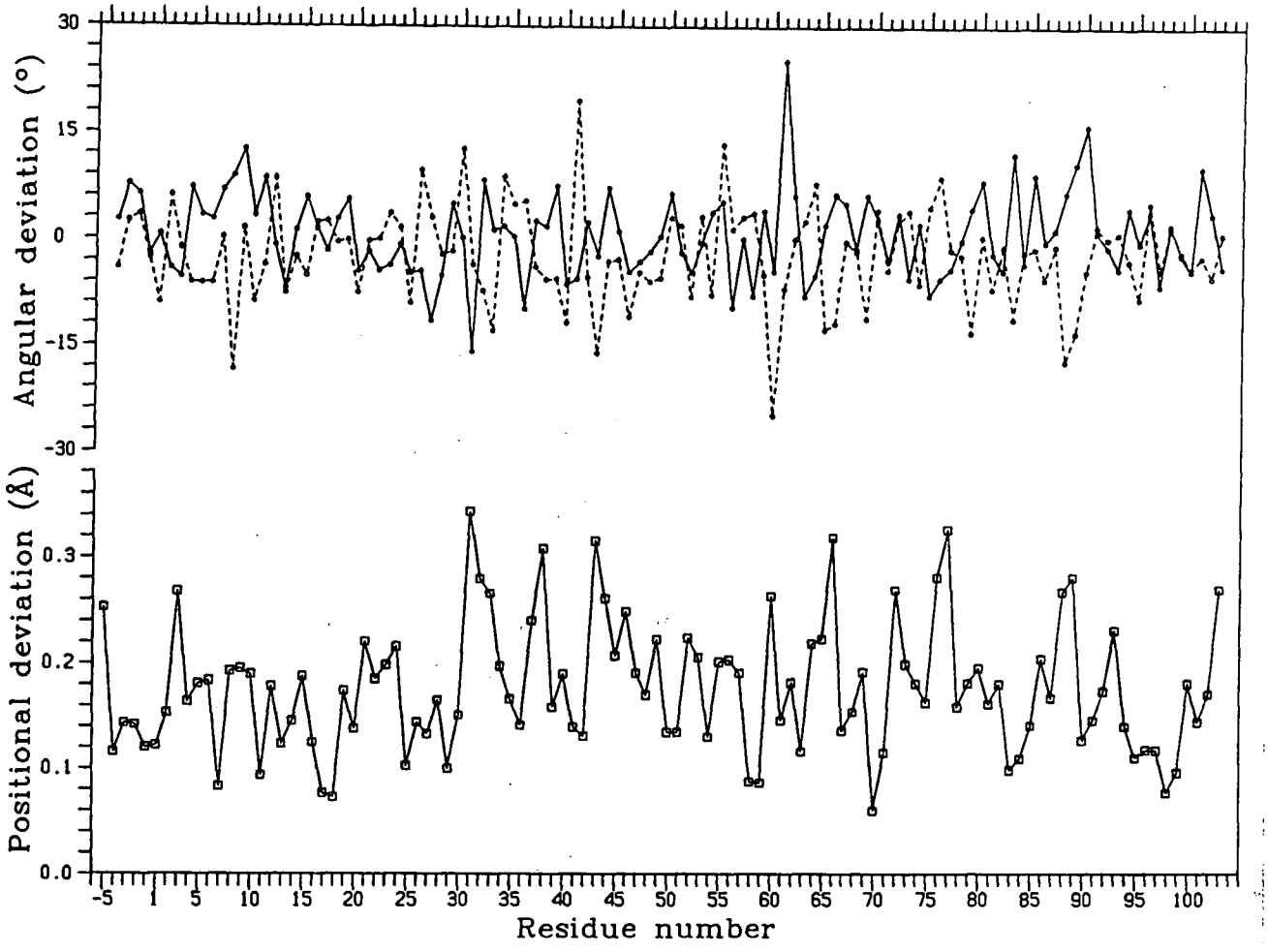

Figure IX.2. Comparison of the polypeptide backbone conformations of the Ala38 variant and wild-type iso-1-cytochromes *c*. The lower plot indicates the average deviations observed between the positions of main chain atoms in corresponding residues of the wild-type and Ala38 variant iso-1-cytochromes *c*. The upper plot indicates the deviations in the main chain conformational torsion angles ϕ (—) or ψ (---) of corresponding residues of the two proteins.

residues. The overall r.m.s. difference between positions of corresponding main chain atoms is 0.20 Å, and the overall r.m.s. differences between phi and psi angles are 6.2° and 7.2°, respectively.

The most notable positional shifts observed in the vicinity of the substitution site are associated with rearrangements within the network of hydrogen bonds formed by the surrounding polypeptide chain (see Figure IX.3). Wat 121, an internal water molecule formerly hydrogen bonded to the guandinium group of Arg38, and the side chain of Tyr48 shift closer together to

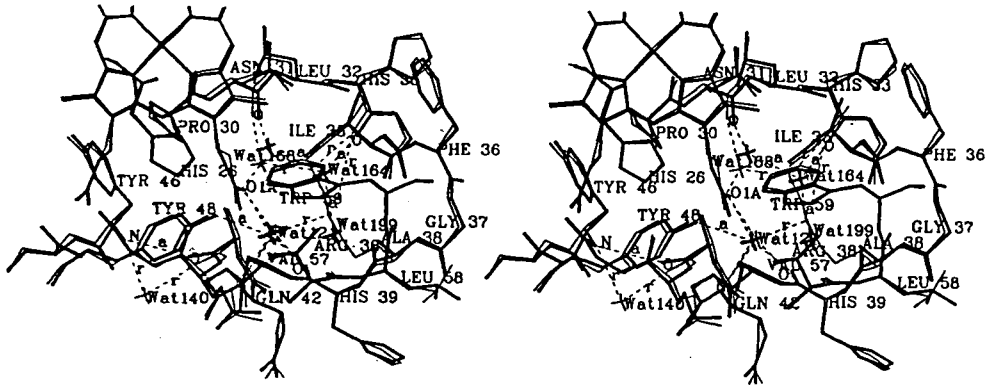

Figure IX.3. Stereo drawing of the region around residue 38 in the Ala38 variant (thick lines) and wild-type (thin lines) iso-1-cytochromes *c*. The heme groups and the polypeptide chain of residues 30 through 48 and 56 through 59 of the two proteins are shown superimposed. Also shown are the water molecules Wat121 and 168 in both proteins; and Wat164 and 199 in the Ala38 variant, which fill the space vacated by the side chain of Arg38. Hydrogen bond interactions are shown as thin dashed lines, with those unique to either the Ala38 variant ('a') or the wild-type protein ('r') labelled.

allow the formation of a Tyr48 OH-Wat121 hydrogen bond. An adjustment in the geometry of the nearby β-turn (residues 43 to 46) is also observed. In the wild-type molecule, this turn is distorted as a bridging water molecules mediates the interaction between Ala43 N and Tyr46 O, whereas a direct hydrogen bond between these groups occurs in the Ala38 variant. It was proposed earlier that packing of the side chain of Tyr46 against Pro30 is partly responsible for the distortion of this β-turn in the wild-type protein (see Section III.B.3). The observation of Moore and Williams (1980b) that the spatially adjacent aromatic rings of Tyr46 and Tyr48 interact conformationally suggests that the adjustment in the β-turn is brought about by the movement of the tyrosyl ring of Tyr48, which in turn influences the conformation of Tyr46. The observed adjustments in the hydrogen bonding pattern suggest that there is significant flexibility in the formation of the hydrogen bonding interactions within the bottom loop of the molecule. This flexibility is also evident in the variations in conformation observed for the side chain of Arg38 itself in yeast iso-1-, tuna and rice cytochromes *c* (see Section IV.B.2 and Figure IV.7).

Wallace's group (Proudfoot *et al.*, 1986; also see Table I.2) has shown that amino acid replacements at Arg38 lower the overall stability of the cytochrome *c* molecule. Based on the

observation that the side chain of Arg38 extends upward from the bottom loop of the polypeptide chain to form hydrogen bonds with groups in the main body of the molecule, Proudfoot and Wallace (1987) have suggested that the role of this residue is in conformational stabilization of this bottom loop. Replacement of Arg38 might therefore be expected to perturb the structure of this loop (residues 36 to 59). However, apart from the localized deviations described above, the overall folding of these residues does not differ significantly between the Ala38 variant and wild-type iso-1-cytochromes *c* (see Figure IX.2). Nor is there any substantial dissimilarity between the two proteins in the flexibility of this loop of polypeptide chain (as assessed by atomic temperature factors). Thus comparison of the crystal structures of the Ala38 variant and wild-type iso-1-cytochromes *c* gives no obvious indication that Arg38 directly influences the conformational properties of the bottom loop. It should be noted, however, that structural destabilization effected by replacement of Arg38 may arise from an increased frequency of transient unfolding. This possibility cannot be assessed in a crystallographic study, which cannot provide information on such high frequency processes (Ringe and Petsko, 1985).

2. Propionic acid group of heme pyrrole ring A

The heme groups of the Ala38 variant and wild-type iso-1-cytochromes *c* do not differ significantly in either overall positioning or conformation. The mean difference in position of the 43 heme atoms is only 0.18 Å, and of the 5 atoms of the rear propionic acid group is 0.26 Å. Thus, the positioning of this propionic acid group of the heme is not significantly affected by the absence of the interaction between this group and the guanidinium group of Arg38. This result is in agreement with NMR studies which show that there are no differences between the Ala38 variant and wild-type iso-1-cytochromes *c* in NOE's involving this propionic acid group (Cutler *et al.*, 1989). The hydrogen bonding interactions formed by this propionic acid in the Ala38 variant are identical to those occurring in the wild-type molecule (Figure III.11).

3. Solvent structure in the vicinity of the substitution site

The replacement of Arg38 by an alanine residue has a marked effect on the solvent structure in the vicinity of the substitution site (see Figure IX.3). Most notably, the loss of the bulky arginyl side chain renders the adjacent propionic acid group of the heme accessible to the external solvent medium (Figure IX.4). In addition, several other groups normally buried in the interior of the protein molecule become exposed to solvent, including the side chains of Ile35 and Tyr48. Not unexpectedly, the number of bound water molecules occurring in this region of the molecule is increased in the Ala38 variant. The depression within the molecular surface formerly occupied by the Arg38 side chain is filled by two water molecules, Wat164 and Wat199, which are hydrogen bonded to each other. The two (formerly) internal water molecules (Wat121 and Wat168) which in the wild-type structure form hydrogen bonds to both the rear propionic acid oxygen O1A and the Arg38 guanidinium group are retained in the Ala38 structure. Wat121 and Wat168 form the same hydrogen bonds to protein groups in both the Ala38 variant and wild-type molecules, with the exception of those to Arg38. In the absence of these latter hydrogen bonds,

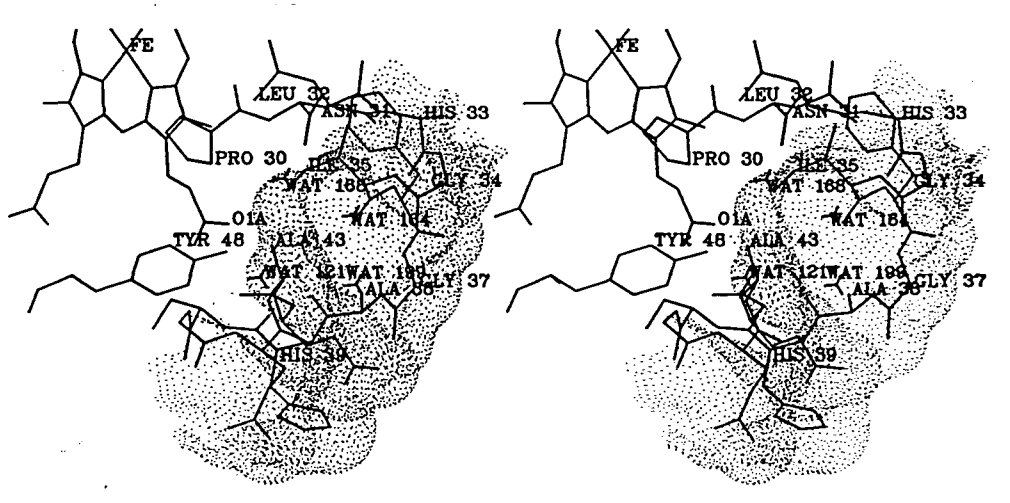

Figure IX.4. Molecular surface in the region of residue 38 in the Ala38 variant iso-1-cytochrome *c*. The dot surface represents the region of the protein molecule which is accessible to solvent, and was calculated (Connolly, 1983) using a probe sphere of radius 1.4 Å. Note the cavity occupied by the water molecules Wat121, 164, 168 and 199. The inner surface at the base of this cavity belongs to the O1A atom of the rear propionic acid group of the heme. In the wild-type protein, the cavity is largely filled by the side chain of Arg38.

Wat168 shifts closer to one of its hydrogen bonding partners Asn31 O, and Wat121 undergoes a slight movement to enable it to hydrogen bond to Tyr48 OH. Wat168 also forms an additional hydrogen bond to Wat164, whose position coincides approximately with that of Arg38 NH1 in the wild-type protein.

C. Influence of Arg38 on reduction potential

The various position 38 variants of iso-1-cytochrome *c* all have lowered reduction potentials (see Section I.G), thus indicating that Arg38 has an important role in establishing a high reduction potential in cytochrome *c*. Cutler *et al.* (1989) have noted that for the individual variants, the decrease in reduction potential correlates with the decrease in the electron-withdrawing capacity (or basicity) of the side chain group introduced at position 38. This result indicates that the primary mechanism by which Arg38 controls the reduction potential is electrostatic: the guanidinium group of this residue places a positive charge ~12 Å from the heme iron, thereby destabilizing the positive charge resident on a ferric heme iron and favoring the reduced state of the protein. This is unlikely to be the only mechanism, as evidenced by the differing reduction potentials of the Arg38 and Lys38 proteins, in which the side chains of residue 38 would be positively charged, as well as of the Leu38 and Ala38 proteins, in which the side chains are uncharged. It is notable that as an approximation, the reduction potential of iso-1-cytochrome *c* decreases as the volume of the side chain at position 38 decreases.

Proudfoot and Wallace (1987) have proposed that the extensive network of hydrogen bonding involving the side chain of Arg38 serves to stabilize the conformation of the bottom loop of the cytochrome *c* molecule, and consequently to induce a cooperative tightening of the global structure. They suggest that amino acid substitutions at residue 38, through disruption of the global structural integrity of the molecule, may directly effect an increase in the overall solvent accessibility to the heme, and thus in the overall polarity of the heme environment. Their proposal is in agreement with the observed pattern of decrease in reduction potential in the series of position 38 variants. The decrease is largest for those variants in which the side chain at position 38 lacks the capacity to form hydrogen bonds homologous to those formed by the Arg38

side chain. In addition, as discussed above loss of the bulky arginyl side chain creates a large void in the surface of the protein molecule, which would likely be structurally destabilizing. The degree of destabilization would be expected to be correlated with the size of the cavity created. These considerations are consistent with the reduction potential decreasing as the volume of the side chain at position 38 decreases.

The intricate hydrogen bonding network formed by the side chain of Arg38 may function in an additional, previously unemphasized, manner to effect a high reduction potential in cytochrome *c*. The network may serve to delocalize the negative charge on the ionized heme propionate group onto the resonance-stabilized guanidinium group of Arg38. Lowered electronegativity of the heme propionate would favor the reduced state of the heme iron. These considerations may explain the preference for an arginine at position 38. The groups within the network are held in a fixed orientation, and thus may be particularly suited to this role (Barlow and Thornton, 1983; Singh *et al.*, 1987).

As discussed earlier, the positioning of the internal water molecule Wat166 at the lower left of the heme crevice is responsive to an electropositive charge at the heme iron. Consistent with the ~50mV drop in reduction potential of the Ala38 variant, the positional shift of Wat166 in this protein is very similar to that observed in the Gly82 and Ser82 variants, which also have reduction potentials ~50mV lower than normal. Wat166 in the Ala38 variant has moved 0.8 Å with respect to its position in the wild-type protein, thereby decreasing the distance between it and the heme iron atom from 6.4 to 5.8 Å.

X. SUMMARY

The three-dimensional atomic structure of yeast (*Saccharomyces cerevisiae*) iso-1-cytochrome *c* has been determined through molecular replacement techniques. This structure has been refined against X-ray diffraction data with $F \geq 2\text{-}3\sigma_F$ in the resolution range 6.0-1.23 Å (12513 reflections in total). The crystallographic R-factor is 0.192 and the coordinate accuracy is estimated to be better than 0.18 Å. The atomic model includes 1 sulfate ion and 116 water molecules, 4 of which are buried in the interior of the protein molecule. Yeast iso-1-cytochrome *c* has the typical cytochrome *c* fold, with the polypeptide chain organized into five α-helical segments and a series of loop structures which serve to enclose almost completely the heme prosthetic group within a hydrophobic pocket. The conformational torsion angles occurring within secondary structural elements, the good agreement with ideality in conformations adopted by side chains, the high degree of saturation of hydrogen bonding groups, and the degree of solvation of the molecular surface by water molecules are all consistent with patterns observed in other high resolution protein structures. Notable structural features of yeast iso-1-cytochrome *c* are the localized distortion of secondary structural elements by intramolecular interactions occurring within the final folded state of the protein molecule, and the structural rigidity of the interior of the molecule. The accurate structural determination for yeast iso-1-cytochrome *c* provides a firm basis for both the analysis of structural perturbations and the interpretation of altered functional properties of the numerous variant proteins available.

The structure of yeast iso-1-cytochrome *c* has been compared in detail with the high resolution structures of the related tuna and rice cytochromes *c*. Although localized conformational differences caused by amino acid replacements do occur, the fold of the polypeptide backbone, intramolecular hydrogen bonding, and the conformation of side chains are found to be very similar in the three proteins. The strongly conserved packing of polypeptide chain groups against the heme moiety is particularly striking and is consistent with an important role of the heme in determining the fold of the surrounding polypeptide chain. This packing of protein groups, in establishing the identity of the heme environment, is also likely a determinant of the functional

properties of the heme moiety.

X-ray crystallographic techniques have also been used to determine the structures of several variant iso-1-cytochromes *c* bearing single-site amino acid substitutions at Phe82 and Arg38. Structural differences between the variant and wild-type proteins were identified initially from difference Fourier maps. The structures of the various variant iso-1-cytochromes *c* were subsequently refined at nominal resolutions in the range 2.8 to 1.76 Å, with R-factors in the range 0.15 to 0.20. Comparison of the structure of wild-type iso-1-cytochrome *c* with those of the position 82 variant proteins shows that substitutions of serine, glycine, tyrosine or isoleucine for Phe82 are accommodated by slight alterations in the conformation of the polypeptide chain, but cause no gross rearrangements in the overall conformation of the cytochrome *c* molecule. These results are consistent with the similarity in intrinsic CD spectra of the wild-type and variant proteins, which indicates that the global secondary structure in the variant proteins is unaltered, and also with NMR studies of the variants, which show that chemical shifts occur only for protein groups in the direct vicinity of the substitution site. Nonetheless, conspicuous structural perturbations in the neighborhood of the substituted side chain are evident in all of the variant proteins. In wild-type iso-1-cytochrome *c*, the phenyl ring of Phe82 is positioned adjacent and approximately parallel to the heme group, and occupies a hydrophobic cavity within the heme crevice. In the Ser82 variant, a solvent channel is created which extends from the surface of the molecule down into the heme crevice. In the Gly82 variant, the polypeptide backbone refolds into the space formerly occupied by the phenyl ring of Phe82. The phenolic ring of Tyr82 cannot be entirely accommodated within the pocket normally occupied by a phenyl ring, and thus is displaced further toward the exterior of the molecule. In the Ile82 variant, branching at the β-carbon of residue 82 generates steric contacts between the isoleucine side chain and adjacent groups in the heme crevice. In wild-type iso-1-cytochrome *c*, the side chain of Arg38 is involved in a network of hydrogen bonding, through which it interacts with one of the buried propionic acid groups of the heme. In the Ala38 variant, the propionic acid group is no longer occluded from external solvent, and a slight reorganization of the hydrogen bonding network occurs.

Comparison of the structures of reduced yeast iso-1- and tuna cytochromes *c* with those of oxidized tuna and rice cytochromes *c* shows that significant oxidation-state dependent conformational changes occur only in the vicinity of an internal water molecule located within the heme crevice below the Met80 ligand. These results are in agreement with those of Takano and Dickerson (1981b), who considered only the two redox forms of tuna cytochrome *c*. The same conformational changes are also observed upon structural comparison of the wild-type reduced iso-1-cytochrome *c* with the Gly82(Thr102) variant, whose structure was determined in the oxidized form. It is notable, however, that the positioning of the sensitive water molecule (Wat166 in the iso-1-cytochrome *c* structures) varies not only with oxidation state, but also with reduction potential. Within the series of position 82 variants whose structures have been determined in the reduced state, a decrease in reduction potential is correlated with the closer proximity of Wat166 to the heme iron. The movement of Wat166 is interpreted as arising from an attraction to an increased electropositive charge residing on the heme iron.

The altered functional properties of the position 82 variant proteins have been interpreted with respect to the structural perturbations caused by the amino acid substitutions. The drop in reduction potential, most notably for the Ser82 and Gly82 variants, can be explained by an increased polarity of heme environment. This increase in polarity can be attributed to the accessibility of the heme pocket to either solvent (in the Ser82 variant) or polar groups of the polypeptide backbone (in the Gly82 variant). These results thus confirm that residue 82 directly influences the characteristics of the heme environment, and also that the dielectric constant of the heme environment is an important determinant of the heme reduction potential. The decreased structural stability of the position 82 variants is indicated by their susceptibility to denaturation during crystallization trials, by their lowered pK_a's toward alkaline isomerization, and by the resemblance of their CD spectra to spectra of thermally denatured wild-type protein. That amino acid substitutions at Phe82 decrease the stability of the cytochrome *c* molecule is typical of numerous other substitutions at other conservative residues. These results indicate that the forces maintaining the tertiary structure of the cytochrome *c* molecule are highly cooperative and finely

tuned, with many amino acid residues contributing to stabilization of the native conformation of the protein. For the position 82 variants, the reduced stability of the heme crevice is likely due to the disruption of stabilizing packing forces formed by the Phe82 phenyl ring within its hydrophobic cavity. The lowered activity, in comparison to the wild-type protein and the Tyr82 variant, for electron transfer with $^{+}$Zn-cytochrome *c* peroxidase is attributed to the loss of an aromatic group positioned adjacent to the heme group. The coupled π-electron systems of these two groups may serve as a direct conduit for movement of the transferred electron. Alternatively, the presence of an aromatic group at residue 82 may induce an electronic distribution within the π-orbitals of the heme which is particularly favorable for the electron transfer reaction. The electron transfer reaction between iso-1-cytochrome *c* and cytochrome *c* peroxidase may be further inhibited by the altered surface topography of the variant proteins (particularly the Gly82, Tyr82 and Ile82 variants), which hinders formation of a productive electron transfer complex.

In conclusion, these results suggest that the invariant Phe82 contributes in at least three ways to the proper functioning of cytochrome *c*. It has an important structural role in maintaining the integrity of the heme crevice, in preventing the intrusion of polar groups into the heme pocket, and in establishing the appropriate heme environment. The aromaticity of the phenyl ring of Phe82 is important for efficient movement of an electron to and from the heme of cytochrome *c*. Finally, Phe82 may have a role in forming the appropriate intermolecular interactions with enzymic redox partners of cytochrome *c*. A phenylalanine at position 82 is clearly optimally suited for these roles in terms of its steric size and its aromatic and non-polar chemical nature.

XI. REFERENCES

Adar, F. (1979). In *The Porphyrins* (Dolphin, D., ed.), Vol. III, Part A, pp. 167–209, Academic Press, New York.

Alber, T.A., Bell, J.A., Dao-pin, S., Nicholson, H., Wozniak, J.A., Cook, S.P. & Matthews, B.W. (1987). *Science* **239**, 631–635.

Alber, T.A., Dao-pin, S., Wilson, K., Wozniak, J.A., Cook, S.P. & Matthews, B.W. (1987). *Nature (London)* **330**, 41–46.

Armstrong, G.D., Chambers, J.A. & Sykes, A.G. (1986). *J. Chem. Dalton Trans.* **182**, 755–758.

Arnott, S. & Dover, S.D. (1967). *J. Mol. Biol.* **30**, 209–212.

Baker, E.N. & Hubbard, R.E. (1984). *Prog. Biophys. Mol. Biol.* **44**, 97–179.

Bandekar, J. & Krimm, S. (1985). *Int. J. Peptide Protein Res.* **26**, 407–415.

Barlow, D.J. & Thornton, J.M. (1983). *J. Mol. Biol.* **168**, 867–885.

Bechtold, R. & Bosshard, H.R. (1985). *J. Biol. Chem.* **260**, 5191–5200.

Bechtold, R., Kuehn, C., Lepre, C. & Isied, S.S. (1986). *Nature (London)* **320**, 286–288.

Bernstein, F.C., Koetzle, T.F., Williams, G.J.B., Meyer, E.F., Bruce, M.D., Rodgers, J.R., Kennard, O., Shimanouchi, T. & Tasumi, M. (1977). *J. Mol. Biol.* **112**, 535–542.

Bhatia, G.E. (1981). *Ph. D. Thesis, Univ. of California at San Diego.*

Blake, C.C.F., Pulford, W.C.A., & Arthymiuk, P.J. (1983). *J. Mol. Biol.* **167**, 693–723.

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

Boon, P.J., Van Raay, A.J.M., Tesser, G.I. & Nivard, R.J.F. (1979). *FEBS Lett.* **108**, 131–135.

Boots, H.A. & Tesser, G.I. (1987). In *Peptides 1986* (Theodoropoulos, D., ed.), pp. 211–214, Walter de Gruyter, Berlin.

Boswell, A.P., Moore, G.R., Williams, R.J.P., Wallace, C.J.A., Boon, P.J., Nivard, R.J.F. & Tesser, G.I. (1981). *Biochem. J.* **193**, 493–502.

Bott, R.G. & Sarma, R. (1976). *J. Mol. Biol.* **106**, 1037–1048.

Brems, D.N., Cass, R. & Stellwagen, E. (1982). *Biochemistry* **21**, 1488–1493.

Bryant, C., Stottmann, J.M. & Stellwagen, E. (1985). *Biochemistry* **24**, 3459–3464.

Burley, S.K. & Petsko, G.A. (1988). *Adv. Prot. Chem.* **39**, 125–189.

Butler, J., Chapman, S.K., Davies, D.M., Sykes, A.G., Speck, S.H., Osheroff, N. & Margoliash, E. (1983). *J. Biol. Chem.* **258**, 6400–6404.

Chambers, J.L. & Stroud, R.M. (1979). *Acta Crystallogr.* **B35**, 1861–1874.

Chothia, C. & Lesk, A.M. (1985). *J. Mol. Biol.* **182**, 151–158.

Churg, A.K. & Warshel, A. (1986). *Biochemistry* **26**, 1675–1681.

Connolly, M.L. (1983). *Science* **221**, 709–713.

Corradin, G. & Harbury, H.A. (1974). *Biochem. Biophys. Res. Commun.* **61**, 1400–1406.

Corthesy, B.E. & Wallace, C.J.A. (1986). *Biochem. J.* **236**, 359–364.

Cowgill, R.W. & Clark, W.M. (1952). *J. Biol. Chem.* **198**, 33–61.

Creighton, T.E. (1984). *Proteins: structures and molecular principles*, W.H. Freeman, New York.

Crowther, R.A. (1973). In *The Molecular Replacement Method* (Rossman, M.G., ed.), pp. 173–178, Gordon and Breach, New York.

Crowther, R.A. & Blow, D.M. (1967). *Acta. Cryst.* **23**, 544–548.

Cutler, R.L., Pielak, G.J., Mauk, A.G. & Smith, M. (1987). *Protein Eng.* **1**, 95–99.

Cutler, R.L., Davies, A.M., Creighton, S., Warshel, A., Moore, G.R., Smith, M. & Mauk, A.G. (1989). *Biochemistry* **28**, 3188–3197.

Das, G., Hickey, D.R., Principio, L., Conklin, K.T., Short, J., Miller, J.R., McLendon, G. & Sherman, F. (1988). *J. Biol. Chem.* **263**, 18290–18297.

Das, G., Hickey, D.R., McLendon, D., McLendon, G. & Sherman, F. (1989). *Proc. Natl. Acad. Sci. USA* **86**, 496–499.

Davis, L.A., Schejter, A. & Hess, G.P. (1974). *J. Biol. Chem.* **249**, 2624–2632.

Dickerson, R.E. (1980). In *Structure and Evolution of Enzymes* (Sigman, D.S. & Brazier, M.A.B., eds.), pp. 171–202, Academic Press, New York.

Dickerson, R.E. & Timkovich, R. (1975). In *The Enzymes* (Boyer, P.D., ed.), 3rd ed., Vol. 11, Part A, pp. 397–547, Academic Press.

Dickerson, R.E., Takano, T., Eisenberg, D., Kallai, O.B., Samson, L., Cooper, A. & Margoliash, E. (1971). *J. Biol. Chem.* **246**, 1511–1532.

Dickerson, R.E., Timkovich, R. & Almassy, R.J. (1976). *J. Mol. Biol.* **100**, 473–491.

DiMaria, P., Polastro, E., DeLange, R.J., Kim, S. & Paik, W.K. (1979). *J. Biol. Chem.* **254**, 4645–4652.

Dixon, D.W. (1988). In *Mechanistic Principles of Enzyme Activity* (Liebman, J.F. & Greenberg, A., eds.), pp. 169–224, VCH Publishers, New York.

Dumont, M., Ernst, J.F. & Sherman, F. (1988). *J. Biol. Chem.* **263**, 15928–15937.

Eden, D., Matthew, J.B., Rosa, J.J. & Richards, F.M. (1982). *Proc. Natl. Acad. Sci. USA* **79**,

815–819.

Edsall, J.T. & McKenzie, H.A. (1983). *Adv. Biophys.* **16**, 53–183.

Eley, C.G.S. & Moore, G.R. (1983). *Biochem. J.* **215**, 11–21.

Erecinska, M. & Vanderkooi, J.M. (1978). *Methods Enzymol.* **53**, 165–181.

Ernst, J.F., Hampsey, D.M., Stewart, J.W., Rackovsky, S., Goldstein, D. & Sherman, F. (1985). *J. Biol. Chem.* **260**, 13225–13236.

Errede, B.J. & Kamen, M.D. (1976). *Biochemistry* **17**, 1015–1027.

Everest, A.M., Wallin, S.A., Stemp, E.D.A., Nocek, J.M., Mauk, A.G. & Hoffman, B.M. (1990). *J. Am. Chem. Soc.* **submitted.**

Falk, J.E. (1964). *Porphyrins and Metalloporphyrins*, Elsevier, Amsterdam.

Falk, K.E., Jovell, P.A. & Angstrom, J. (1981). *Biochem. J.* **193**, 1021–1024.

Farooqui, J., DiMaria, P., Kim, S. & Paik, W.K. (1981). *J. Biol. Chem.* **256**, 5041–5045.

Feng, Y., Roder, H. & Englander, S.W. (1990). *Biochemistry* **29**, 3494–3504.

Ferguson–Miller, S., Brautigan, D.L. & Margoliash, E. (1976). *J. Biol. Chem.* **251**, 1104–1115.

Ferguson–Miller, S., Brautigan, D.L. & Margoliash, E. (1978). *J. Biol. Chem.* **253**, 149–159.

Ferguson–Miller, S., Brautigan, D.L. & Margoliash, E. (1979). In *The Porphyrins* (Dolphin, D., ed.), Vol. VII, Part B, pp. 149–240, Academic Press, New York.

Fetrow, J.S., Cardillo, T.S. & Sherman, F. (1989). *Proteins: Struc. Funct. Genet.* **6**, 372–381.

Fisher, W.R., Taniuchi, H. & Anfinsen, C.B. (1973). *J. Biol. Chem.* **248**, 3188–3195.

Fitzgerald, P.M.D. (1988). *J. Appl. Cryst.* **23**, 273–278.

Frey, M., Sieker, L., Payan, F. Haser, R., Bruschi, M., Pepe, G. & LeGall, J. (1987). *J. Mol. Biol.* **197**, 525–541.

Gadsby, P.M.A., Peterson, J., Foote, N., Greenwood, C. & Thomson, A.J. (1987). *Biochem. J.* **246**, 43–54.

George, P., Hanania, G.I.H. & Eaton, W.A. (1966). In *Hemes and Hemoproteins* (Chance, B., Estabrook, R.W. & Yonetani, T., eds.), pp. 267–270, Academic Press, New York.

Hampsey, D.M., Das, G. & Sherman, F. (1988). *FEBS Lett.* **231**, 275–283.

Hampsey, D.M., Das, G. & Sherman, F. (1986). *J. Biol. Chem.* **261**, 3259–3271.

Harbury, H.A. (1977). In *Semisynthetic Peptides and Proteins* (Offord, R.E. and DiBello, C., eds.), pp. 73–89, Pergamon Press, Oxford.

Harbury, H.A. & Marks, R.H.L. (1973). In *Inorganic Biochemistry* (Eichhorn, G., ed.), pp. 902–954,

Elsevier, Amsterdam.

Harris, D.E. & Offord, R.E. (1977). *Biochem. J.* **161**, 21–25.

Hazzard, J.T., Poulos, T.L. & Tollin, G. (1987). *Biochemistry* **26**, 2836–2848.

Hazzard, J.T., McLendon, G., Cusanovich, M.A., Das, G., Sherman, F. & Tollin, G. (1988a). *Biochemistry* **27**, 4445–4451.

Hazzard, J.T., McLendon, G., Cusanovich, M.A. & Tollin, G. (1988b). *Biochem. Biophys. Res. Commun.* **151**, 429–434.

Hendrickson, W.A. (1985). *Methods Enzymol.* **115**, 252–270.

Hendrickson, W.A. & Konnert, J. (1981). In *Biomolecular Structure, Function, Conformation and Evolution* (Srinivasan, R., ed.) Vol. 1, pp. 43–57, Pergamon Press, Oxford.

Hendrickson, W.A. & Ward, K.B. (1976). *Acta Crystallogr.* **A32**, 778–780.

Hickey, D.R., McClendon, G. & Sherman, F. (1988). *J. Biol. Chem.* **263**, 18298–18305.

Hill, G.C. & Pettigrew, G.W. (1975). *Eur. J. Biochem.* **57**, 265–270.

Ho, P.S., Sutoris, C., Liang, N., Margoliash, E. & Hoffman, B.M. (1986). *J. Am. Chem. Soc.* **107**, 1070–1071.

Hoffman, B.M. & Ratner, M.A. (1987). *J. Am. Chem. Soc.* **109**, 6237–6242.

Hol, W.G.J. (1985). *Prog. Biophys. Mol. Biol.* **45**, 149–195.

Holzschu, D., Principio, L., Conklin, K.T., Hickey, D.R., Short, J., Rao, R., McLendon, G. & Sherman, F. (1987). *J. Biol. Chem.* **262**, 7125–7131.

Honzatko, R.B., Hendrickson, W.A. & Love, W.E. (1985). *J. Mol. Biol.* **184**, 147–164.

Howell, E.E., Villafranca, J.E., Warren, M.S., Oatley, S.J. & Kraut, J. (1986). *Science* **231**, 1123–1128.

Jack, A. & Levitt, M. (1978). *Acta Crystallogr.* **A34**, 931–935.

James, M.N.G. & Sielecki, A.R. (1983). *J. Mol. Biol.* **163**, 299–361.

Jones, T.A. (1978). *J. Appl. Crystallogr.* **11**, 268–272.

Juillerat, M. & Homandberg, G.A. (1981). *Int. J. Peptide Protein Res.* **18**, 335–342.

Juillerat, M.A. & Taniuchi, H. (1986). *J. Biol. Chem.* **261**, 2697–2711.

Kamphuis, I.G., Kalk, K.H., Swarte, M.B.A. & Drenth, J. (1984). *J. Mol. Biol.* **179**, 233–256.

Kamphuis, I.G., Drenth, J. & Baker, E.N. (1985). *J. Mol. Biol.* **182**, 317–329.

Kang, C.H., Ferguson-Miller, S. & Margoliash, E. (1977). *J. Biol. Chem.* **252**, 919–926.

Karplus, P.A. & Schulz, G.E. (1985). *Naturwissenschaften* **72**, 212–213.

Karplus, P.A. & Schulz, G.E. (1987). *J. Mol. Biol.* **195**, 701–729.

Kassner, R.J. (1972). *Proc. Natl. Acad. Sci. USA* **69**, 2263–2267.

Kassner, R.J. (1973). *J. Am. Chem. Soc.* **95**, 2674–2677.

Koppenol, W.H. & Margoliash, E. (1982). *J. Biol. Chem.* **257**, 4426–4437.

Korszun, Z.R. & Salemme, F.R. (1977). *Proc. Natl. Acad. Sci. USA* **74**, 5244–5247.

Koul, A.K., Wasserman, G.F. & Warme, P.K. (1979). *Biochem. Biophys. Res. Commun.* **89**, 1253–1259.

Kreil, G. (1963). *Hoppe-Seyler's Z. Physiol. Chem.* **334**, 153–166.

Kreil, G. (1965). *Hoppe-Seyler's Z. Physiol. Chem.* **340**, 86–87.

Lederer, F., Simon, A.M. & Verdiere, J. (1972). *Biochem. Biophys. Res. Commun.* **47**, 55–58.

Leszczynski, J.S. & Rose, G.D. (1986). *Science* **234**, 849–855.

Leung, C.J., Nall, B.T. & Brayer, G.D. (1989). *J. Mol. Biol.* **206**, 783–785.

Liang, N., Kang, C.H., Margoliash, E., Ho, P.S. & Hoffman, B.M. (1986). *J. Am. Chem. Soc.* **108**, 4665–4666.

Liang, N., Pielak, G.J., Mauk, A.G., Smith, M. & Hoffman, B.M. (1987). *Proc. Natl. Acad. Sci. USA* **84**, 1249–1252.

Liang, N., Mauk, A.G., Pielak, G.J., Johnson, J.A., Smith, M. & Hoffman, B.M. (1988). *Science* **240**, 311–313.

Louie, G.V. & Brayer, G.D. (1989). *J. Mol. Biol.* **210**, 313–322.

Louie, G.V., Hutcheon, W.L.B. & Brayer, G.D. (1988a). *J. Mol. Biol.* **199**, 295–314.

Louie, G.V., Pielak, G.J., Smith, M. & Brayer, G.D. (1988b). *Biochemistry* **27**, 7870–7876.

Ludwig, M.L., Pattridge, K.A., Powers, T.B., Dickerson, R.E. & Takano, T. (1982). In *Electron Transport and Oxygen Utilization* (Ho, C., ed.), pp. 27–32, Elsevier, Amsterdam.

Lum, V.R., Brayer, G.D., Louie, G.V., Smith, M. & Mauk, A.G. (1987). In *Protein Structure, Folding and Design* (Oxender, D., ed.), pp. 143–150, Alan R. Liss, New York.

Luntz, T.L., Schejter, A., Garber, E.A.E. & Margoliash, E. (1989). *Proc. Natl. Acad. Sci. USA* **86**, 3524–3528.

Luzzati, V. (1952). *Acta. Crystallogr.* **5**, 802–810.

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

Marcus, R.A. & Sutin, N. (1985). *Biochim. Biophys. Acta* **811**, 265–322.

Margoliash, E. & Bosshard, H.R. (1983). *Trends Biochem. Sci.* **35**, 316–320.

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

Margoliash, E., Ferguson-Miller, S., Tulloss, J., Kang, C.H., Feinberg, B.A., Brautigan, D.L. & Morrison, M. (1973). *Proc. Natl. Acad. Sci. USA* **70**, 3245–3249.

Mathews, F.S. (1985). *Prog. Biophys. Mol. Biol.* **45**, 1–56.

Matsuura, Y., Hata, Y., Yamaguchi, T., Tanaka, N. & Kakudo, M. (1979). *J. Biochem. (Tokyo)* **85**, 729–737.

Matsuura, Y., Takano, T. & Dickerson, R.E. (1982). *J. Mol. Biol.* **156**, 389–409.

Matthews, B.W. (1968). *J. Mol. Biol.* **33**, 491–497.

Matthews, B.W. (1977). In *The Proteins* (Neurath, H. & Hill, R.L., eds.), Vol. 3, 3rd edit., pp. 403–590, Academic Press, New York.

Matthews, B.W., Nicholson, H. & Becktel, W.J. (1987). *Proc. Natl. Acad. Sci. USA* **84**, 6663–6667.

Mauk, M.R., Mauk, A.G., Weber, P.C. & Matthew, J.B. (1986). *Biochemistry* **25**, 7085–7091.

Mayo, S.L., Ellis, W.R., Crutchley, R.J. & Gray, H.B. (1986). *Science* **233**, 948–952.

McGregor, M.J., Islam, S.A. & Sternberg, M.J.E. (1987). *J. Mol. Biol.* **198**, 295–310.

McLendon, G. (1988). *Acc. Chem. Res.* **21**, 160–167.

McPherson, A. (1982). *The preparation and analysis of protein crystals.* John Wiley and Sons, New York.

Meade, T.J., Gray, H.B. & Winkler, J.R. (1989). *J. Am. Chem. Soc.* **111**, 4353–4356.

Meyer, T.E. & Kamen, M.D. (1982). *Adv. Protein Chem.* **35**, 105–212.

Milner-White, E.J., Ross, B.M., Ismail, R., Belhadj-Mostefa, K. & Poet, R. (1988). *J. Mol. Biol.* **204**, 777–782.

Moench, S.J. & Satterlee, J.D. (1989). *J. Biol. Chem.* **264**, 9923–9931.

Moench, S.J., Satterlee, J.D. & Erman, J.E. (1987). *Biochemistry* **26**, 3821–3826.

Moore, G.R. (1983). *FEBS Lett.* **161**, 171–175.

Moore, G.R. & Williams, R.J.P. (1977). *FEBS Lett.* **79**, 229–232.

Moore, G.R. & Williams, R.J.P. (1980a). *Eur. J. Biochem.* **103**, 523–532.

Moore, G.R. & Williams, R.J.P. (1980b). *Eur. J. Biochem.* **103**, 533–541.

Moore, G.R., Harris, D.E., Leitch, F. & Pettigrew, G.W. (1984). *Biochim. Biophys. Acta* **764**,

331-342.

Moore, G.R., Pettigrew, G.W. & Rogers, N.K. (1986). *Proc. Natl. Acad. Sci. USA* **83**, 4998-4999.

Mori, E. & Morita, Y. (1986). *J. Biochem. (Japan)* **87**, 249-266.

Morize, I., Surcouf, E., Vaney, M.C., Epelboin, Y., Buehner, M., Fridlansky, F., Milgrom, E. & Mornon, J.P. (1987). *J. Mol. Biol.* **194**, 725-739.

Motonaga, K., Katano, H. & Nakanishi, K. (1965). *J. Biochem. (Tokyo)* **57**, 29-33.

Myer, Y.P. (1968). *J. Biol. Chem.* **243**, 2115-2122.

Myer, Y.P. & Pande, A. (1979). In *The Porphyrins* (Dolphin, D., ed.), Vol. III, Part A, pp. 271-322, Academic Press, New York.

Myer, Y.P., Thallum, K.K., Pande, J. & Verma, B.C. (1980). *Biochem. Biophys. Res. Commun.* **94**, 1106-1112.

Nall, B.T. & Landers, T.A. (1981). *Biochemistry* **20**, 5403-5411.

Nall, B.T., Zuniga, E.H., White, T.B., Wood, L.C. & Ramdas, L. (1989). *Biochemistry* **28**, 9834-9839.

Narita, K. & Titani, K. (1969). *J. Biochem. (Tokyo)* **65**, 259-267.

Nemethy, G. & Printz, M.P. (1972). *Macromolecules* **5**, 755-758.

Nocera, D.G., Winkler, J.R., Yocom, K.M., Bordignon, E. & Gray, H.B. (1984). *J. Am. Chem. Soc.* **106**, 5145-5150.

North, A.C.T., Phillips, D.C. & Mathews, F.S. (1968). *Acta Crystallogr.* **A24**, 351-359.

Northrup, S.H., Pear, M.R., McCammon, J.A. Karplus, M. & Takano, T. (1980). *Nature (London)* **287**, 659-660.

Northrup, S.H., Pear, M.R., Morgan, J.D. & McCammon, J.A. (1981). *J. Mol. Biol.* **153**, 1087-1109.

Northrup, S.H., Boles, J.O. & Reynolds, C.D. (1988). *Science* **241**, 67-70.

Ochi, H., Hata, Y., Tanaka, N., Kakudo, M., Sakuri, T., Achara, S. & Morita, Y. (1983). *J. Mol. Biol.* **166**, 407-418.

Offord, R.E. (1981). *Semi-synthetic Proteins*, Academic Press, New York.

Offord, R.E. (1987). *Protein Eng.* **1**, 151-157.

Osheroff, N., Borden, D., Koppenol, W.H. & Margoliash, E. (1980). *J. Biol. Chem.* **255**, 1689-1697.

Paik, W.K., Polastro, E. & Kim, S. (1980). In *Current Topics in Cellular Regulation*, (Horecker, B.L. & Stadtman, E.R., eds.), Vol. 16, pp. 87-111, Academic Press, New York.

Paik, W.K., Cho., Y.-B., Frost, B. & Kim, S. (1989). *Biochem. Cell. Biol.* **67**, 602-611.

Pande, J. & Myer, Y.P. (1980). *J. Biol. Chem.* **255**, 11094–11097.

Pearce, L.L., Gartner, A.L., Smith, M. & Mauk, A.G. (1989). *Biochemistry* **28**, 3152–3156.

Peters, P. & Peters, J. (1985). *Biopolymers* **24**, 491–508.

Pettigrew, G.W. & Moore, G.R. (1987). *Cytochrome* c: *Biological Aspects*, Springer–Verlag, Heidelberg.

Pettigrew, G.W., Leaver, J.L., Meyer, T.E. & Ryle, A.P. (1975). *Biochem. J.* **147**, 291–302.

Pflugrath, J.W. & Quiocho, F.A. (1985). *Nature (London)* **314**, 257–260.

Pielak, G.J., Mauk, A.G. & Smith, M. (1985). *Nature (London)* **313**, 152–154.

Pielak, G.J., Oikawa, K., Mauk, A.G., Smith, M., & Kay, C.M. (1986). *J. Am. Chem. Soc.* **108**, 2724–2727.

Pielak, G.J., Concar, D.W., Moore, G.R. & Williams, R.J.P. (1987). *Protein Engineering* **1**, 83–88.

Pielak, G.J., Atkinson, R.A., Boyd, J. & Williams, R.J.P. (1988a). *Eur. J. Biochem.* **177**, 179–185.

Pielak, G.J., Boyd, J., Moore, G.R. & Williams, R.J.P. (1988b). *Eur. J. Biochem.* **177**, 167–177.

Poerio, E., Parr, G.R. & Taniuchi, H. (1986). *J. Biol. Chem.* **261**, 10976–10989.

Polastro, E., Deconinck, M.M., Devogel, M.R., Mailier, E.L., Looze, Y., Schneck, A.G. & Leonis, J. (1978). *FEBS Lett.* **86**, 17–20.

Ponder, J.W. & Richards, F.M. (1987). *J. Mol. Biol.* **193**, 775–791.

Poulos, T.L. & Finzel, B.C. (1984). *Peptide Protein Rev.* **4**, 115–171.

Poulos, T.L. & Kraut, J. (1980). *J. Biol. Chem.* **255**, 10322–10330.

Poulos, T.L., Finzel, B.C. & Howard, A.J. (1987). *J. Mol. Biol.* **195**, 687–700.

Poulos, T.L., Sheriff, S. & Howard, A.J. (1987). *J. Biol. Chem.* **262**, 13881–13884.

Proudfoot, A.E.I. & Wallace, C.J.A. (1987). *Biochem. J.* **248**, 965–967.

Proudfoot, A.E.I., Wallace, C.J.A., Harris, D.E. & Offord, R.E. (1986). *Biochem. J.* **239**, 333–337.

Proudfoot, A.E.I., Rose, K. & Wallace, C.J.A. (1989). *J. Biol. Chem.* **264**, 8764–8770.

Rackovsky, S. & Goldstein, D.A. (1984). *Proc. Natl. Acad. Sci. USA* **81**, 5901–5905.

Rafferty, S.P., Pearce, L.L., Barker, P.D., Guillemette, G., Kay, C.M., Smith, M. & Mauk, A.G. (1990). *Biochemistry* **29**, 9365–9369.

Ramakrishnan, C. & Ramachandran, G.N. (1965). *Biophys. J.* **5**, 909–933.

Ramdas, L. & Nall, B.T. (1986). *Biochemistry* **25**, 6959–6964.

Ramdas, L., Sherman, F. & Nall, B.T. (1986). *Biochemistry* **25**, 6952–6958.

Raphael, A.L. & Gray, H.B. (1989). *Proteins: Struc. Funct. Genet.* **6**, 338–340.

Read, R.J. & James. M.N.G. (1988). *J. Mol. Biol.* **200**, 523–551.

Read, R.J., Fujinaga, M., Sielecki, A.R. & James. M.N.G. (1983). *Biochemistry* **22**, 4420–4433.

Reeke, G.N. Jr. (1984). *J. Appl. Crystallogr.* **17**, 125–130.

Rees, D.C. (1980). *J. Mol. Biol.* **141**, 323–326.

Richardson, J.S. (1981). *Adv. Protein Chem.* **34**, 167–339.

Richardson, J.S. & Richardson, D.C. (1988). *Science* **240**, 1648–1652.

Ridge, J.A., Baldwin, R.L. & Labhardt, A.M. (1981). *Biochemistry* **20**, 1622–1630.

Ringe, D. & Petsko, G.A. (1985). *Prog. Biophys. Mol. Biol.* **45**, 197–235.

Robinson, M.N., Boswell, A.P., Huang, Z.-X., Eley, C.G.S. & Moore, G.R. (1983). *Biochem. J.* **213**, 687–700.

Roder, H., Elove, G.A. & Englander, S.W. (1988). *Nature* **335**, 700–704.

Rodkey, F.L. & Ball, E.G. (1950). *J. Biol. Chem.* **182**, 17–28.

Rossman, M.G. (1973). *The Molecular Replacement Method*, Gordon and Breach, New York.

Saigo, S. (1986). *J. Biochem. (Tokyo)* **100**, 157–165.

Salemme, F.R. (1972). *Arch. Biochem. Biophys.* **151**, 533–539.

Salemme, F.R. (1976). *J. Mol. Biol.* **102**, 563–568.

Salemme, F.R. (1977). *Ann. Rev. Biochem.* **46**, 299–329.

Salemme, F.R., Freer, S.T., Xuong, N.H., Alden, R.A. & Kraut, J. (1973) *J. Biol. Chem.* **248**, 3910–3921.

Santucci, R., Brunori, M. & Ascoli, F. (1987). *Biochim. Biophys. Acta* **914**, 185–189.

Satterlee, J.D., Moench, S.J. & Erman, J.E. (1987). *Biochim. Biophys. Acta* **912**, 87–97.

Scheidt, W.R. (1979). In *The Porphyrins* (Dolphin, D., ed.), Vol. III, Part A, pp. 463–511, Academic Press, New York.

Schlauder, G.S. & Kassner, R.J. (1979). *J. Biol. Chem.* **254**, 4110–4113.

Schoenborn, B.P. (1988). *J. Mol. Biol.* **201**, 741–749.

Schulz, G.E. & Schirmer, R.H. (1979). *Principles of Protein Structure*, Springer-Verlag, New York.

Schweingruber, M.E., Stewart, J.W. & Sherman, F. (1979). *J. Biol. Chem.* **254**, 4132–4143.

Senn, H., Eugster, A. & Wuthrich, K. (1983). *Biochim. Biophys. Acta* **743**, 58–68.

Shellnutt, J.A., Rousseau, D.L., Dethmers, J.K. & Margoliash, E. (1979). *Proc. Natl. Acad. Sci. USA* **76**, 3865–3869.

Shellnutt, J.A., Rousseau, D.L., Dethmers, J.K. & Margoliash, E. (1981). *Biochemistry* **20**, 6485–6497.

Sheriff, S., Hendrickson, W.A. & Smith, J.L. (1987). *J. Mol. Biol.* **197**, 273–296.

Sherwood, D. (1976). *Crystals, X-rays and Proteins*, Halsted Press, New York.

Sherwood, C. & Brayer, G.D. (1985). *J. Mol. Biol.* **185**, 209–210.

Singh, J., Thornton, J.M., Snarey, M. & Campbell, S.F. (1987). *FEBS Lett.* **224**, 161–171.

Smith, H.T., Staudenmayer, N. & Millett, F. (1977). *Biochemistry* **16**, 4971–4974.

Smith, W.W., Burnett, R.H., Darling, G.D. & Ludwig, M.L. (1977). *J. Mol. Biol.* **117**, 195–225.

Smith, M., Leung, D.W., Gillam, S., Astell, C.R., Montgomery, D.L. & Hall, B.D. (1979). *Cell* **16**, 753–761.

Sorrell, T.N., Martin, P.K. & Bowden, E.F. (1989). *J. Am. Chem. Soc.* **111**, 766–767.

Stellwagen, E. (1978). *Nature (London)* **275**, 73–74.

Stout, G.H. & Jensen, L.H. (1968). *X-ray Structure Determination -- A Practical Guide*, Macmillan, New York.

Stuart, D.I. & Phillips, D.C. (1985). *Methods Enzymol.* **115**, 117–142.

Summers, N.L., Carlson, W.D. & Karplus, M. (1987). *J. Mol. Biol.* **196**, 175–198.

Swanson, R., Trus, B.L., Mandel, N., Mandel, G., Kallai, O.B. & Dickerson, R.E. (1977). *J. Biol. Chem.* **252**, 759–775.

Takano, T. & Dickerson, R.E. (1980). *Proc. Natl. Acad. Sci. USA* **77**, 6371–6375.

Takano, T. & Dickerson, R.E. (1981a). *J. Mol. Biol.* **153**, 79–94.

Takano, T. & Dickerson, R.E. (1981b). *J. Mol. Biol.* **153**, 95–115.

Tanaka, N., Yamane, T., Tsukihara, T., Ashida, T. & Kakudo, M. (1975). *J. Biochem. (Tokyo)* **77**, 147–162.

Teeter, M.M. (1984). *Proc. Natl. Acad. Sci. USA* **81**, 6014–6018.

ten Kortenaar, P.B.W., Adams, P.J.H.M. & Tesser, G.I. (1985). *Proc. Natl. Acad. Sci. USA* **82**, 8279–8283.

Theorell, H. & Akesson, A. (1941). *J. Am. Chem. Soc.* **63**, 1804–1820.

Thiessen, W.E. & Levy, H.A. (1973). *J. Appl. Crystallogr.* **6**, 309.

Timkovich, R. (1979). In *The Porphyrins* (Dolphin, D., ed.), Vol. VII, Part B, pp. 241–289, Academic Press, New York.

Timkovich, R. & Dickerson, R.E. (1976). *J. Biol. Chem.* **251**, 4033–4046.

Trewhella, J., Carlson, V.A.P., Curtis, E.H. & Heidorn, D.B. (1988). *Biochemistry* **27**, 1121–1125.

Tsunasawa, S., Stewart, J.W. & Sherman, F. (1985). *J. Biol. Chem.* **260**, 5382–5391.

Waldemeyer, B. & Bosshard, H.R. (1985). *J. Biol. Chem.* **260**, 5184–5190.

Waldemeyer, B., Bechtold, R., Bosshard, H.R. & Poulos, T.L. (1982). *J. Biol. Chem.* **257**, 6073–6076.

Wallace, C.J.A. (1987). *J. Biol. Chem.* **262**, 16767–16770.

Wallace, C.J.A. & Corthesy, B.E. (1986). *Protein Eng.* **1**, 23–27.

Wallace, C.J.A. & Harris, D.E. (1984). *Biochem. J.* **217**, 589–594.

Wallace, C.J.A. & Proudfoot, A.E.I. (1987). *Biochem. J.* **245**, 773–779.

Wallace, C.J.A. & Rose, K. (1983). *Biochem. J.* **215**, 651–658.

Wallace, C.J.A., Corradin, G., Marchiori, F. & Borin, G. (1986). *Biopolymers* **25**, 2121–2132.

Wallace, C.J.A., Mascagni, P., Chait, B.T., Collawn, J.F., Paterson, Y., Proudfoot, A.E.I. & Kent, S.B.H. (1989). *J. Biol. Chem.* **264**, 15199–15209.

Ward, K.B., Wishner, B.C., Lattman, E.E. & Love, W.E. (1975). *J. Mol. Biol.* **98**, 161–177.

Wendoloski, J.J., Matthew, J.B., Weber, P.C. & Salemme, F.R. (1987). *Science* **238**, 794–797.

White, T.B., Berget, P.B. & Nall, B.T. (1987). *Biochemistry* **26**, 4358–4366.

Williams, R.J.P. (1988). *Z. phys. Chemie Leipzig* **269**, 387–402.

Williams, G., Moore, G.R, Porteous, R., Robinson, M.N., Soffe, N. & Williams, R.J.P. (1985a). *J. Mol. Biol.* **183**, 409–428.

Williams, G., Moore, G.R. & Williams, R.J.P. (1985b). *Comments Inorg. Chem.* **4**, 55–98.

Wilson, A.J.C. (1942). *Nature (London)* **150**, 151–152.

Wlodawer, A. & Hendrickson, W.A. (1982). *Acta Crystallogr.* **A38**, 239–243.

Wood, L.C., Muthukrishnan, K., White, T.B., Ramdas, L. & Nall, B.T. (1988a). *Biochemistry* **27**, 8554–8561.

Wood, L.C., White, T.B., Ramdas, L. & Nall, B.T. (1988b). *Biochemistry* **27**, 8562–8568.

Wooten, J.B., Cohen, J.S., Vig, I. & Schejter, A. (1981). *Biochemistry* **20**, 5394–5402.

Wright, C.S. (1987). *J. Mol. Biol.* **194**, 501–529.

Wu, T.T. & Kabat, E.A. (1970). *J. Exp. Med.* **132**, 211–250.

Zuniga, E.H. & Nall, B.T. (1983). *Biochemistry* **22**, 1430–1437.