Structural Studies of

Cytochrome *c*

or

Cytochrome *c* Allsorts

by *Jonathan C. Parrish*

Submitted in partial fulfilment of the requirements for the degree of

PhD

at

Dalhousie University

Halifax, Nova Scotia, Canada

July, 1997

DALHOUSIE UNIVERSITY

FACULTY OF GRADUATE STUDIES

The undersigned hereby certify that they have read and recommend to the Faculty of Graduate Studies for acceptance a thesis entitled "Structural Studies of Cytochrome *c* "

by Jonathan C. Parrish

in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

Dated: August 15, 1997

External Examiner

Research Supervisor

Examining Committee

DALHOUSIE UNIVERSITY

DATE: July, 1997

AUTHOR: Jonathan C. Parrish

TITLE: Structural Studies of Cytochrome *c*

DEPARTMENT: Department of Biochemistry

DEGREE: PhD CONVOCATION: October YEAR: 1997

SIGNATURE OF AUTHOR

And then there is painting. What a red rag is to a bull, Turner's "Slave Ship" was to me before I studied art. Mr. Ruskin is educated in art up to a point where that picture throws him into as mad an ecstasy of pleasure as it used to throw me into one of rage, last year, when I was ignorant. His cultivation enables him - and me, now - to see water in that glaring yellow mud, and natural effects in those lurid explosions of mixed smoke and flame, and crimson sunset glories. It reconciles him - and me, now - to the floating of iron cable-chains and other unfloatable things. It reconciles us to fishes swimming around on top of the mud - I mean the water. The most of the picture is a manifest impossibility - that is to say, a lie. And only rigid cultivation can enable a man to find a truth in a lie. But it enabled Mr. Ruskin to do it, and it has enabled me to do it. A Boston Newspaper reporter went and took a look at the Slave Ship floundering about in that fierce conflagration of red and yellows and said it reminded him of a tortoise-shell cat having a fit in a platter of tomatoes. In my then uneducated state that went home to my nonCultivation and I thought here is a man with an unobstructed eye. Mr Ruskin would have said: This person is an ass. That is what I would say, now.

Mark Twain "A Tramp Abroad"

Table of Contents

TABLE OF CONTENTS (CONT) PAGE

List of Figures and Tables

Figures

FIGURES (CONT.)

TABLES

TABLES (CONT.)

Abstract

Cytochrome *c* is a small heme-containing protein, found in the intermembrane space of mitochondria, which forms a key link in the process of oxidative phosphorylation by transferring electrons between Complexes III and IV, in addition to a number of other mitochondrial electron-transport proteins. Studies of this critical protein are aimed at elucidating the nature of electron transport, intermolecular interactions, and protein folding and structure. In the first part of this work, a number of mutants of *S. cerevisiae* iso-1 cytochrome *c* have been examined involving mutations of two surface charged residues, Arg13 and Asp90, and two partially buried residues, Phe82 and Leu85. Modifications of Arg13 leads to major deviations in electron transport while those of Asp90 lead to more structural effects. For the two hydrophobic residues, structural disruptions appear to be related to the size of residue 85 while functional disruptions are related to the hydrophobicity, suggesting that conservation of this residue maintains aspects of both. In the second part, a binding model of ATP to horse cytochrome *c* has been developed using a computer simulation with energy minimisation, and suggests that the nucleotide binds to the protein surface proximal to the invariant Arg91, a number of lysine residues, and Glu69. In the third part of this study, computer modelling of a number of "methionine scan" mutants of *S. cerevisiae* iso-1 cytochrome *c* was carried out to complement a number of structural and functional studies. The results show that methionine insertions can be tolerated in many different locations but may have unexpected effects based on the evolutionary variability of an amino acid, as replacement of the variable Lys55 apparently leads to structural and functional disruptions.

List of Abbreviations and Symbols Used

8-N_3-ATP	8-azidoadenosine 5' triphosphate
Ala	Alanine
Asn	Asparagine
Asp	Aspartic Acid (Aspartate)
Arg	Arginine
ATP	Adenosine 5' triphosphate
CCP	*S. cerevisiae* Cytochrome *c* Peroxidase
CDNP	4-carboxy-2,6-dinitrophenyl
Cys	Cysteine
DTT	Dithiothreitol
EDTA	Ethylenediaminetetraacetate
Gln	Glutamine
Glu	Glutamic Acid (Glutamate)
Gly	Glycine
HEPES	N-2-Hydroxyethylpiperazine-N'-2-ethanesulphonic acid
His	Histidine
HPLC	High-Performance Liquid Chromatography
Ile	Isoleucine
LB	Luria-Bertani Medium
Leu	Leucine
Lys	Lysine

Met	Methionine
mg	Milligrams
mol	Moles
MOPS	3-(N-Morpholino)propanesulphonic acid
PDB	(Brookhaven) Protein Data Bank
Phe	Phenylalanine
PMSF	Phenylmethylsulphonyl flouride
Pro	Proline
Ser	Serine
ssDNA	single-stranded DNA
TE	Tris-EDTA
Thr	Threonine
TNP	2,4,6-Trinitrophenyl
TMPD	Tetramethyl-p-phenylene diamine
TNP-8-N_3-ATP	2'(3')-*O*-(2,4,6-trinitrophenyl)-8-azidoadenosine 5' triphosphate
TRIS	Tris(hydroxymethyl)aminomethane
Trp	Tryptophan
Tyr	Tyrosine
Val	Valine
YPD	Yeast extract/peptone/dextrose media

ACKNOWLEDGEMENTS

FIRST AND FOREMOST I WOULD LIKE TO THANK CARMICHAEL FOR SUPPORTING ME AND PROVIDING ME WITH A PLACE TO WORK AND PLAY.

THANKS ALSO TO DAVE BYERS, MIKE GRAY, RICK SINGER AND JACOB VERPOORTE FOR SERVING ON MY COMMITTEE AND TO PAUL LIU FOR GETTING ME STARTED.

Special thanks to Angela Brigley for assistance in the lab, to Guy Guillemette for help with mutagenesis, to Jack Kornblatt for supplying some of the mutants, and to Terry Lo and Natalie Strynadka for aiding in computer modelling and software acquisition.

Super special thanks to:

Anthony Woods, Anne Rich, Ken Niguma, Doug Craig, Ian Dawe and Scott Beeser

Super-Duper Special Thanks to:

Alexo, Andrew, Barb, Bart, Björn, Bob, Dave, Donna, E--Rat, Gwen, Homer, Iphigenia, Julie, Kathryn, Lisa, Little Eagle, Maggie, Marel, Marg, Marg, Marge, Moira, Patrick, Professor, Ptit Gris, Roisin, Space Moose, Spencer, Theodore, Trina, Yauri, Zoe.

Boomcastic Super-Duper Special Thanks to:

The Messrs Plowe and Scomergy Scomickle. Without whom none of this would have been possible. Thank you Kings of Cornish Carnage.

CHAPTER 1 : INTRODUCTION

1.1 THE CENTRAL DOGMA AND PROTEIN STRUCTURE

When Crick proposed the central dogma in 1958, he was describing what at that time was a hypothesis concerning the flow of "information" in biological systems. The simple summary of his hypothesis was that information flowed from nucleic acid to nucleic acid or from nucleic acid to protein, but not from protein to nucleic acid or protein to protein. Specifically, the information flow was from the deoxyribonucleic acid (DNA) sequence to the ribonucleic acid (RNA) sequence to the amino acid sequence (Fig. 1.1A). While this described the most common cases encountered in gene expression and protein translation, the central dogma can be expanded to that of Fig. 1.1B, which includes the ability to reproduce RNA from RNA (RNA replicase) or DNA from RNA (reverse transcriptase). In addition, other processes are involved in protein production; editing or other genetic rearrangements which fundamentally affect the *information* used to generate the amino acid sequence and post-translational modifications which occur and which may be necessary for a functional protein. The result, if we include all of these steps in the central dogma, is summarised in Fig. 1.1C. In each case, the arrows indicate when the information encoded in each molecule is used to direct the modification to the next step. However, in all of the models the fundamental point is the same: *information* encoded in

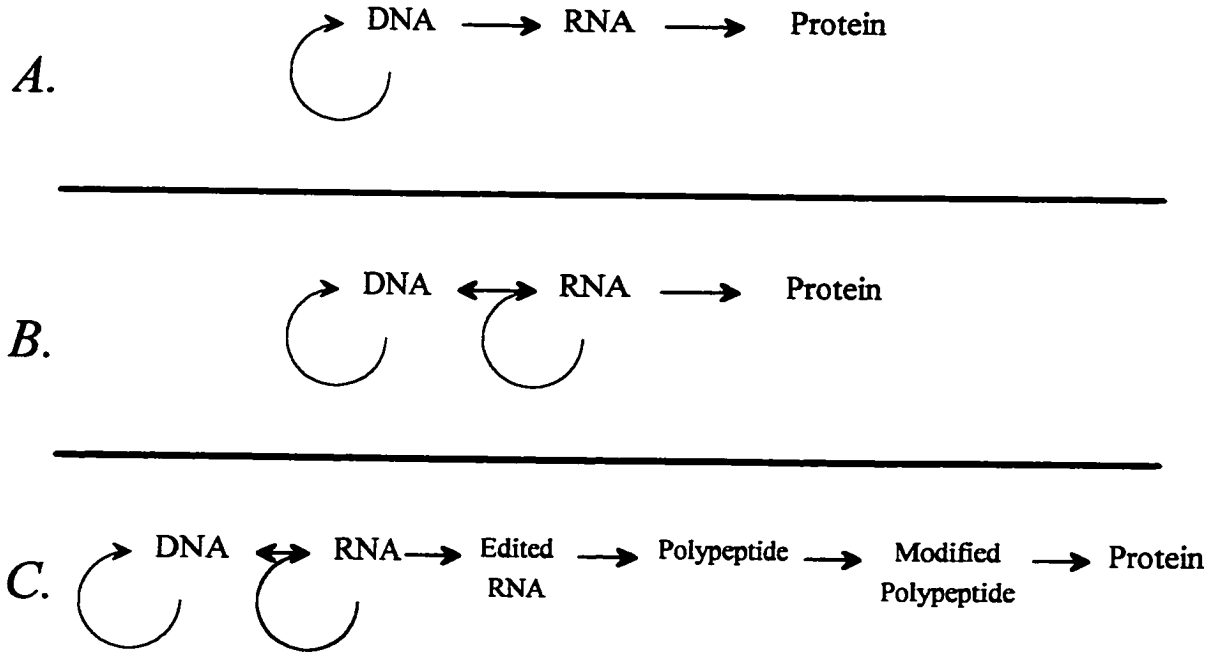


Figure 1.1: The "Central Dogma" and Variations

The Central Dogma, as originally proposed by F. Crick in 1958 is shown in panel A. Panel B shows the expanded dogma to include additional nucleic acid polymerases. Panel C shows a further expanded model to include RNA editing and polypeptide dependent steps.

the linear sequence is translated into *function* in the three dimensional arrangement. Thus a pivotal question in biochemistry is concerned with how the sequence relates to the functional and structural properties of the final protein. Any attempt to address this question requires knowledge of what the modifications are at each step and what the features of the completed whole are. This is no longer a linear problem in the same sense as that of transcription or translation.

Experience shows that different amino acid sequences can produce essentially the same structure and/or function in different proteins (Finkelstein et al., 1993). As a result, we cannot realistically expect to derive a "dictionary" of protein structure which tells us that a particular sequence or amino acid produces a specific structure or function because the behaviour of each amino acid in the sequence is context dependent. What we actually need to determine could perhaps be more accurately described as a "thesaurus" of protein structure. One of the most informative ways of addressing the relationship between amino acids and structure or function is within a single protein context; this can be approached by examining the natural variants of the protein or by specifically altering the protein through engineering.

1.2 Cytochrome c as a Model Protein

Cytochrome *c* is a popular model for protein structure and function studies for a number of reasons. In addition to its æsthetically pleasing red colour, other useful characteristics include its ease of production and purification, solubility, structural stability, a number of distinct measurable spectroscopic markers, and its ubiquitousness amongst all mitochondria. Additionally, cytochrome *c* interacts with a number of different physiological partners allowing for functional studies under a number of different circumstances. All of these features make it an excellent model protein for the study of electron transport.

The proteins that interact with mitochondrial cytochrome *c* are shown schematically in Fig. 1.2. The most important enzymes for oxidative phosphorylation that interact directly with cytochrome *c* are Complexes III and IV (coenzyme Q-cytochrome *c* oxidoreductase and cytochrome oxidase respectively) shown in Fig. 1.2a. Complex III receives electrons from Complex I (NADH dehydrogenase) or Complex II (succinate Q reductase) via coenzyme Q, Complex IV completes the electron transport pathway with concomitant reduction of molecular oxygen to water. Phosphorylation of ADP to ATP is coupled to proton translocation across the inner membrane at Complexes I, III, and IV. In animals, the additional enzymes cytochrome b_5 and sulphite oxidase donate electrons to cytochrome *c* (Fig. 1.2b). Sulphite oxidase transfers electrons from the reaction of SO_3^{2-} to SO_4^{2-} and cytochrome b_5 obtains electrons from cytosolic NADH oxidation via

Figure 1.2: Mitochondrial electron donors and acceptors for cytochrome *c (over)*

Coenzyme Q (Q) passes electrons from Complexes I and II to Complex III (coenzyme Q-cytochrome *c* oxidoreductase) which then flow to cytochrome *c*. These complexes, with Complex IV (cytochrome *c* oxidase), form the main respiratory chain for oxidative phosphorylation; Complexes III and IV are shown in (A). Sulphite oxidase (SO) and cytochrome b_5 are found specifically in animals (B) while cytochrome b_2 and cytochrome *c* peroxidase (CCP) are found specifically in yeast (C). IM, IMS, and OM refer to inner mitochondrial membrane, intermembrane space and outer membrane respectively. Arrows indicate electron flow. See text for details.

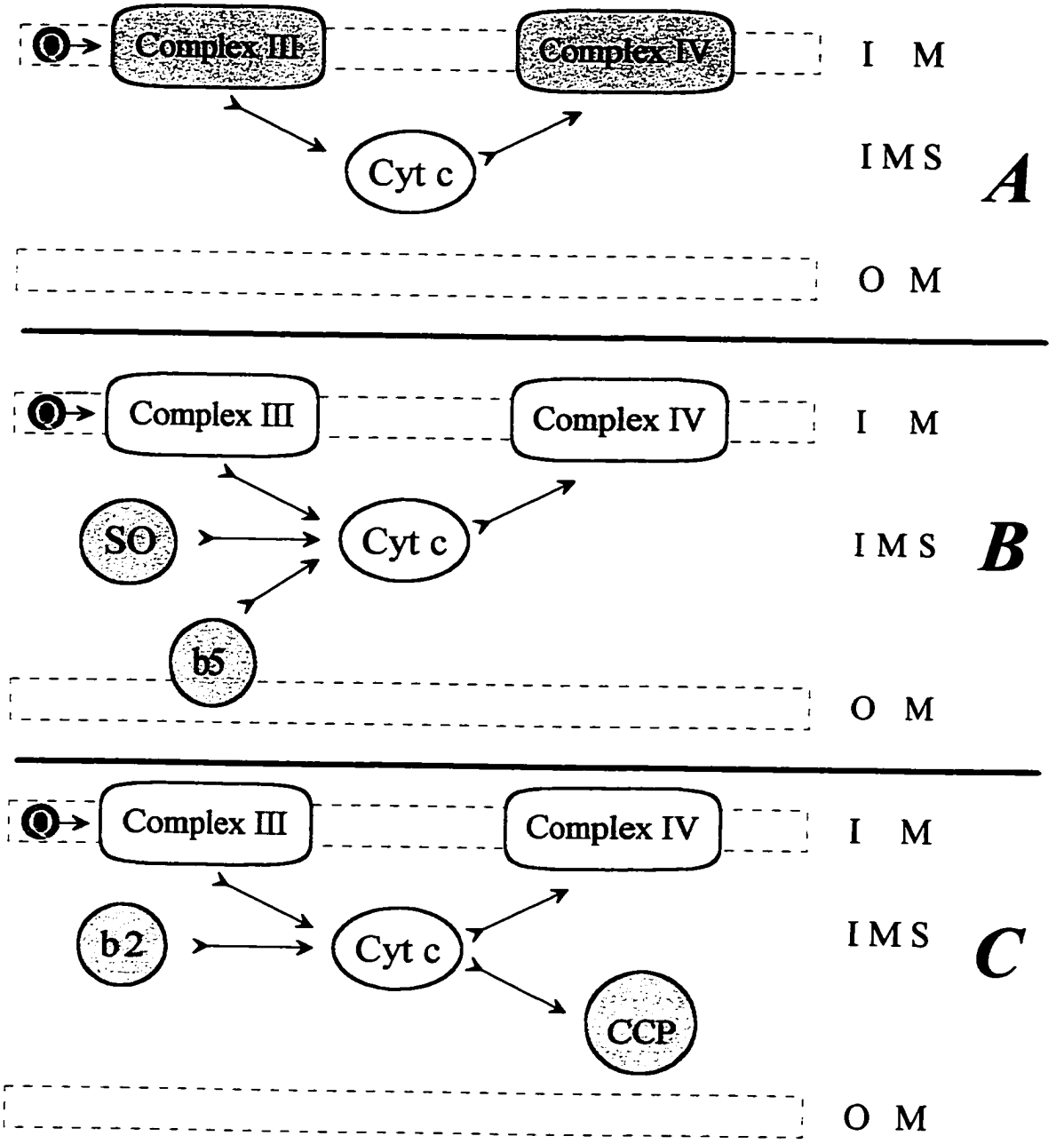


Figure 1.2: Mitochondrial electron donors and acceptors for cytochrome *c*

NADH-cytochrome b_5 reductase (located in the outer membrane). Fig. 1.2c shows cytochrome b_2 (lactate dehydrogenase) and cytochrome *c* peroxidase, two enzymes specific to yeast mitochondria. Cytochrome b_2 passes electrons obtained from oxidation of lactate to pyruvate in the inter membrane space while cytochrome *c* peroxidase accepts electrons from cytochrome *c* and reduces hydrogen peroxide to water.

1.3 Cytochromes c: Amino acid sequences and structure

Mitochondrial cytochrome *c* has a number of different physiological partners and fulfils most of the same basic functions within many organisms; as a result, it is reasonable to expect that different natural variants will share many conserved features. At least 88 distinct mitochondrial cytochrome *c* sequences have been determined (Moore and Pettigrew, 1990) and the examples which have had three-dimensional structural determinations published are shown aligned in Table 1.1. Across this diverse group of sequences are numerous conserved amino acids which are presumably necessary for maintaining either the structure or the function of the native protein. Fig. 1.3 shows a superposition of the alpha-carbon traces for cytochrome *c* structures solved to a resolution of 2 Å or less; it is clear that, despite the sequence differences (44% identity for all five), all the proteins adopt essentially the same structure.

One of the most important functions of the protein chain of cytochrome *c* is to maintain the environment of the heme group, the iron-containing central core. The heme

```
                      #    #   #   # ####   #   # #### # ###  #   #  #
ALBACORE  :          @GDVAKGKKTFVQKCAQCHTVENGGKHKVGPNLWGLFGRKTGQAEGYSYT
BONITO    :          @GDVAKGKKTFVQKCAQCHTVENGGKHKVGPNLWGLFGRKTGQAEGYSYT
HORSE     :          @GDVEKGKKIFVQKCAQCHTVEKGGKHKTGPNLHGLFGRKTGQAPGFTYT
S.c.iso1  :     TEFKAGSAKKGATLFKTRCLQCHTVEKGGPHKVGPNLHGIFGRHSGQAEGYSYT
S.c.iso2  : AKESTGFKPGSAKKGATLFKTRCQQCHTIEEGGPNKVGPNLHGIFGRHSGQVKGYSYT
RICE      : @ASFSEAPPGNPKAGEKIFKTKCAQCHTVDKGAGHKQGPNLNGLFGRQSGTTPGYSYS
                     |    |   |     |||          || | |   |  |      |
Consensus :       eappGDVEKGKKIFVQKCAQCHTVEKGGKHKTGPNLHGLFGRKTGQAAGFSYT
                -5    1   5    1    1    2    2    3    3    4    4
                                   0    5    0    5    0    5    0    5

                 ##      #        ## ## ######## # #  #    # ##  ##
ALBACORE  :  DANKSKGIVWNNDTLMEYLENPKKYIPGTKMIFAGIKKKGERQDLVAYLKSATS
BONITO    :  DANKSKGIVWNENTLMEYLENPKKYIPGTKMIFAGIKKKGERQDLVAYLKSATS
HORSE     :  DANKNKGITWKEETLMEYLENPKKYIPGTKMIFAGIKKKTEREDLIAYLKKATNE
S.c.iso1  :  DANIKKNVLWDENNMSEYLTNP*KYIPGTKMAFGGLKKEKDRNDLITYLKKATE
S.c.iso2  :  DANINKNVKWDEDSMSEYLTNP*KYIPGTKMAFAGLKKEKDRNDLITYMTKAAK
RICE      :  TANKNMAVIWEENTLYDYLLNP*KYIPGTKMVFPGL*KPQERADLISYLKEATS
               |      |        |  | |  || || | |      |
Consensus :  DANKNKGITWGEDTLMEYLENPKKYIPGTKMIFAGLKKPQERADLIAYLKKATA
             5    5    6    6    7    7    8    8    9    9    1
             0    5    0    5    0    5    0    5    0    5    0
                                                               0
```

Table 1.1: Alignment of mitochondrial cytochrome *c* sequences

Sequence alignment includes all the natural variants of mitochondrial cytochrome *c* with published crystal structures. Numbering is based on standard horse numbering, @ and * indicate an acetyl group (blocked N-termini) and the modified lysine residue trimethyllysine respectively. *S.c.*iso1 and *S.c.*iso2 refer to the *S. cerevisiae* mitochondrial cytochrome *c* isoforms 1 and 2 respectively. Identical residues in all five proteins are indicated with a # above the residue position. The consensus sequence shows the most common amino acids at each position based on 96 natural variants and is adapted from Moore and Pettigrew (1990); lower case letters indicate N-terminal extensions where an amino acid is usually present, but no one amino acid is prevalent. Vertical lines mark invariant residues in all the known sequences.

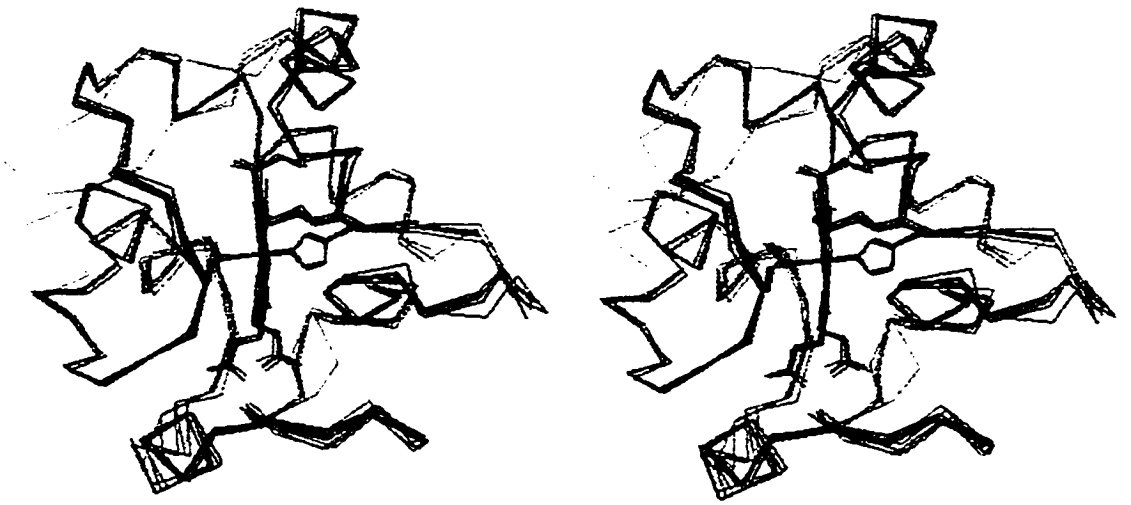


Figure 1.3: Superposition of five cytochrome *c* structures

A stereo view of the backbone traces for five different cytochrome *c* structures, all solved to 2.0 Å resolution or better. Structures shown are for mitochondrial cytochrome *c* from: Albacore tuna (1.5 Å, Takano and Dickerson, 1981a); *S. cerevisiae* iso-1 (1.23Å, Berghuis and Brayer, 1992) and iso-2 (1.5 Å, Murphy et al., 1992); Horse (1.94 Å, Bushnell et al., 1990); and Rice (2.0 Å, Ochi et al., 1983).

Darker shading indicates portions of the protein which are closer to the viewer. The heme group is also shown (visible in the centre of the protein) and side chains involved in co-ordinating the iron atom and covalent attachment of the heme.

is attached through thioether linkages to cysteines 14 and 17 (with the exception of a few kinetoplastids where only cysteine 17 is present); the iron atom is co-ordinated by an imino nitrogen of histidine 18, the sulphur of methionine 80, and four nitrogens from the heme ring system. These residues are shown explicitly in Fig. 1.3. The protein fold maintains a hydrophobic milieu around the heme group, in this way the redox potential of the group is maintained.

Another conserved feature of the protein is a number of positively charged lysine and arginine residues (and the positively charged non-standard amino acid trimethylysine). These residues form a ring of positive charge on one side of the molecule; it is this arrangement of charges which is responsible for the electrostatic interactions of cytochrome *c* with its physiological partners. It is interesting to note that one positive residue, the invariant residue arginine 91, has only recently been shown to have an explicit function in ATP binding (Craig and Wallace, 1991), an aspect which will be discussed in more detail later. The other invariant arginine residue at position 38 is involved in an electrostatic interaction with one of the heme propionate groups.

Both invariance and conservative substitutions of amino acids across a diverse group of organisms is a persuading indication of necessity; this provides us with a good first insight into understanding the contribution of each member to the whole. With this as a framework, it is possible to examine potential variations on a single theme through

engineered changes to a single protein structure. Both of these aspects concerning those specific residues addressed in this study will be discussed in more detail.

1.3.1 RESIDUES 13 AND 90

Two conserved residues which are of interest are those found at positions 13 and 90 in the alignment (Table 1.1); residue 13 is either a lysine or an arginine, 90 is either aspartic acid, glutamic acid, asparagine, or glutamine. The association matrix in Table 1.2 shows that the two residues are not covariant, and that the Lys/Glu pair is the most common. The two residues are found fairly closely associated in the structure (Fig. 1.4), which raises the possibility that the two may form an ionic interaction (with the exception of those with Asn or Gln at 90). Other possibilities for the conservation of these surface exposed amino acids include maintenance of the ionic character of the cytochrome, engaging in specific interactions with physiological partners, or involvement in other structural and/or mechanistic aspects.

The idea that residues 13 and 90 may be involved in an electrostatic interaction with each other was supported by early structural work. The two appeared to form an ionic interaction in the crystal structure solutions for albacore (Tanaka et al., 1975), bonito (Takano and Dickerson, 1981a and 1981b), rice (Ochi et al., 1983) and the original structure for horse (Dickerson, 1972) cytochromes *c*; in these structures the charged groups are on the order of 3.0 Å apart (Table 1.3). This combined with an examination of

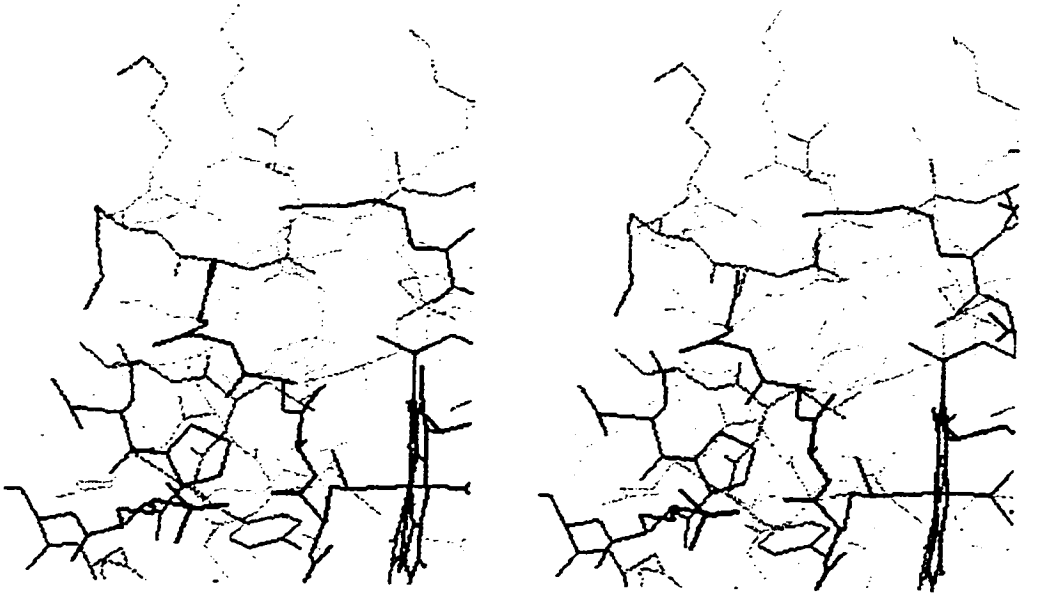


Figure 1.4: Stereoview of *S. cerevisiae* iso-1 cytochrome *c* showing side chains

Shown are those amino acids in one portion of the protein. The five amino acids shown in colour are Lys13 (dark blue), Phe82 (brown), Leu85 (green), Asp90 (red), and Arg91 (light blue). The heme group is visible in the lower right corner. Orientation of the molecule is identical to that of Figure 3.

electrostatic interactions suggested that the two residues 13 and 90 formed one of two salt bridges that stabilised the heme crevice (Osheroff et al., 1980). Later structures, however, for the *S. cerevisiae* iso-1 and iso-2 cytochromes *c* and a new structural determination for horse did not continue to show this interaction, even though the distances between α-carbons remained essentially invariant (Table 1.3) (Bushnell et al., 1990; Berghuis and Brayer, 1992; Murphy et al., 1992). While it is possible to suggest that crystallisation conditions may influence this interaction, the electrostatic association is certainly not absolute.

If the two residues do not interact with each other, then the conservation of ionic character needs to be explained in some other fashion. As mentioned in Section 1.3, one side of cytochrome *c* is characterised by a ring of basic residues, including residue 13. The arrangement of a number of positive residues on one side of the molecule and negative residues on the other generates a dipole which may assist in the association of cytochrome *c* with physiological partners (Osheroff et al. 1980). The potential ramifications of the dipole include providing a specific interaction surface for physiological partners and the more general effect of orienting cytochrome *c* so that intermolecular collisions are consistently productive. In the latter case, where a charged residue may not form a specific contact between two partners, replacement with a similar charged residue would not be expected to significantly alter the interaction.

If a conserved residue is involved in a specific intermolecular contact, then changing it to another residue, even one chemically similar, may have a significant effect on activity. While there is not a lot of data concerning residue 90, chemical modification of residue 13 results in decreases of cytochrome *c* activity with cytochrome oxidase, cytochrome reductase, cytochrome b_5, sulphite oxidase, and cytochrome *c* peroxidase (reviewed in Pettigrew and Moore, 1987; Kornblatt et al., 1992). Residue 13 is also found in the intermolecular interface between cytochrome *c* and yeast cytochrome *c* peroxidase (CCP): in the *S. cerevisiae* cytochrome *c* / CCP complex Arg13 forms an ionic bond with Asp34 of CCP, in the horse cytochrome *c* / CCP complex Lys13 interacts with Glu90 ($N\zeta_{Lys13}$ - $O\varepsilon_{Glu90}$ = 2.67 Å) (Pelletier and Kraut, 1992). Clearly residue 13 is intimately involved in the intermolecular associations of cytochrome *c* with its physiological partners.

The final possibility to explore is that these residues are conserved in order to preserve other structural or mechanistic aspects. Chemical modification of Lys 13 in horse cytochrome *c* with CDNP or TNP destabilises the heme crevice (Osheroff et al., 1980). Mutation of Arg13 in *S. cerevisiae* iso-1 cytochrome *c* to isoleucine showed an opening of the heme crevice and a concomitant increase in electron transfer rates with CCP (Hazzard et al., 1988). Electrochemistry experiments with this same mutant (Arg13Ile) suggest that the electrostatic character of the surface is critical to maintenance of protein conformation in the presence of other polar interactions (Burrows et al., 1991). Work done with both Arg13 and Asp90 in *S. cerevisiae* iso-1 cytochrome *c* shows that mutation

		Residue 13		
		Lys	Arg	Total
Residue 90	Glu	42	12	54
	Asp	13	11	24
	Gln	1	1	2
	Glx	3	0	3
	Asx	2	0	2
	Total	61	24	85

Table 1.2: Amino Acid Association Matrix for Residues 13 and 90 in Mitochondrial Cytochromes *c*

Numbers were collected by probing the SWISSPROT protein database using the PhD alignment program (Rost and Sander, 1994) with the sequence for *S.c.* iso-1 cytochrome *c*. Glx and Asx refer to those residues which could be either Glu or Gln and Asp or Asn respectively.

	Structure	Oxid.	Res.	13	90	$C\alpha_{13}$-$C\alpha_{90}$	N_{13} - O_{90}
A	Albacore I §	O	1.8 Å	Lys	Glu	11.12 Å	2.64 Å
B	Albacore II	O	1.8 Å	Lys	Glu	11.21 Å	2.86 Å
C	Albacore	R	1.5 Å	Lys	Glu	11.29 Å	2.78 Å
D	Bonito	R	2.3 Å	Lys	Glu	11.08 Å	3.82 Å
E	Horse	O	1.94 Å	Lys	Glu	11.15 Å	5.64 Å
F	Rice	O	1.5 Å	Lys	Glu	11.09 Å	3.36 Å
G	*S.c.* (iso1)	O	1.9 Å	Arg	Asp	11.02 Å	7.22 Å
H	*S.c.* (iso1)	R	1.23 Å	Arg	Asp	11.06 Å	6.50 Å
I	*S.c.* (iso2)	R	1.5 Å	Arg	Asp	11.30 Å	6.50 Å

Table 1.3: Interatomic Distances for Residues 13 and 90

Distances between atoms were measured from data files deposited with the PDB. Oxid. refers to the oxidation state of the structure, Res. to the structural resolution, 13 and 90 to the amino acids found at each of these positions, $C\alpha_{13}$-$C\alpha_{90}$ to the distance between α-carbons in residues 13 and 90, and N_{13} - O_{90} to the distance between heavy atoms of charged groups on the side chains of 13 and 90. All distances were measured using MSI Technologies molecular modelling software Insight II.

References: A & B, Takano and Dickerson, 1981b; C, Takano and Dickerson, 1981a; D, Tanaka et al., 1975; E, Bushnell et al., 1990; F, Ochi et al., 1983; G, Berghuis and Brayer, 1992; H, Louie and Brayer, 1990; I, Murphy et al., 1992.

§ For the Albacore, the structures I and II reflect the two structures which form the repeating unit in the crystal.

Cytochrome *c* is a roughly spherical protein with a diameter of approximately 30 Å.

of Arg13 to cysteine results in little change in stability but lower activity, changing Asp90 to cysteine leads to lower stability but little change in activity, and a double mutant with cysteine at both positions exhibited both lower stability and function suggesting that residue 13 has a primarily functional role while residue 90 has a structural one (Huang et al., 1994). Finally, the structural solutions for *S. cerevisiae* iso-1 cytochrome *c* shows that upon oxidation Arg13 forms a hydrogen bond to the carbonyl of Gly84 which is not present in the reduced form suggesting a possible mechanistic role (Berghuis and Brayer, 1992). Further exploration of these residues will aid in our understanding of their contribution to the structure and function of the whole protein.

1.3.2 Residues 82 and 85

Spatially located near residue 13 in a partially buried position adjacent to the heme group are the residues 82 and 85 (Fig. 1.4). Residue 82 is an invariant phenylalanine while residue 85 is conserved as a bulky hydrophobic residue, primarily isoleucine and leucine with single examples of methionine and phenylalanine (Moore and Pettigrew, 1990).

Phe82 has been extensively studied through mutation analysis, and was one of the first residues studied in cytochrome *c* (Pielak et al., 1985). It is interesting that, although it is invariant, it is possible to make a wide range of substitutions at this position without significantly altering gross function of the protein, the only exceptions being the Cys82

mutant which shows significantly reduced activity *in vivo* and the His82 mutant which leads to replacement of the Met80 heme ligand with the new histidine (Inglis et al., 1991; Hawkins et al., 1994). The role of Phe82 in cytochrome *c* structure and function appears to be at least threefold: maintenance of the redox potential of the heme group through both shielding against solvent molecules and its inherent hydrophobicity, involvement in protein-protein interactions, and the providing of an efficient channel for electron transfer owing to its aromaticity. (Pielak et al, 1985; Liang et al., 1988; Rafferty et al., 1990). Additionally, Phe82 replacements with serine, glycine, isoleucine and leucine lowered protein stability suggesting a further structural role (Pearce et al. 1989).

Residue 85, although not as extensively studied, forms a close interaction with Phe82. A number of structural studies have been done with these two residues, and crystal structures are available for all the *S. cerevisiae* iso-1-cytochrome *c* mutants examined in this study; these include Leu85Met, Leu85Phe, Leu85Cys, Leu85Ala, Phe82Tyr, and the double mutant Phe82Tyr/Leu85Ala (Lo et al., 1995a and b). Local structural perturbations are noted for all the structures, and the results are summarised in Table 1.4.

It is interesting to note that the introduction of water molecules into the Leu85Ala/Phe82Tyr mutant does not significantly alter the redox potential as might be expected (Lo et al., 1995a), suggesting that conservation of residue 85 as both a bulky and hydrophobic group is important for other reasons besides redox potential determination.

Leu85Met §	-Positional shift of Leu9 -Sulphur positioned far from Phe82 to prevent unfavourable interaction
Leu85Phe §	-Aromatic/Aromatic interaction between Phe82 and Phe85 -Leu9 and Arg13 shift away from Phe85 to compensate for increased bulk
Leu85Cys §	-Introduction of water molecule (bulk maintained) -Sulphur atom forms favourable non-polar interaction with Phe82 -Movement of Leu9 away from new water
Phe82Tyr †	-Tyr82 rotates into more solvent exposed position -Shift in Arg13 and Leu9 to account for added bulk
Leu85Ala †	-Arg13 and Asp90 form a salt bridge -Shifts in Leu9 and Leu94 maintain hydrophobicity and fill the new space formed
Leu85Ala/Phe82Tyr †	-Two water molecules bound in new hydrophilic pocket -Arg13 and Asp90 form a salt bridge -Shift in Leu9

Table 1.4: Summary of Structural Changes in Mutants of *S. cerevisiae* iso-1 cytochrome *c*

Major structural changes that occur in six crystallised mutants are listed. Data summarised from Lo et al. 1995a (§) and Lo et al. 1995b (†).

Table 1.4 indicates a number of rearrangements that occur in the mutant forms, suggesting that it may fulfil a largely structural role. One proposed role for the conservation of this residue is to prevent the Arg13 - Asp90 interaction so that Arg13 is more available to interact with physiological partners (Lo et al., 1995a).

1.3.3 Arginine 91 and ATP Binding

The residue Arginine 91 is invariant (Table 1.1) yet its role remained elusive until approximately 10 years ago. Replacement or modification of this residue results in no apparent diminishing of biological activity (Corthésy and Wallace, 1986), thus its activity has to be involved in some other aspect of cytochrome *c* function to explain its evolutionary conservation.

Modification of Arg91 does affect the binding of anions to cytochrome *c*, of particular interest and importance is the binding of ATP (Corthésy and Wallace, 1986). Since electron transport is intimately coupled to ATP synthesis, it is not unreasonable to suggest that ATP binding might have a modulating activity on cytochrome *c* in the sense of feedback inhibition. However, such an effect is only possible if the following conditions are met: ATP must bind in the physiological concentration range and it must affect cytochrome *c* activity to diminish electron transport activity.

At physiological ionic strength, horse cytochrome *c* does bind ATP within the physiological range with a K_d on the order of 3.5 mM (Craig and Wallace, 1991), making it appropriate for a regulatory role. In attempting to assess the effect of the nucleotide on activity, it is necessary to separate the effect of ATP on cytochrome *c* from any effects on other electron transport proteins, since ATP is also proposed to modulate cytochrome *c* oxidase activity (Kadenbach, 1986). Using the ATP analogue $8N_3$-ATP to covalently modify horse cytochrome *c* led to the discovery that the modified protein had a significantly reduced electron transport rate with both the reductase and the oxidase, and a decreased association with the inner mitochondrial membrane despite no apparent change in the physical properties of cytochrome *c* (Craig and Wallace, 1993; Craig and Wallace, 1995). Since both the reductase and cytochrome *c* oxidase are integral inner mitochondrial membrane proteins dissociation would have a significant impact on the number of intermolecular collisions, as it would involve three-dimensional as opposed to two-dimensional (membrane bound) diffusion.

Presumably Arg91 is involved in the binding of ATP through interaction with the γ-phosphate. The ATP analogues TNP-$8N_3$-ATP and $8N_3$-ATP attach to the protein through activation of the azido group, the resulting product is covalently linked to a lysine residue through the 8-position of the nucleotide (Craig and Wallace, 1993; McIntosh et al., 1995). Mapping of those lysines in horse cytochrome *c* which were modified by TNP-$8N_3$-ATP revealed that lysines 72, 86 and 87 were the major identifiable attachment sites

of the analogue. All of these lysines are located near the binding site(s) proposed for other physiological electron transport partners (Pettigrew and Moore, 1987)

In order to fully explore the nature of the ATP binding site, it is desirable to have an explicit structural model to rationalise the data obtained and to aid in the design of further experiments. One of the objectives of this study was to develop this model using computer modelling of horse cytochrome *c*/ ATP binding. The results for the ATP analogues were used as a reference, however it is also important to note the effect substitution of the hydrogen attached to the C8 atom with a heavier atom may induce. The conformation of ATP is normally the *anti* conformation, which is defined by the six-membered ring of the nucleotide being essentially *trans* to the sugar ring. Substitution at the C8 position of the ring creates a steric conflict that results in the *syn* conformation being preferred, this time with the six membered ring essentially *cis* to the sugar group (rotation of ~180° about the C1'-N9 bond) (Saenger, 1984). As a result the physiological binding of ATP would be expected to be in the *anti* conformation while that of the nucleotide analogues TNP-$8N_3$-ATP or $8N_3$-ATP would be expected to be in the *syn* conformation owing to the azido groups at position 8.

1.4 Protein Engineering Studies

Examining natural variants of protein sequences can be illuminating, however it also suffers from a number of limitations. The first limitation is one of isolation; usually

natural variants have multiple amino acid differences which may confound the interpretation of each residue's contribution to the whole. The second is that natural variants will tend to fall into a range of operative differences, that is to say that ineffective or grossly altered forms are not likely to be recovered, particularly if the protein is necessary for the survival of the organism. The third limitation is that only standard amino acids will be substituted in the sequences, as these constitute the vocabulary of protein sequences in vivo. Protein engineering is an effective way of circumventing each of these restrictions, and different approaches can be loosely divided into three categories: site-directed mutagenesis, covalent modification, and synthetic methods. These techniques can be used singly, but when combined they are also highly complementary.

1.4.1 Site-Directed Mutagenesis

By far the most popular and also the most accessible protein engineering technique is site-directed mutagenesis. This process involves altering the gene responsible for encoding a protein, and then expressing the gene in a suitable organism to produce the final protein. Cytochrome *c* was one of the first proteins employed for this technique (Pielak et al., 1985) and has subsequently been exploited for a number of other studies in order to address the relationship between sequence and structure in electron transport (Pearce et al., 1989; Inglis et al., 1991; Huang et al., 1994; Lo et al., 1995a and 1995b; Rafferty et al., 1996). The advantages of this technique, because protein synthesis is carried out *in vivo*, include harnessing of cellular protein synthesis and post-translational

modification machinery (for example in cytochrome *c* this includes attachment of the heme group, modification of lysine to trimethyllysine, or acetylation of the N-terminus) and also relatively large yields. The drawbacks, which also stem from production in the organism, include again being limited to the natural set of 20 amino acids, the possibility of protein degradation by resident proteinases, and the requirement of a suitable organism for expression which will carry out any necessary post-translational modifications.

1.4.2 Covalent Modification

Another protein engineering technique which can be useful involves covalent modification of a protein through chemical methods. In work of this sort there are a number of chemical techniques which can target essentially all the natural functional groups found in proteins including amino, carboxylic acid, guanidyl, hydroxyl and sulphur groups. While a discussion of all these methods is beyond the scope of this thesis; examples of these kinds of modifications in cytochrome *c* have largely revolved around modification of the basic amino acids lysine and arginine (Corthésy and Wallace, 1986; Kornblatt et al., 1992; many reviewed in Pettigrew and Moore, 1987). The advantages of these techniques include the ability to introduce larger changes than simple amino acid substitutions and the lack of dependence on organismal production; the disadvantages can include the requirement for the protein to undergo harsh treatment and still be able to retain a folded conformation, often large quantities are necessary to obtain suitable yields

of product, and the potential difficulty in specifically isolating different modified forms from each other and from the unmodified forms.

1.4.3 Synthetic Methods

The final types of protein engineering to discuss are the synthetic methods and semisynthetic methods, these techniques involve either total synthesis of a protein or of a portion of a protein respectively. Since neither method relies on genetic translation, both allow the construction of sequences containing non-coded amino acids; both methods have been used to study cytochrome *c* (Sano, 1972; reviewed in Wallace, 1993). The major disadvantage to peptide synthesis is the yield of pure product obtained, particularly for long chain synthesis; this is the major advantage of semisynthetic methods. Semisynthetic methods, as mentioned earlier, involve the synthesis of a portion of a protein followed by reassembly of this synthetic portion with the balance of the protein which was obtained from the original source (these fragments are usually generated by cleavage of the native protein). Thus the length of the synthesised portion is reduced considerably, allowing for higher yields and purity. Also, the natural fragment will retain all of the post-translational modifications, simplifying reconstruction of the intact novel protein.

One semisynthetic technique of particular importance to this study is the use of peptide cleavage at methionine residues by cyanogen bromide. Under these conditions, this reaction produces two important results: the peptide chain is cleaved on the

carboxy-terminal side of the methionine with concomitant conversion of the methionine to homoserine lactone. This amino acid is unique in that it can spontaneously undergo peptide bond formation with a proximal amino group, reforming the intact peptide chain with a homoserine (-CH_2CH_2OH) substituted for methionine at the same position (Corradin and Harbury, 1974). The utility of this technique has been successfully combined with site-directed mutagenesis to introduce different cleavage sites illustrating the complementarity of these different methods (Wallace et al., 1991; Woods et al., 1996).

1.5 Computer Modelling of Protein Structure

One of the best aids to protein engineering and protein structure analysis, from aspects of both visualisation and molecular simulations of energetics, is the ability to render computer models of protein structure. Its uses extend from simple familiarisation with structures to rationalisation of the observed behaviour of proteins in experimental work to prediction of the results of future work. Besides visualisation of protein structure, the two major molecular simulation methods employed in this work are energy minimisation and molecular dynamics.

Cytochrome *c* provides an ideal situation in which to use modelling methods, owing to its small size and hydrophilic nature. It has been used in a number of computer modelling simulations from dynamics studies of isolated protein to complexation models with the soluble electron transport proteins cytochrome *c* peroxidase and ferrocytochrome

b_5 to predict binding interactions between physiological partners (Northrup et al., 1981; Wong et al., 1993; Lum et al., 1987; Guillemette et al., 1994). The small size of the protein ensures that computation times are minimal, since the complexity of these calculations is exponentially related to the number of atoms present in the system.

Both energy minimisation and molecular dynamics rely on two basic conditions: how each atom is defined and how that definition is treated. Each atom in a system is characterised by its element, what other atoms it is bonded to, and the type of bonds involved; this is referred to as the atom's *potential type*. How each atom is treated is defined by a set of empirical data collected into a *force field*; this includes information on bond lengths, bond angles, van der Waals information, and electrostatics which apply to the different potential types and their interactions with each other.

1.5.1 Energy Minimisation

Energy minimisation involves manipulation of a given structure through an iterative process with the single goal of potential energy reduction. Essentially, once a protein system is defined by both potential types and force field, it is possible to calculate a potential energy surface. The process of minimisation then attempts to adjust all the given atoms to positions of lower potential energy, after which the iterative process is begun again. Ideally, the process continues until the system comes to a stable point where the sum of all attractive and repulsive forces acting on every atom equals zero. It should be

noted that this condition will be satisfied by any local minima, and not necessarily a global one.

Two common algorithms are used in minimisation processes, these are "steepest descents" and "conjugate gradient". Steepest descents is a useful algorithm when the potential energies involved are changing rapidly; for instance, when replacement of an amino acid leads to large steric contributions owing to van der Waals overlap. Its advantage is that it is capable of calculating each iteration rapidly; however, its disadvantage is that it is slow to reach the equilibrium point (when potential energy gradients are small). Conjugate gradient, on the other hand, is more effective near equilibrium positions. Simply stated, it employs more calculations per iteration, but because it approaches the equilibrium more effectively it requires fewer iterations. In most minimisations, a combination of both is typically employed to gain the best advantage from both systems.

1.5.2 Molecular Dynamics

The information used in molecular dynamics simulations is no different from that used in minimisations, except that it includes kinetic energy in addition to potential energy. In this way the models are not restricted to energy minima, and are more free to explore conformational space. These simulations are carried out in an attempt to describe a real system, the assumption being that dynamics simulations represent a closer approximation

of molecular behaviour than either crystalline or minimised models; however, they suffer from the restriction that they are extremely CPU intensive compared to energy minimisation simulations.

1.6 Objectives of This Study

The work which is reported in this thesis is roughly divided into three major sections as follows: an examination of site-directed mutants of *S. cerevisiae* iso-1 cytochrome *c*, computer modelling of the ATP binding site of horse cytochrome *c*, and computer modelling of methionine mutants of *S. cerevisiae* iso-1 cytochrome *c*.

1.6.1 Protein Studies of Cytochrome c Mutants

The amino acids found at positions 13, 90, 82, and 85 are of particular interest in cytochrome *c* (see section 1.3). In the *S. cerevisiae* iso-1 cytochrome *c* protein, these residues are Arg13, Asp90, Phe82, and Leu85. The following mutants have been constructed: Arg13Lys, Arg13Cys, Asp90Cys, Arg13Cys/Asp90Cys, Phe82Tyr, Phe82Tyr/Leu85Ala, Leu85Ala, Leu85Cys, Leu85Phe, and Leu85Met. Of these, previous work has been carried out in other labs on all these mutants with the exception of Arg13Lys. The purpose of this work is to extend the understanding of the contribution of these residues to the structure and function of cytochrome *c* through analysis of physical properties and functionality within electron transport systems.

1.6.2 Computer Modelling of ATP binding to Cytochrome *c*

As discussed in section 1.3.3, horse cytochrome *c* has a single high affinity ATP binding site which has been localised to that region near residues 91, 72, 86, and 87. To further explore this and to generate an explicit structural model of the binding interaction between horse cytochrome *c* and ATP, computer modelling including energy minimisation was employed. This included simulations of both the *syn* and *anti* conformations of ATP.

1.6.3 Computer Modelling of Methionine Mutants

In order to create new and varied sites for cyanogen bromide cleavage in *S.cerevisiae* iso-1 cytochrome *c*, a number of site-directed mutants were constructed with methionine residues introduced to replace the residues Pro25, Val28, Ile35, Lys55, Leu68, and Ile75. In order for these mutants to be successfully employed in semisynthesis, it is necessary to completely characterise them beforehand. Part of this characterisation was undertaken through computer modelling of the different mutants through both minimisation and dynamics simulations to examine any particular structural perturbations which may arise from the different substitutions.

CHAPTER 2: MATERIALS AND METHODS

2.1 MOLECULAR BIOLOGY

2.1.1 PLASMID STRAINS

Two types of plasmid systems were used for the expression of the *S. cerevisiae* iso-1 cytochrome *c* gene. The first system involved transfer of the gene between two plasmids, pEMBL8 (Dente et al., 1983) and YEp213 (Broach et al., 1979), for the purposes of mutagenesis and propagation in yeast respectively, the second system involved the use of a single shuttle vector pING 4 (Inglis et al., 1991). pEMBL8 is a 4 kb bacterial plasmid containing the gene for ampicillin resistance, an origin of replication, the F1 origin for phage replication, and the β-galactosidase gene which is disrupted upon insertion of DNA into the polycloning site. The YEp213 plasmid is a 10.7 kb shuttle vector containing ampicillin and tetracycline resistance genes, origins of replication for yeast and *E. coli*, and the leu2 gene for selection in yeast auxotrophs (tetracycline resistance is destroyed upon insertion of DNA in the polycloning site). In both of these plasmids the inserted DNA was a 2kb BamHI - HindIII fragment containing the cyc1 gene (iso-1 cytochrome *c* including control sequences) with the cysteine at position 102

replaced by threonine (cyc1 C102T). The pING 4 shuttle vector which was used for most of the mutagenesis contains the ampicillin resistance gene, the F1 origin, and an origin of replication for growth in *E. coli*, the leu2 gene and 2μ sequences for growth in yeast, and the cyc1 C102T gene.

The replacement of cysteine 102 with threonine prevents the dimerization of the protein during manipulation, particularly under oxidising conditions. Replacement of this residue as a background mutation does not change the fundamental properties of the protein, and this form of the protein is the established "wild type" used in this and other studies (Cutler et al., 1987; Wallace et al., 1992).

2.1.2 Mutagenesis of Plasmids

Mutagenesis was carried out using a variation of the Kunkel method (Kunkel, 1985). Plasmids were transformed (see 2.1.3) into the *E. coli* strain RZ1032 (dut-ung-), resulting in the production of DNA with significant quantities of uracil in place of thymidine. After mutagenesis the plasmid is transformed into E. coli strain HB2151 (dut+ung+), improving selection for the mutagenized strand.

2.1.2.1 Preparation of ssDNA

Fresh transformants of pING4 in RZ1032 (an identical procedure is carried out for pING4 in HB2151) were inoculated into 2 mL of LB containing 100 μL/ml ampicillin and grown with shaking at 37°C. When the cell density of the culture reached an A_{600} of 0.1, 5 μL of the helper phage R408 was added. Cell growth was continued at 37°C for at least 5 hours at which point 1.5 mL was transferred to a microcentrifuge tube and centrifuged at high speed for five minutes to remove bacterial cells. The supernatant was transferred to another tube and the centrifugation repeated, 1 mL of this supernatant was added to 200 μL of 2.5 M NaCl / 20 % PEG 8000 solution and left on ice for 1 hour. Centrifugation of this solution for 30 minutes at 4°C yielded a small pellet of phage particles which was resuspended in 200 μL of TE and extracted once with 200 μL of phenol then twice with 200 μL of 1:1 phenol/chloroform. DNA was then precipitated by addition of 20 μL of sodium acetate (3 M, pH 5) and 500 μL of 95% ethanol . After 1 hour at -20°C samples were centrifuged at 4°C for 30 minutes and the supernatant removed, pellets were washed with 70% cold ethanol and dried. Samples of single stranded DNA (ssDNA) produced were then redissolved in 50 μL of TE and checked on a 1% agarose gel in TBE against a standard preparation of ssDNA.

2.1.2.2 Mutagenesis

Oligonucleotide primers were obtained from commercial sources complete with a 5' phosphate and were used without further purification. To anneal the primer and DNA, 5μL of ssDNA solution were mixed with 5 pmol of oligonucleotide, 1μL of buffer (500mM NaCl, 100 mM TRIS (pH 8.0), 100 mM $MgCl_2$, 10 mM DTT), and 3 μL water. Mixtures were heated to 55°C for 5 minutes and allowed to cool slowly to 30°C. At this point 4 μL of 2.5 mM dNTPs (2.5 mM dATP, dTTP, dCTP, and dGTP), 1μL of ligase buffer (660 mM TRIS (pH 7.5), 50 mM $MgCl_2$, 50 mM DTT), 1 μL 10 mM rATP, 2 μL of water, 1 unit of Klenow DNA polymerase, and 1 unit of T4 DNA ligase were added and incubated at 37°C. After two hours, 1 more unit of both Klenow and T4 ligase were added and the solution incubated for another two hours. 5 μL of this mixture was used to transform the *E. coli* strain HB2151.

2.1.3 Transformation of E. coli

All *E. coli* strains were transformed using the same technique. Cells were grown at 37°C in LB overnight, and the overnight culture used to inoculate fresh media. Cells were grown to a density of A_{600}=0.3 - 0.6 at which point they were placed on ice for 10 minutes. The sample was centrifuged at 4°C for 10 minutes at 4000g and the resulting bacterial pellet was resuspended in 1/10 th volume of cold transformation storage buffer

(TSB) (1% bactotryptone, 0.5% yeast extract, 1% NaCl, 10% PEG 3350, 5% DMSO, 10mM $MgCl_2$, 10mM $MgSO_4$). After 10 minutes on ice, 100 µL of competent cells were mixed with 5 µL or less of DNA and left on ice. After 30 minutes 900 µL of TSB with 20mM glucose was added and samples transferred to 37 °C and incubated for another 30 minutes. From these preparations, 200-500 µL of cells were plated on LB plates containing 100 µmol/mL ampicillin and grown at 37°C.

2.1.4 Sequencing of Mutants

Preparation of ssDNA for sequencing from transformed E.coli strain HB2151 was identical to that of section 2.1.2.1 for RZ1032. Sequencing of DNA was carried out using the commercially available T7 sequencing kit (Pharmacia) using a primer specific to the cyc1 C102T gene. All mutants were sequenced before and after transformation into *S. cerevisiae*.

2.1.5 Transformation of *S. cerevisiae*

Plasmids were transformed into *S. cerevisiae* using electroporation. Plasmid DNA was prepared from *E. coli* by growing transformants overnight in 2 mL LB + 100 mM ampicillin. Cultures were harvested by centrifugation in microcentrifuge tubes, the supernatant decanted and the pellets resuspended in 350 µL STET (8% sucrose, 5%

Triton X-100, 50 mM EDTA, 50 mM TRIS pH 8.0) with 1mg/mL lysozyme and 1 mg/mL RNAse. The mixture was boiled for 1 minute, placed immediately on ice, and then centrifuged for 15 minutes at 4°C. The pellet was removed and 700 µL of 95% ethanol and 35 µL of sodium acetate (3M, pH 5) was added to the supernatant. After 1 hour at -20°C samples were centrifuged at 4°C, the supernatant removed and the pellet washed with cold 70% ethanol. Pellets were resuspended in 100 µL TE, extracted once with phenol and three times with phenol:chloroform (1:1), and DNA precipitated with the addition of 60 µL of 2.5M NaCl/20% PEG 8000. Following centrifugation, pellets were redissolved in 50 µL TE and analysed on a 1% agarose gel in TBE.

Electroporation was carried out using the technique of Becker and Guarente (1992). The yeast strain GM3C-2 was used for transformation because it is a leucine auxotroph and has a deletion of cyc1 and disruption of cyc2 (Iso-2 cytochrome *c*). Cells were grown in 500 ml YPD (1% yeast extract, 2% bactopeptone, 2% glucose) at 30°C to A_{600} = 1.3-1.5, centrifuged (4000g, 10 min), and resuspended in 100 mL YPD with 20 mM HEPES (pH 8) and 25 mM DTT. Shaking continued at 30°C for 15 minutes. The volume was brought to 500 mL with sterile Milli-Q water and the cells were centrifuged again. Pellets were successively resuspended in 500 mL Milli-Q water, 250 mL Milli-Q water, 20 mL 1M sorbitol, and finally in 0.5 mL 1M sorbitol. 40 µL of this suspension was added to < 5µL DNA solution in a 50 µL electroporation cuvette. Cells were pulsed in a Biorad gene-pulser (1.5 kV, 25 µF, 200 Ω), followed by the addition of 1 mL 1M

sorbitol. The entire sample was plated on complete minimal (Leu-, Asp-, Thr-) dropout plates containing 1M sorbitol and grown at 30°C.

2.1.6 Isolation of Plasmids from *S. cerevisiae*

Plasmids were isolated from transformed yeast using glass beads to lyse the cells. Cells were grown in 2 mL cultures and harvested by centrifugation. Cell pellets were resuspended in a 200 µL solution of 100 mM NaCl/ 10 mM TRIS (pH 8)/ 1 mM EDTA/ 0.1% SDS, mixed with glass beads and vortexed for one minute. Another 200 µL of the solution was added and the mixture was phenol extracted twice. DNA was precipitated by addition of 40 µL of sodium acetate (3M, pH 5) and 1 mL of ethanol. After 30 minutes at -20°C the sample was centrifuged at 4°C for one hour and the pellet washed in 70% ethanol before resuspension in TE. This sample was used directly to transform *E. coli* for sequencing.

2.2 Protein Work

2.2.1 Preparation of Protein from S. cerevisiae

S. cerevisiae transformants harbouring even partially functional cytochrome *c* variants are capable of propagation under conditions requiring aerobic growth. Cultures were grown at 30°C in media containing 1% yeast extract, 2% bactopeptone, 0.67 % yeast nitrogen base without amino acids, and 3% carbon source (either glycerol or ethanol). Cell density was monitored at 660 nm until stationary phase was just reached, at this point 1% lactate (pH 5.5) was added to stimulate cytochrome *c* production. After an additional 3 days cells were spun down at 4000 g for 15 minutes. Cells were lysed by the addition of ½ volume of 1M KCl and ¼ volume ethyl acetate in the presence of proteolysis inhibitors (35 mg/mL PMSF, 0.3 mg/mL EDTA, and 0.1% β-mercaptoethanol). After 20 hours the solution was centrifuged for 15 minutes at 4000g, the supernatant brought up to 50% saturated ammonium sulphate, and the solution centrifuged again. The resulting solution was dialysed once against buffer (12.5 mM potassium phosphate (pH 7.0), 35 μg/ml PMSF, 0.3 mg/mL EDTA, and 0.133 mL/mL β-mercaptoethanol) and then twice against distilled water.

In order to further purify the cytochrome *c* produced, ion exchange chromatography was performed. The conductivity of the dialysed solution was reduced

to below that of 80 mM potassium phosphate and loaded onto an SP-Trisacryl Plus-M column (Sigma). Cytochrome *c* was eluted with a potassium phosphate gradient (pH 7.0) from 80 mM to 240 mM and fractions were collected. Peak fractions for intact cytochrome *c* were desalted, lyophilised, and stored at -20°C.

2.2.2 Determination of Redox Potential

Redox potentials were determined using the method of mixtures (Wallace et al., 1986) in 50 mM phosphate solutions (pH 7.0). Solutions of known potential were constructed using a mixture of varying ratios of potassium ferro- and ferri-cyanide. For redox couples, the potential of a solution (E) can be expressed by the equation :

$$E = E^{o'} - 0.06 \log\left\{\frac{reduced}{oxidized}\right\}$$

where $E^{o'}$ is the standard redox potential for the pair. In a mixture of two redox couples, E will be the same for both couples, and the equation for ferri-/ferrocyanide and ferri-/ferrocytochrome *c* in solution is thus:

$$E^{o'}_{fercyanide} - 0.06 \log\left\{\frac{ferrocyanide}{ferricyanide}\right\} = E^{o'}_{cytc} - 0.06 \log\left\{\frac{ferrocytochrome}{ferricytochrome}\right\}$$

The ratio of ferri- to ferrocytochrome c is calculated by measuring absorbance at 550 nm and comparing this to the absorbance of fully oxidised and fully reduced solutions. Plots of logs of the two ratios are extrapolated back to the midpoint of the ferri-/ferrocyanide pair to determine $E^{\circ\prime}_{cytc}$. $E^{\circ\prime}_{ferricyanide}$ is taken to be + 0.43 V (Wallace et al., 1986).

2.2.3 Determination of Alkaline Transition

The alkaline transition was followed by monitoring the disappearance of the absorbance band at 695 nm which is characteristic of the Met80 - Fe^{+3} bond. Solutions of cytochrome *c* in 50 mM phosphate were adjusted to pH 6.0 and the pH was increased to pH 11 in approximately 0.2 pH unit increments by addition of KOH solutions, and the absorbance at 695 nm was measured on either a Beckman DU-60 scanning spectrophotometer or on a Hewlett Packard 8452A diode array spectrophotometer. This was then repeated in the opposite direction from pH 11 to pH 6. The midpoints of the resulting titration curves were then determined from the inflection point in a plot of absorbance versus pH.

2.2.4 ATP Affinity Chromatography

Cytochrome *c* affinity for ATP immobilised on an agarose column was measured by retention times. Samples of cytochrome c (50 μL) were mixed with 5 μL ferricyanide

solution to oxidise the protein and the whole sample was loaded onto a 2 cm column (Sigma Chemicals, A2767) which had been equilibrated with 10 mM phosphate buffer (pH 7.0). Upon elution of the ferricyanide peak (void volume) a gradient was initiated from 10 mM phosphate to 150 mM phosphate (pH 7.0). Detection was carried out using a flow-through absorbance detector (Pharmacia). Elution times measured were the difference between elution of the sample and the void volume elution time. Under these conditions, typical elution times for *S. cerevisiae* cytochrome *c* are 50-60 minutes and for horse cytochrome *c* they are 20-30 minutes.

2.2.5 Biological Activity: Succinate Oxidase Assay

Mitochondria were isolated from rat livers that were homogenised in 8.5% sucrose and centrifuged (600g, 10 min) to remove nuclei and cellular debris. Further centrifugation of the supernatant (8500g, 10 min) yielded a pellet of mitochondria that was resuspended and washed two more times with 8.5% sucrose in similar centrifugations.

Osmotic shock of the mitochondrial preparation was effected by resuspension in 15 mM KCl (10 min at 4°) to remove the outer membrane: samples were then centrifuged (6000g, 10 min) and the pellet resuspended in 150 mM KCl (10 min at 4°) to remove endogenous cytochrome *c* before centrifuging again (5000g, 10 min). The 150 mM KCl

treatment was repeated 3 more times to ensure complete removal of endogenous cytochrome *c*, and the mitochondria were resuspended in 8.5% sucrose and kept on ice.

Biological activity of the mitochondria with cytochrome *c* was measured in a Clark-type oxygen electrode. 4.1 mL of buffer (300 mM sucrose, 75 mM glucose, 1mM $MgCl_2$, 12 mM K_2HPO_4, 11 mM succinate, 2 mM ATP and 15mg/132 mL hexokinase (ICN, 100716)) was mixed with 1 mL of mitochondrial suspension in the chamber and allowed to equilibrate for several minutes to achieve a steady baseline oxygen consumption rate. Incremental additions of cytochrome *c* resulted in concomitant increases in oxygen consumption rates, initial rates in the curves of oxygen consumption versus concentration provide a measure of the rate constant for electron transport.

2.2.6 Biological Activity: Cytochrome C Oxidase Assay

Cytochrome *c* oxidase was purified from cow hearts using the technique of Darley-Usmar et al. (1987). Activity was measured using a Clark-type oxygen electrode. The chamber was equilibrated with 5 mL of buffer (50 mM MOPS and 0.3% Tween 80 at pH7.5) , sodium ascorbate was added to a final concentration of 7 mM (70 μL of 0.5 M ascorbate) and TMPD to a final concentration of 0.7 mM (70 μL of a 50 mM TMPD solution, TMPD was recrystallised from 80% acidified ethanol). Addition of cytochrome *c* oxidase at this point gave a baseline oxygen consumption, the chamber was again allowed

to equilibrate for several minutes. Incremental additions of cytochrome *c* resulted in oxygen consumption rate climbing. Initial rates were determined by plotting oxygen consumption versus the amount of cytochrome *c* added and determining the initial slope of the resulting curve, as for the succinate oxidase assay this yields a value for the electron transfer rate.

2.2.7 Determination of Oxidation Half-times

Reduced cytochrome *c*, when exposed to oxygen saturated solutions, will slowly auto-oxidise. Cytochrome *c* was reduced with dithionite and the dithionite removed by passage through a PD-10 desalting column (Pharmacia) using 50 mM phosphate as eluant. Oxidation times are concentration dependent, concentrations of solutions were adjusted to $A_{526.5} \approx 0.18$ (16 mM cytochrome). Oxidation was followed by monitoring the collapse of the 550 nm peak which is characteristic of the reduced state. Readings were taken every five minutes for at least an hour, at the end of this time ferricyanide was added to obtain the fully oxidised reading.

2.2.8 HPLC Retention Times

Cytochrome *c* samples were run through a cation-exchange high performance liquid chromatography (HPLC) column in order to monitor surface charge effects of any

of the mutations. HPLC was carried out on a Waters 600E with an SP5PW column and elution was performed with a gradient from 4 to 400 mM potassium phosphate (pH 7.0) over 20 minutes at 10 mL/minute. Elution times represent the time elapsed between void volume time and elution of the peaks of interest.

2.3 Computer Modelling

2.3.1 Starting Structures

Starting models for cytochrome *c* modelling were obtained either from the Brookhaven Protein Data Bank or directly from the crystallographers. The specific files used for these studies were horse ferrocytochrome *c* (PDB identity code 1HRC) and yeast (*S. cerevisiae*) ferrocytochrome *c* (PDB identity code 1YCC). The starting model for ATP was extracted from the crystal structure of actin/DNase (PDB identity code 1ATN) and was modified to the *syn* conformation for one of the simulations.

All manipulations to the protein structures were executed within the programs Insight II and Discover (Molecular Simulations Inc., San Diego, California). Since X-ray crystallography methods only normally detect the presence of heavy atoms and not hydrogens, explicit hydrogens were added to both protein and explicit water molecules. All bonds were corrected for order and connectivity, this included explicit single bonds

from histidine 18 and methionine 80 to the iron centre. Before proceeding to the simulations hydrogens were minimised for 5000 iterations by conjugate gradient minimisation.

For all calculations in the following sections (2.3.2, 2.3.3) a modified version of the CVFF forcefield was used which included data for the heme group, trimethyllysine, and the methionine and histidine ligands found in cytochrome *c* (supplied by Terry Lo, UBC). Parameters were set automatically and then checked by hand to ensure appropriate atom and residue definitions for the models. Calculations were subjected to cut-offs at 12Å with a switching function applied to the last 1.5 Å. Final models were examined by superposition of the complexes (using the algorithm of InsightII) which superimposes two given structures and provides the root mean square (RMS) deviation between the two structures.

2.3.2 ATP and Horse Ferricytochrome *c* Modelling

Using the crystal structure of the reduced form of Horse cytochrome *c* as a starting model, an energy minimisation simulation of a cytochrome *c*/ATP complex was carried out. Since ATP binds preferentially to the oxidised form of cytochrome *c*, the partial electrostatic charge on the heme iron atom was increased by +1 (Gunner and Honig, 1991). The γ-phosphate of ATP was brought into close proximity to Arginine 91

(3 Å) and the adenine group was placed between lysine residues 86 and 72. Care was taken to avoid any steric clashes between the nucleotide and atoms present in the protein portion of the cytochrome *c* structure. Waters present in the crystal structure were removed where they were in steric conflict with the nucleotide. The entire structure was then surrounded with an 8 Å layer of water to simulate the aqueous environment. This was performed for both simulations including the *syn* and *anti* forms of ATP and in a control model with no ATP.

Minimisation proceeded stepwise, first the sidechains in the area immediately around ATP were allowed to relax (residues within 5 Å, backbone was fixed) for 1000 cycles of steepest descents minimisation. Then all heavy atoms were fixed and the hydrogens were allowed to minimise for 1000 cycles of steepest descents. At this point all atoms were allowed to move, and minimisation proceeded by conjugate gradient minimisation for 10,000 cycles. In the control model, the first cycle of minimisation was skipped.

2.3.3 Modelling Methionine Mutants of Yeast Cytochrome *c*

Starting with the structure of the reduced form of cytochrome *c*, both energy minimisation and molecular dynamics simulations were studied for a variety of mutants of yeast cytochrome *c*. Each of these mutants introduced a methionine residue as a replacement for other residues, specifically these residues were: proline 25, valine 28,

isoleucine 35, lysine 55, leucine 64, and isoleucine 75. A non-substituted model was included of the starting structure, this simulation was repeated twice as a comparison. In all cases the simulations proceeded under the same conditions.

Methionines were replaced using the replacement algorithm internal to InsightII, placing new atoms as close to previous atom positions as possible, then side chains were adjusted by hand to avoid any significant steric clashes. The entire structure was then solvated with a 5Å layer of explicit water molecules. All atoms in the structures were fixed with the exception of the new side chains and minimisation proceeded by steepest descents for 500 iterations. All side chains and the heme were then released and minimised by conjugate gradients for 500 iterations. After this point all atoms were allowed to move and were minimised by conjugate gradients for another 500 iterations. Once the initial rounds of minimisation were complete, the models were subjected to either minimisation by conjugate gradient or dynamics calculations at 300°K, both for 10,000 iterations (10ps for the dynamics runs) with none of the atoms fixed.

CHAPTER 3: RESULTS

3.1 *S. CEREVISIAE* ISO-1 CYTOCHROME *C* MUTANTS

Comparisons between the various mutants of *S. cerevisiae* Iso-1 cytochrome *c* and the wild type (C102T Iso-1 cytochrome *c*; see section 2.1.1) in each of the various assays are summarised in Tables 3.1 to 3.4. The results for each set of mutations are divided into two groups; those involving residues 13 and 90 (Arg13Cys, Arg13Lys, Asp90Cys, Arg13Cys/Asp90Cys) and those at residues 82 and 85 (Phe82Tyr, Phe82Tyr/Leu85Ala, Leu85Ala, Leu85Met, Leu85Cys, and Leu85Phe), and are described in sections 3.1.1 and 3.1.2 respectively.

3.1.1 MUTATIONS OF RESIDUES 13 AND 90

During preparation of the double mutant Arg13Cys/Asp90Cys, anomalously low production of protein occurred, as a result there was not enough material to complete most of the assays. Previous work with this mutant has not shown this behaviour,

Mutant	Redox Potential (±10 mV)	pK 695	t½ Oxidation (hours) (± 50%)
L85C	262	8.6 ± 0.1	3.6
L85F	272	8.7 ± 0.1	8
L85M	267	8.7 ± 0.1	6.3
L85A	277	***7.7 ± 0.1***	2.7
F82Y	266	8.3 ± 0.1	2
F82Y/L85A	271	***7.9 ± 0.1***	2.6
R13C	287	8.9 ± 0.3	3
D90C	276	8.6 ± 0.1	***1.4***
R13C/D90C	N.D.	N.D.	5.4
R13K	278	8.9 ± 0.1	3.9
Wild type	279	8.6 ± 0.1	4

Table 3.1: Physico-chemical Assays of Cytochrome *c* Mutants

Reduction potentials, alkaline transition values (pK 695), and auto-oxidation half times are listed. Wild type refers to *S. cerevisiae* Iso-1-cytochrome *c* with the single amino acid substitution C102T to prevent dimerisation. Error values indicated are standard error and significant values are in bold italics.

Mutant	HPLC retention times (min) (± 0.2)		ATP affinity (min)
	Fe^{2+}	Fe^{3+}	
R13C	***13.7***	***14.6***	***42 ± 1.5***
D90C	15.8	16.8	58 ± 6
R13C/D90C	N.D.	N.D.	N.D.
R13K	***15.5***	***16.5***	55 ± 3.5
Wild Type	15.9	17.0	53 ± 3.5

Table 3.2: Chromatography Characteristics of Cytochrome *c* Mutants (Positions 13 and 90)

Cation-exchange HPLC and ATP affinity chromatography elution times are listed. For the HPLC profiles, Fe^{2+} and Fe^{3+} indicate elution times for the reduced and oxidised forms respectively. Error values indicated are standard error and significant values are in bold italics.

Mutant	HPLC retention times (min) (± 0.2)		ATP affinity (min)
	Fe^{2+}	Fe^{3+}	
L85C	15.1	16.0	51 ± 2
L85F	15.0	16.0	51 ± 2
L85M	15.0	16.0	50 ± 3
L85A	15.0	15.9	51 ± 2
F82Y	15.1	16.1	50 ± 2
F82Y/L85A	14.8	15.4	48 ± 1.5
Wild Type	15.0	16.0	53 ± 3.5

Table 3.3: Chromatography Characteristics of Cytochrome *c* Mutants (Positions 82 and 85)

Cation-exchange HPLC and ATP affinity chromatography elution times are listed. For the HPLC profiles, Fe^{2+} and Fe^{3+} indicate elution times for the reduced and oxidised forms respectively. Error values indicated are standard error.

Mutant	Succinate Oxidase	Oxidase
L85C	56 ± 5	59 ± 4
L85F	95 ± 14	77 ± 3
L85M	78 ± 7	58 ± 3
L85A	92 ± 5	67 ± 5
F82Y	95 ± 6	81 ± 5
F82Y/L85A	48 ± 7	97 ± 4
R13C	34 ± 1	26 ± 3
D90C	72 ± 5	78 ± 6
R13C/D90C	6 ± 1	> 3
R13K	68 ± 6	72 ± 5

Table 3.4: Biological Assays of Cytochrome *c* Mutants.

Numbers represent the relative oxygen consumption in either Succinate Oxidase (depleted mitochondria) or Cytochrome *c* Oxidase systems, and are relative to the oxygen consumption of the wild type (C102T Iso-1 cytochrome *c*). Error values indicated are standard error.

suggesting that this is not to be taken as an indicator of any *in vivo* characteristics of the protein (Huang et al., 1994).

3.1.1.1 Physical Aspects of Mutation at Positions 13 and 90

Structural indicators for these mutants are shown in Table 3.1. The redox potentials are essentially unchanged for all three mutants, results for Arg13Cys and Asp90Cys are in agreement with the results of Huang et al. (1994) who found the potentials in the range 275-280 mV, indicating no significant perturbation of the heme environment or change in solvent exposure to the heme. The midpoint of the alkaline transition (pK_{695}) , an indicator of heme environment integrity reflecting a rearrangement from methionine to lysine coordination of the iron, does not alter noticeably, and marginally increased in the mutants with Arg13 replaced by Cys or Lys. Auto-oxidation times also reflect heme-crevice stability; in these experiments results vary widely, but Asp90Cys shows a significantly lowered time compared to all others. This stability decrease is supported by urea denaturation experiments performed by Huang et al. (1994), who showed no decrease in stability for Arg13Cys but a drop for Asp90Cys. The fact that the pK_{695} transition midpoint does not decrease for Asp90Cys may result from the restoration of the negative charge during the assay. Cysteine -SH has a pKa in the range of 7.0 to 9.0 in proteins (Lundblad and Noyes, 1984; Creighton,1993) which may mean residue 90 is deprotonated; if the negative charge is significant in the stabilizing role of Asp90, then the pH titration may restore the structural integrity.

Chromatography results for all of these mutants show significant deviations, which is expected due to the charged nature of the residues. As seen in Table 3.2 Arg13Lys showed a moderate decrease in cation exchange affinity while Arg13Cys shows a substantial decrease. It is interesting to note that another mutant, Lys55Met, which has a lysine replaced by an uncharged group, shows only a modest decrease in retention time (17.3 and 18.2 minutes compared to 17.6 and 18.6 for wild type (Fe2+ and Fe3+ respectively) (Woods et al., 1996); while the shift in this case is on the order of minutes. The elution times for Asp90Cys were surprisingly unchanged from that of the wild type, suggesting either that this group is deprotonated at pH 7.0 or that the charge difference does not alter the chromatographic characteristics. These differences may indicate that the cation-exchange column interacts with cytochrome *c* in a specific manner, and that differences in retention times reflect modulation of a particular area of the cytochrome *c* surface as opposed to net charge. ATP affinity chromatography reflects a similar series with Arg13Cys eluting significantly faster than the wild type while the other mutants elute at the same time. Asp90Cys values indicate a high range of error, this is due to the relatively low concentrations available; higher loadings may yield more precise results.

3.1.1.2 Biological Aspects of Mutation at Positions 13 and 90

Table 3.4 shows the biological activity of all four 13/90 mutants in the succinate oxidase and cytochrome *c* oxidase assays. The succinate oxidase assay potentially reflects

the interaction of cytochrome *c* and cytochrome *c* reductase (Complex III) (Wallace, 1984). In all of these cases, both assays were affected similarly in each mutation, however modification of Arg13 had a more significant effect than that of Asp90. The relatively conservative mutation Arg13Lys yielded the same activity that the non-conservative mutation Asp90Cys showed, while the non-conservative mutation Arg13Cys showed a more dramatic drop in activity. The double mutant showed a decrease in activity larger than just an additive one for the single mutations (which would be on the order of 25%) indicating a synergistic effect of the two mutants. This is confirmed by the work of Huang et al. who showed similar levels of activity in the cytochrome *c* oxidase assay (1994).

3.1.2 Mutations of Residues 82 and 85

3.1.2.1 Physical Aspects of Mutation at Positions 82 and 85

Again referring to Table 3.1, the physico-chemical assays for the six mutants at positions 82 and 85 reveal a number of differences relative to the wild type. Redox potentials are not significantly changed, this agrees well with the redox potentials determined for these same mutants in other studies (Lo et al., 1995a and b; values determined varied from 280-290 mV at pH 6.0). The alkaline transition midpoint indicates a modest decrease in stability for Phe82Tyr, and more significant drops for

Leu85Ala and Phe82Tyr/Leu85Ala. For the auto-oxidation experiments, none of these six showed a significant increase in the rate of oxidation.

Table 3.3 shows the chromatographic properties for these six mutants. Since none of the residue mutations involve the introduction of charge differences, it is not surprising that none of the retention times are significantly different (repetition of the apparently low Phe82Tyr/Leu85Ala runs in a different set yielded values of 16.7 min and 17.8 min for the reduced and oxidised forms respectively compared to 16.5 min and 17.7min for the C102T wild type reduced and oxidised forms). For the ATP binding test the double mutant shows a marginally lowered affinity; it may be that this assay could be improved to show more subtle effects on ATP binding characteristics through the use of a modified gradient or buffer system..

3.1.2.2 Biological Aspects of Mutation at Positions 82 and 85

The activity of the six mutants in the succinate oxidase and cytochrome *c* oxidase assays indicates an interesting series (Table 3.4). In terms of ranking, activity in the succinate oxidase assay for the six mutants can be ordered as follows:

Wild Type = Leu 85Phe, Phe82Tyr, Leu85Ala > Leu85Met > Leu85Cys, Phe82Tyr/Leu85Ala

while in the cytochrome *c* oxidase assay the series is:

Wild Type = Phe82Tyr/Leu85Ala > Leu85Phe, Phe82Tyr > Leu85Ala > Leu85Met = Leu85Cys.

One aspect to be considered when comparing these two assays is that the cytochrome *c* oxidase assay was carried out at pH 7.5 compared to pH 7.0 for the succinate oxidase assay. For most of the mutants this will not make a difference, however for both Leu85Ala and Phe82Tyr/Leu85Ala it may. The fraction of protein in the native form can be determined from the following equation:

$$\log\left[\frac{1-x}{x}\right] = pH - pK_{695}$$

where x is the fraction of protein found in the native form (Pearce et al., 1989). Under these conditions Leu85Ala will have an x = 0.61 or 61% (± 6%) native at pH 7.5 compared to 0.83 or 83% (± 6%) at pH 7.0. If we assume that the activity in the cytochrome *c* oxidase assay is entirely due to the native form then the activity will be adjusted to the range of 90 to 110 %. This is supported by an experiment where this assay was carried out at 12°C, under these conditions the activity of Leu85Ala is indistinguishable from wild type. For the double mutant Phe82Tyr/Leu85Ala the native fraction will be 0.72 at pH 7.5 and 0.89 at pH 7.0, indicating a potential activity greater than wild type of 120% for the cytochrome *c* oxidase assay. It should be noted, however, that this correlation is not necessarily dependable since the pK_{695} titration and these biological assays are not performed under identical buffer conditions and other factors may affect activity, particularly the presence of detergent in the cytochrome *c* oxidase assay. The largest relative differences between the two different assays are the drop in Leu85Met and the increase in Phe82Tyr/Leu85Ala activities.

3.2 ATP and Horse Ferricytochrome c Modelling

The final models for the *syn* and *anti* simulations are shown in Figures 3.1 and 3.2, both superimposed over the wild type, while Figures 3.3 and 3.4 show the amino acids near the binding region more clearly for the two simulations. Tables 3.5 and 3.6 list the major movements observed for the charged atoms, as well as the side chains that are found proximal to the nucleotides. It is apparent that the lysine side chains nearby move to enfold the phosphate groups; all three of the lysine residues 86, 87, and 88 move 2-4 Å towards the negative charges, although the contacts for the *anti* model are much closer, particularly for Lys88. Glu69 twists during the simulations, allowing for both the hydrogen bond to form to the ribose hydroxyl group and to decrease the electrostatic repulsion with the phosphate groups. For Arg91, the η-nitrogens move approximately 1 Å toward the γ-phosphate in the simulations; however, the α-carbon does not move significantly. This reflects the general aspects revealed in both models: binding of the nucleotide causes conformational shifts of surface side chains and not of the backbone.

The larger number of proximal residues in the *anti* model can be attributed primarily to the presence of the six-membered ring packing close to the protein surface. In the *syn* model, this group projects away from the protein surface, making no close

A: Movement of Close Charged Sidechains

Side Chain:Atom	ATP atom	Final Distance	Initial Distance	Distance Moved
Glu69:Oε1	Ribose O2'	2.58 Å	3.54 Å	1.38 Å
Lys86:Nζ	(Pα)-O	5.69 Å	7.79 Å	3.40 Å
Lys87:Nζ	(Pα)-O	3.91 Å	5.27 Å	4.02 Å
Lys88:Nζ	(Pγ)-O	2.64 Å	6.55 Å	3.98 Å
Arg91:Nη1	(Pγ)-O	2.62 Å	3.37 Å	1.23 Å
Arg91:Nη2	(Pγ)-O	3.44 Å	4.61 Å	1.19 Å

Glu 69:Oε2 moves 3.46 Å away from the phosphate groups

B: Proximal Residues to the Nucleotide

Glu69	Ala83	Lys86	Arg91
Asn70	Gly84	Lys87	
Pro71	Ile85	Lys88	

Table 3.5: Side Chain Movement and Proximity: *Anti* ATP Model

A) Distances are listed between atom positions in the control (no ATP) and in test model. Distances were calculated by superimposing the control and test models, then measuring distances between atoms in the nucleotide and the two protein models. Distance moved refers to the distance between the same atom in the control and test models. B) Side chains are listed as being proximal if a non-hydrogen atom is found within 5 Å of the nucleotide.

A: Movement of Close Charged Sidechains

Side Chain:Atom	ATP atom	Final Distance	Initial Distance	Distance Moved
Glu69:Oε1	Ribose O3'	2.44 Å	4.88 Å	2.73 Å
Lys86:Nζ	(Pα)-O	7.42 Å	7.63 Å	2.42 Å
Lys87:Nζ	(Pα)-O	4.51 Å	6.84 Å	3.44 Å
Lys88:Nζ	(Pγ)-O	5.32 Å	7.67 Å	3.68 Å
Arg91:Nη1	(Pγ)-O	2.82 Å	3.12 Å	0.42 Å
Arg91:Nη2	(Pγ)-O	2.77 Å	3.87 Å	1.12 Å

Glu 69:Oε2 moves 3.01 Å away from the phosphate groups

B: Proximal Residues to the Nucleotide

Glu69	Ala83	Lys86	Lys88
Asn70	Gly84	Lys87	Arg91

Table 3.6: Side Chain Movement and Proximity: *Syn* ATP Model

Information listed is the same as for Table 3.5.

contacts to the protein. The difference in orientation also allows for an explanation of the major covalent adducts formed during reactions of the ATP analogue with cytochrome *c*. It is possible to manipulate the lysine residues manually in each of the models to their closest approach distance from the C8 position of the nucleotide (where the reactive azido group is), side views of the nucleotide binding site are shown in Figures 3.5-3.6. Only when the nucleotide is in the *anti* conformation is it possible to bring lysines 86 and 87 near this position. Lys86 is much closer to a reactive position (3.12 Å, Fig. 3.5a), while Lys87 is blocked from a very close approach by steric clashes with Lys86 (7 Å, Fig. 3.5b). In the *syn* conformation, Lys72 is close enough to react with an azido group at this position (3.5 Å, Fig. 3.6). This suggests that the different adducts are dependent on the conformation of the nucleotide. Presumably in the natural nucleotide the conformation would be much more biased toward the *anti* conformation since the *syn* conformation is favourable only in the presence of a substituent at the 8 position (Saenger, 1984). One other aspect of the TNP-8N_3-ATP analogue is the attachment of the fluorescent group to the ribose hydroxyls, this group could be accommodated with a slight shift of the ribose away from Glu 69 with necessary loss of the hydrogen bond, but this may be compensated for by a charge interaction between the TNP group (which bears a negative charge) and Lys73. The electrostatic interactions between the phosphates of ATP and the positively charged residues appears to be the most crucial mediator of binding, as previously predicted (Craig and Wallace, 1991).

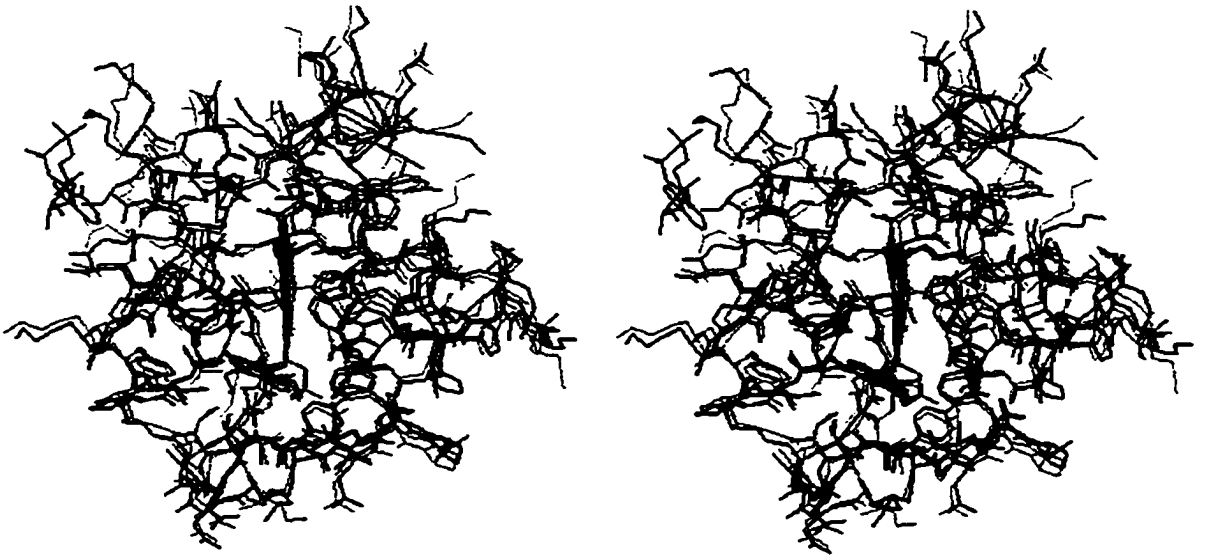


Figure 3.1: Final Models for ATP binding: *Anti* ATP

Full models of the protein including side chains (water shell is not shown) and nucleotide are superimposed over the control run. The simulation including ATP is shown in black, the control structure in grey. The nucleotide is visible in the upper left-hand portion of the picture.

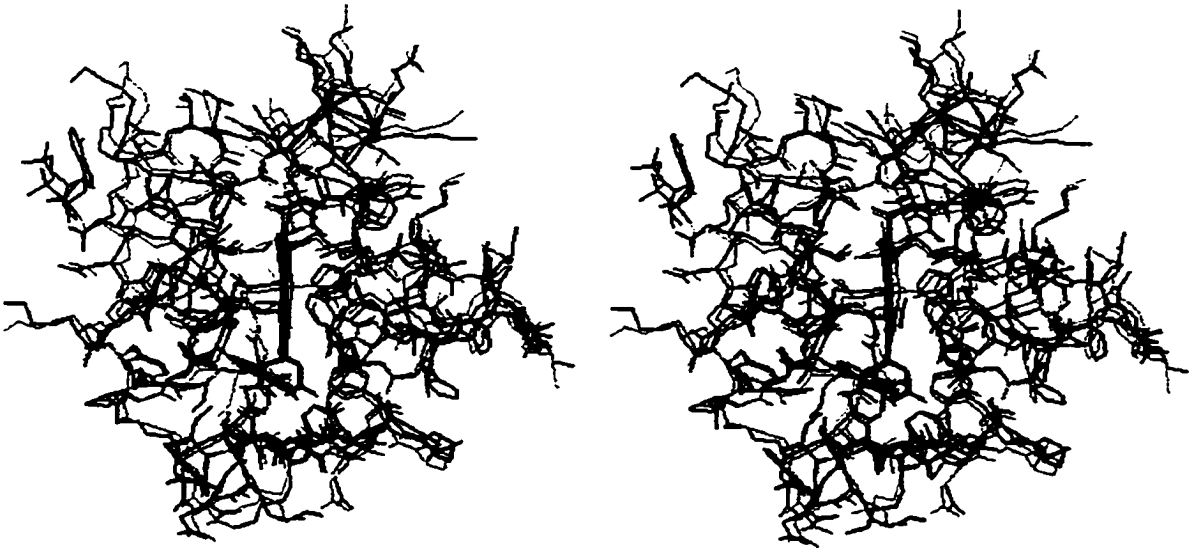


Figure 3.2: Final Models for ATP binding: *Syn* ATP

The figure is identical to 3.1 in description, this time ATP is present in the *syn* conformation.

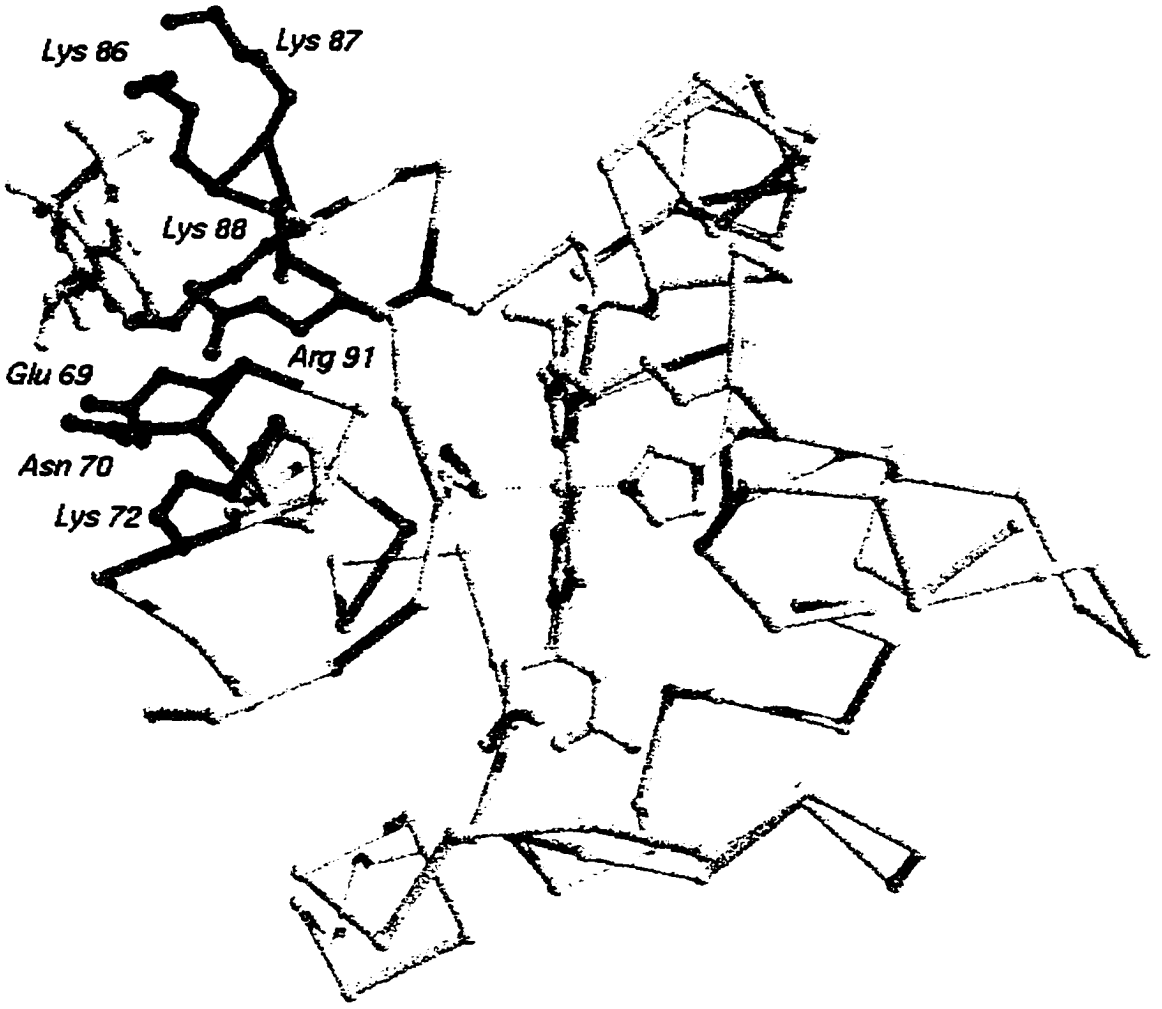


Figure 3.3: Final Models for ATP binding: *Anti* ATP

Side chains involved in nucleotide binding are shown in black and the side chains are shown explicitly. The rest of the protein is shown as an α - carbon trace with the exception of the heme group and the heme binding residues Cys14, Cys17, His18, and Met80.

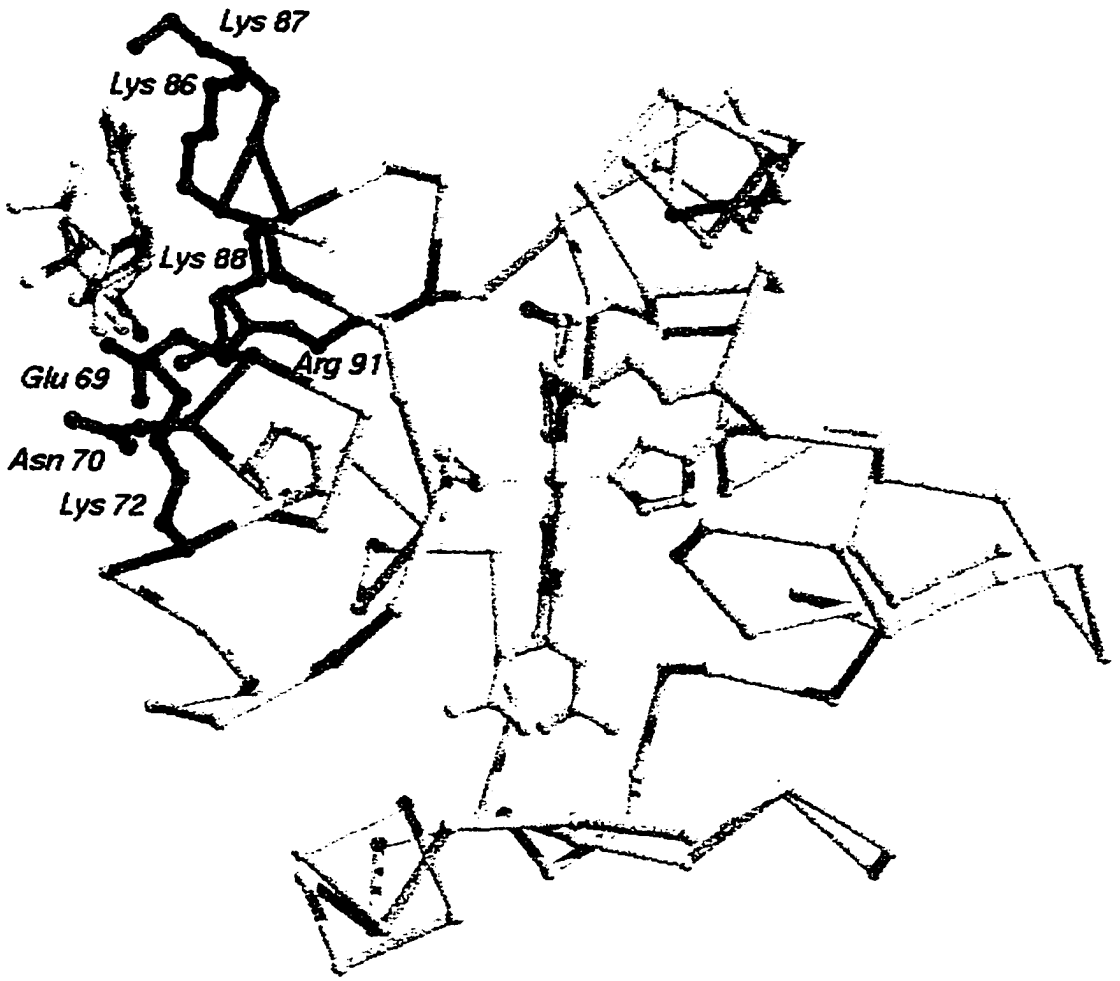


Figure 3.4: Final Models for ATP binding: *Syn* ATP

Cytochrome *c* and ATP are shown for the *syn* nucleotide conformation as in Figure 3.3.

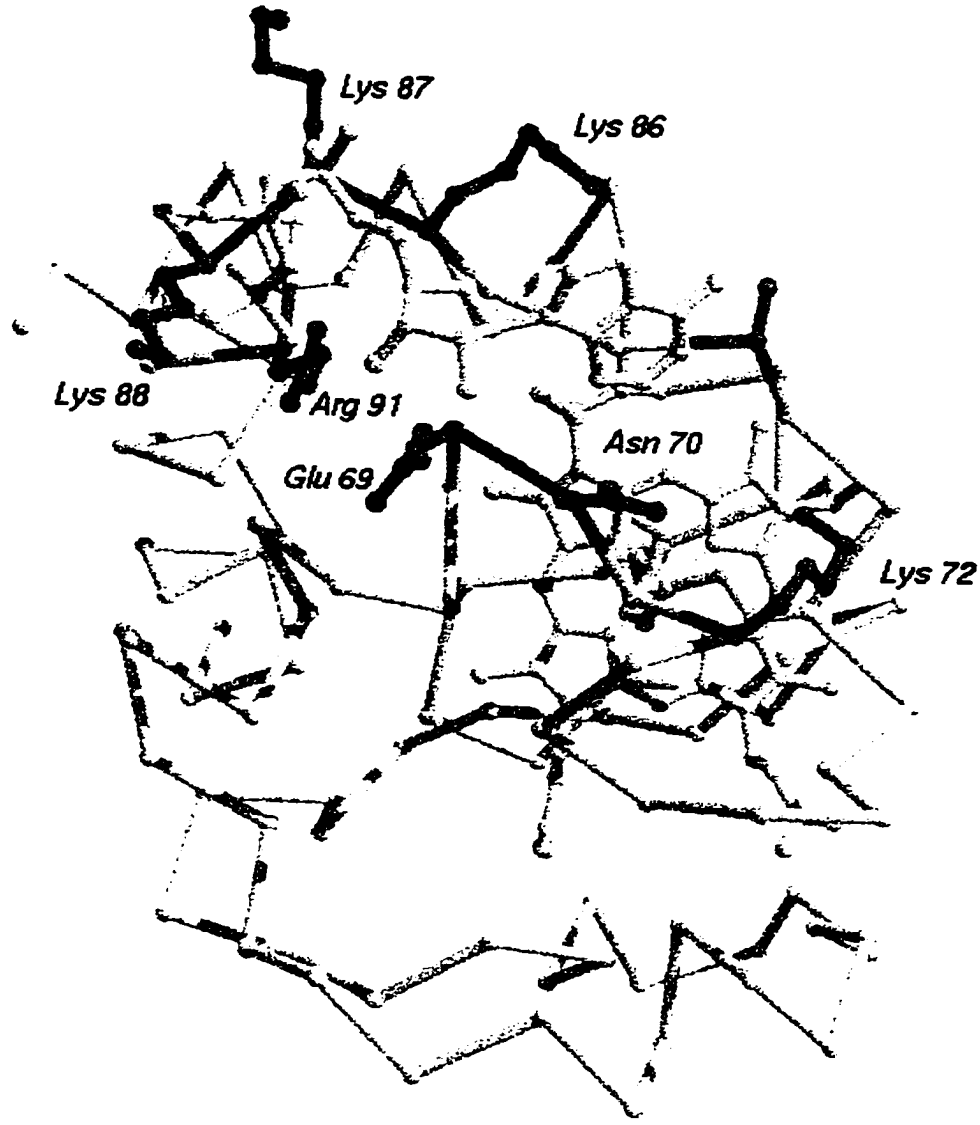


Figure 3.5a: *Anti* ATP and cytochrome *c* with Lys86

Side view (rotated 90° to Figure 3.3) of the nucleotide binding in the *anti* conformation. Lysine 86 has been rotated to its closest approach to the C8 atom of ATP. Cytochrome *c* side chains are shown as in Figure 3.3.

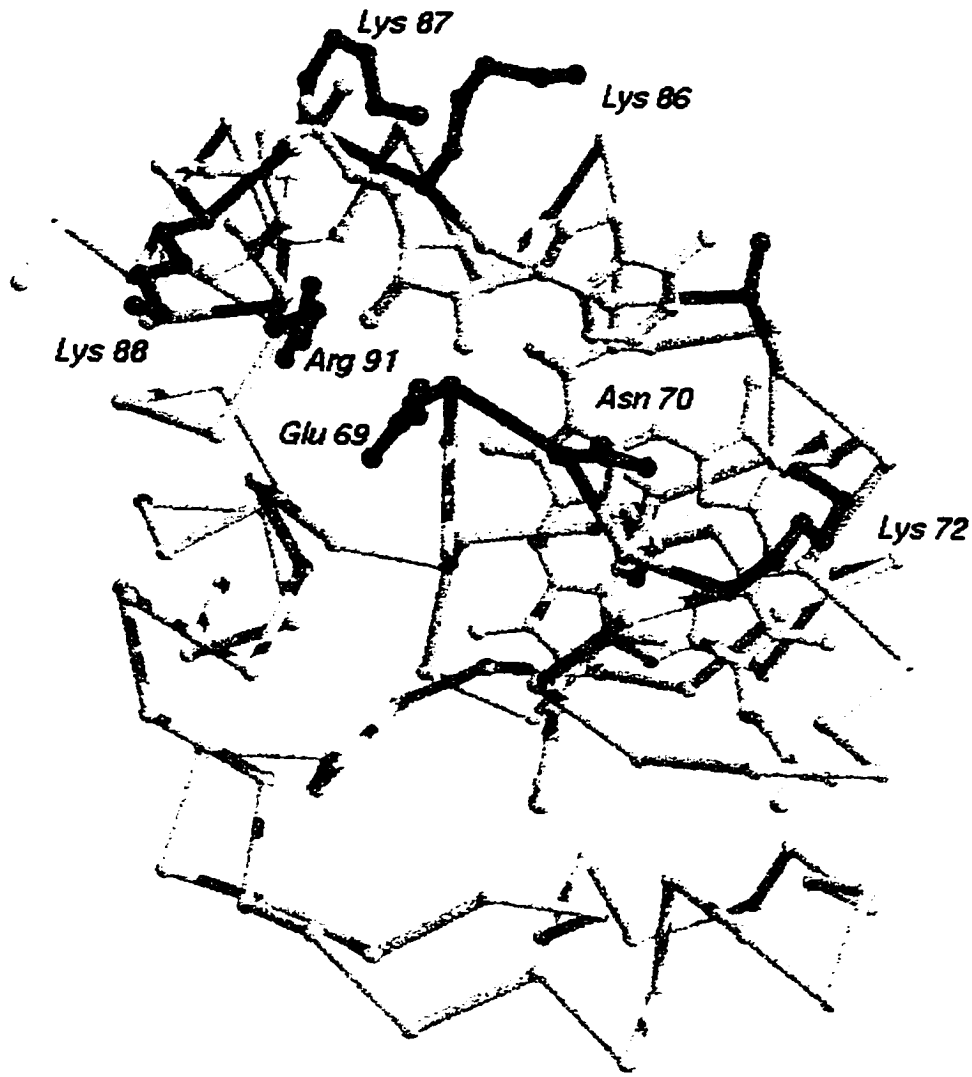


Figure 3.5b: *Anti* ATP and cytochrome *c* with Lys87

Same view is shown as for Figure 3.7, showing closest approach of Lys87 to the C8 of ATP.

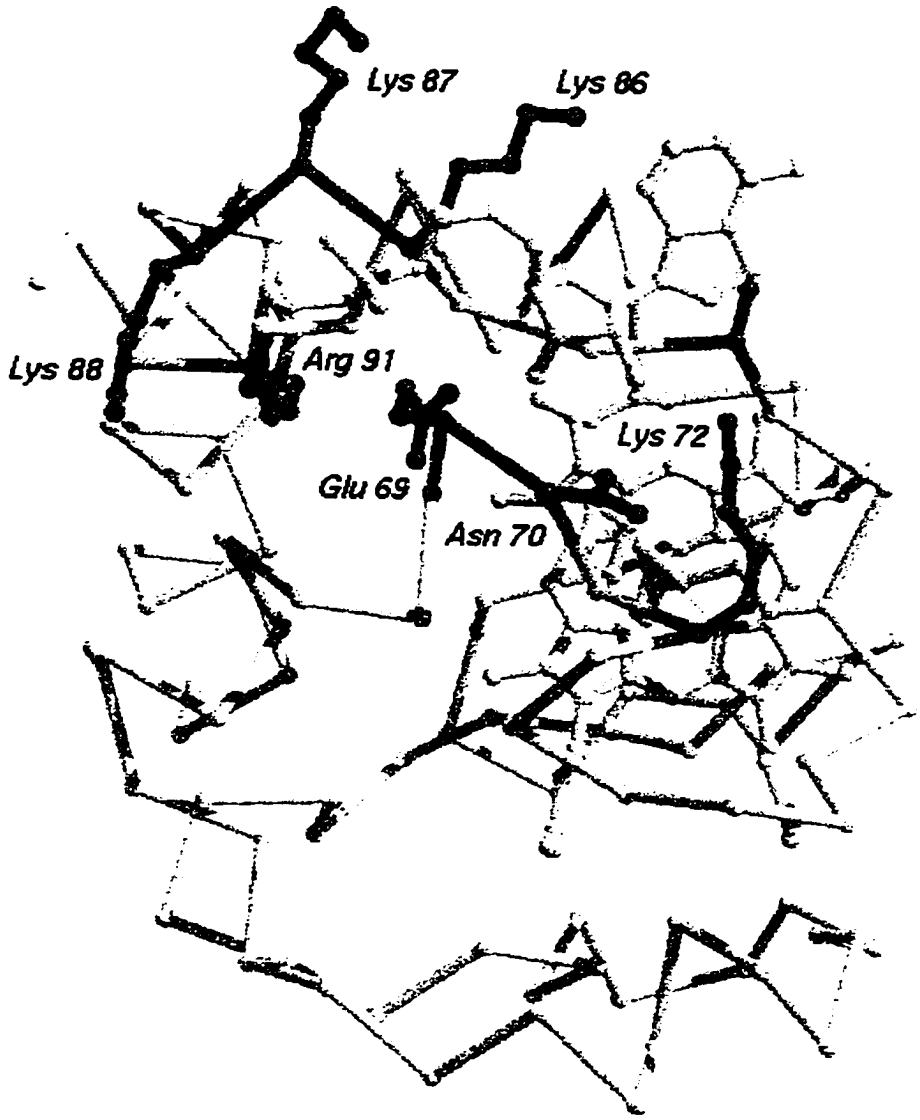


Figure 3.6: *Syn* ATP and cytochrome *c* with Lys72

Side view (rotated 90° to Figure 3.4) of the nucleotide binding in the *syn* conformation, side chains of cytochrome *c* shown as in Figure 3.4. Lysine 72 has been rotated to its closest approach to the C8 atom of ATP.

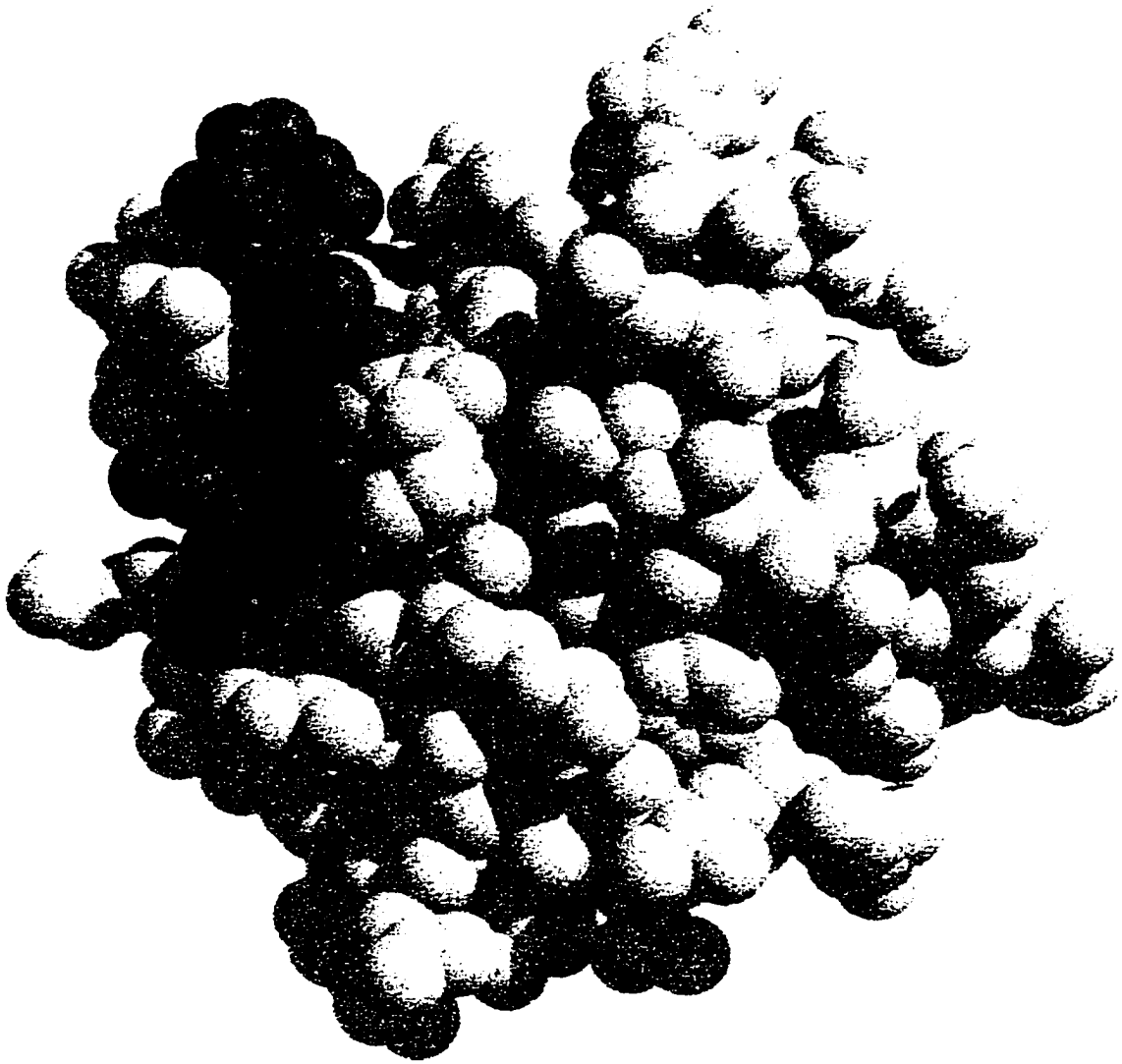


Figure 3.7: Space Filling Model of *anti* ATP and cytochrome *c*

The same view of cytochrome *c* and *anti* ATP as Figure 3.3 with all non-hydrogen atoms shown as spheres. Orientation and colouring scheme are the same as that of Figure 3.3.

A space filling model of the nucleotide in the anti-conformation and cytochrome *c* is seen in Figure 3.7. The two fit snugly together with ATP falling in between the alpha helix from 61-70 and the amino acids which form a ridge from 83-88. This places the nucleotide just on the edge of the interaction area for cytochrome *c* with both the reductase and cytochrome *c* oxidase, which extends from lysines 86 and 87 (on the other side of these side chains relative to the ATP binding site) to residues 8 and 13 and also including to a lesser extent residue 72 (based on lysine modification studies) (Pettigrew and Moore, 1987). This implies that there may not be a strong steric interference; however, the introduction of the negative charges in the triphosphate tail will clearly modify the electrostatic field in this area. The binding of ATP leads to two major electrostatic effects: first, the overall net charge in the ATP/cytochrome *c* complex is reduced by four relative to cytochrome *c* alone; secondly, the dipole moment of the complex will be shifted relative to the dipole moment of cytochrome *c* alone. Both of these differences could be involved in the proposed mechanisms for ATP regulation of electron transport by interfering with cytochrome *c* interactions with either the inner mitochondrial membrane or with physiological partners (Craig and Wallace, 1995).

3.3 Modelling of Methionine Mutants of Cytochrome c

3.3.1 Locations of Methionine Insertions

Before describing the results of the modelling, it is instructive to understand where all of the different amino acids chosen for mutagenesis were located and what the expectations were. A figure showing all of the positions for new methionines is shown in Fig. 3.8. The different residues which were replaced were: Pro25, Val28, Ile35, Lys55, Leu68, and Ile75. For the purposes of semisynthesis via cyanogen bromide cleavage, these were constructed in a background mutant Met64Leu (with Cys102Thr) to prevent the additional cleavage event; however, for most of the models this mutation was not included as part of the calculation. The exceptions were Ile35Met and Leu68Met which are relatively close to residue 64 in the three-dimensional models and may be expected to have some repercussions, particularly for Leu68Met which is in physical contact as the two residues are on successive turns of an α-helix.

Pro25 is a highly variant residue in most cytochrome *c* sequences (natural residues are Lys, Pro, Ala, Gly, Asp) which is on the surface of a loop from residues 18-32. The variability from the conformationally restricted proline to the highly flexible glycine implies that sidechain effects on the backbone torsions are not necessary to maintain the

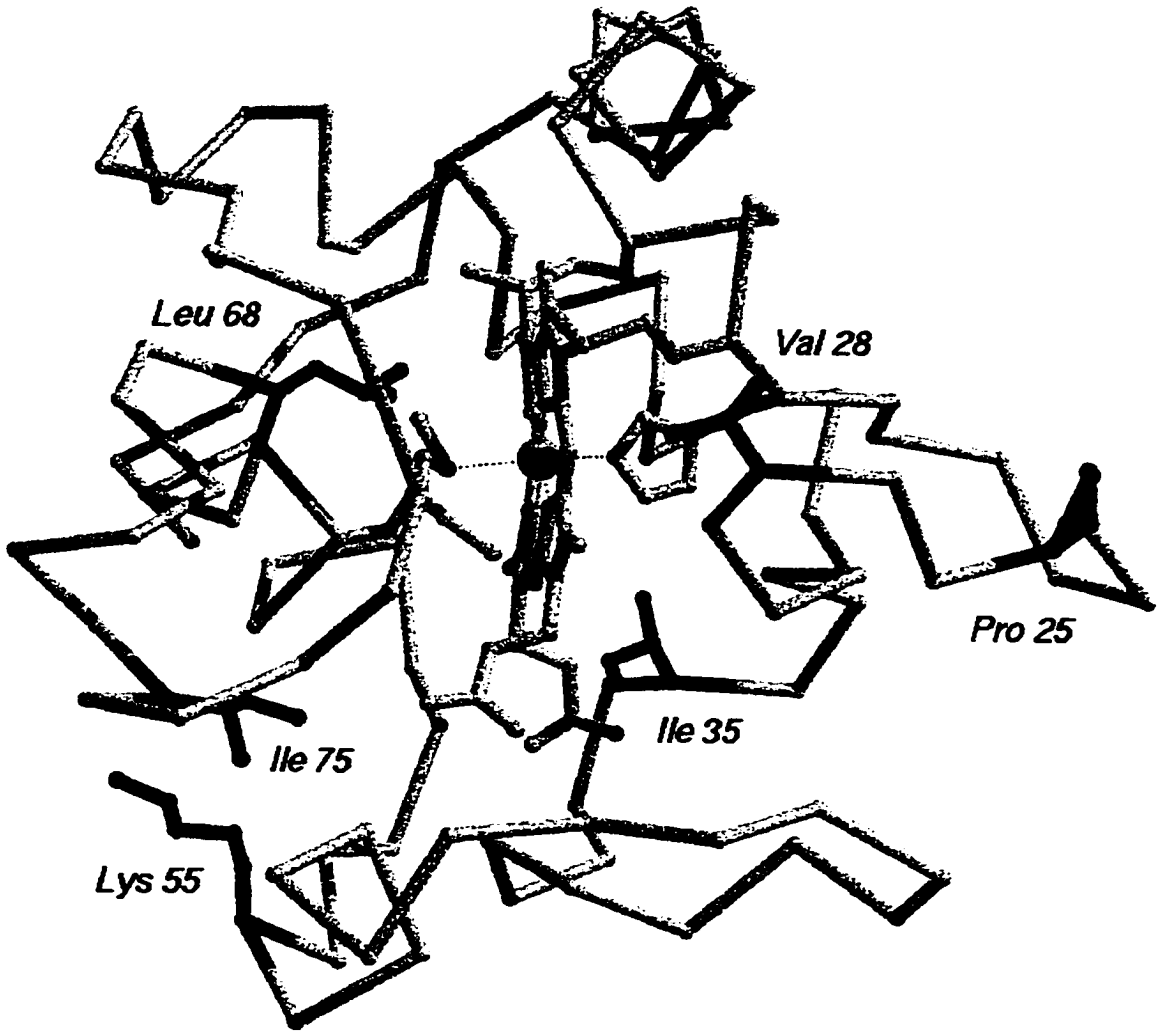


Figure 3.8: Positions of Methionine Insertions

Insertion points are shown as explicit side chains in black and are labelled. The rest of *S.cerevisiae* iso-1 cytochrome *c* is shown as an alpha carbon backbone with the exception of the heme and associated residues Cys14, Cys17, His18, and Met80.

conformation; thus substitution of proline with methionine may not result in any loss of backbone structural integrity.

Val28 is a surface exposed residue near the "front" face of cytochrome *c*. It is close to the heme and the amino acids which appear naturally at this position are isoleucine, threonine, and glutamine. The side chain itself would not be expected to have significant steric clashes as it is on the surface, however the additional conformational flexibility conferred to the backbone by replacement of the branched β-carbon may have some significant effects on the structure.

Ile35 is buried near the "back" of the heme group, and natural variants of this amino acid are leucine, valine, alanine, and phenylalanine. The side chain forms a portion of the wall of a hydrophobic pocket (a space with the potential capacity to accommodate two water molecules but which is empty) (Lo et al., 1995b); the presence of this pocket means the side chain has some conformational freedom, potentially useful in accommodating small side chain rearrangements, particularly as methionine is a more extended amino acid than isoleucine.

Lys55 is a surface exposed residue in the bottom Ω-loop of cytochrome *c*. Natural variants of this residue include alanine, serine, arginine, and methionine, suggesting that the charged character of this group is not necessary for maintenance of function.

Leu68 is a buried residue near the other side of the hydrophobic pocket near Ile35. It is also located on the interior portion of the 60s alpha helix, which extends from residue 59 to 70. One of the initial sites for cleavage in cytochrome *c* was at position 65 in this same helix (Met65 is a surface residue in horse cytochrome *c*). While it is buried, like Ile35 it is part of the internal hydrophobic pocket allowing for the same potential steric freedom.

Ile75 is another buried residue located near the "bottom" of the heme group. natural variants for this residue include the small hydrophobic residues valine, isoleucine and methionine. It is also located near the internal water switch (Wat166) which is significant in the redox-linked conformational changes which occur in cytochrome *c* (Berghuis et al., 1994). Additionally, a three dimensional structure for this mutant (*S.cerevisiae* iso-1 cytochrome *c* Ile75Met/Cys102Thr) has been determined, and thus will form a control for this work (Rafferty et al., PDB entry code 1IRV).

3.3.2 Modelling Results

Minimisation results for all of the different mutants did not show any significant changes in either side chain or backbone conformations when compared to the wildtype. The three buried side chains Ile35, Leu68, and Ile75 would be expected to have more significant steric rearrangements, however the similar sizes of these residues to methionine allowed accommodation of the new side chains with minimal apparent repercussions. As

well, the hydrophobic cavity may compensate for effects of the buried residues Met35 and Met68. Since minimisation runs the risk of falling into energy wells without exploring a larger range of conformations, dynamics simulations were carried out to see if there might be more subtle ramifications to any of the replacements.

During dynamics simulations proteins approximated energetic equilibrium after approximately 2ps. The side chain conformations and modelling results over the course of the simulations are shown in Figures 3.9-3.19. Tables 3.7 and 3.8 describe the backbone and sidechain dihedral angles respectively, and Table 3.9 shows the superposition results and root mean square (RMS) deviations for the different mutant simulations in comparison to the wild type simulations. The RMS values were calculated for three different sets of atoms. The first two columns compare the entire protein backbone and α-carbon positions respectively, in order to evaluate the global differences in structure. Local perturbations were then evaluated by superimposing the backbones of proximal residues to the methionine residues in either the linear sequence (RMS_{BB-21}) or in three dimensional space ($RMS_{BB-10Å}$). The specific aspects noted for each of the simulations are described below.

3.3.2.1 Pro25Met

The replacement of proline with methionine will allow the value for ϕ in the backbone to explore a much larger conformational space, since ϕ is restricted to $60° \pm 20°$

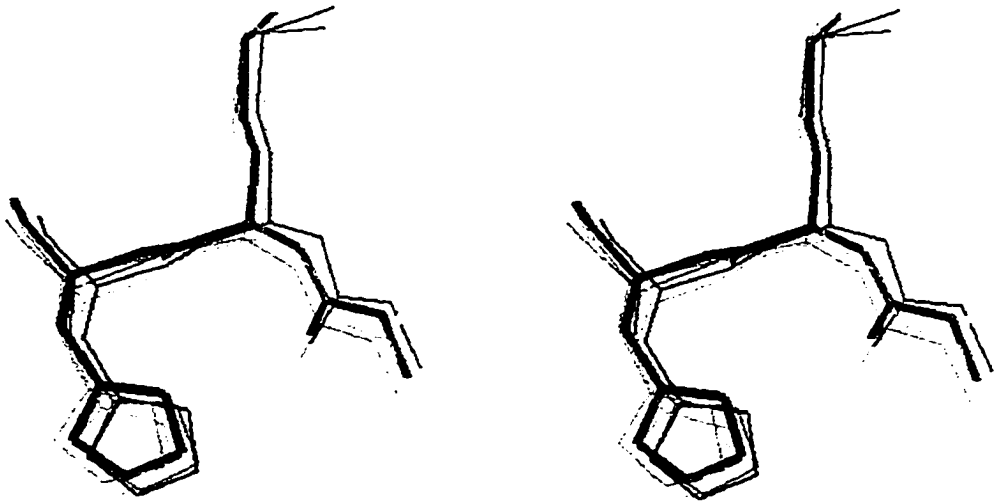


Figure 3.9: Dynamics Conformations for Pro25Met

The conformation of the three residues Gly24-Met25-His26 are shown over the course of the dynamics run. Five time points are indicated explicitly: 2ps, 4ps, 6ps, 8ps, and 10ps; superpositions were carried out for all heavy atoms (non-hydrogen) between each time point and the average. Each time is indicated with the colour range from light grey (2ps) to black (10ps). The average structure (from 2ps to 10ps) is shown as a thick black line.

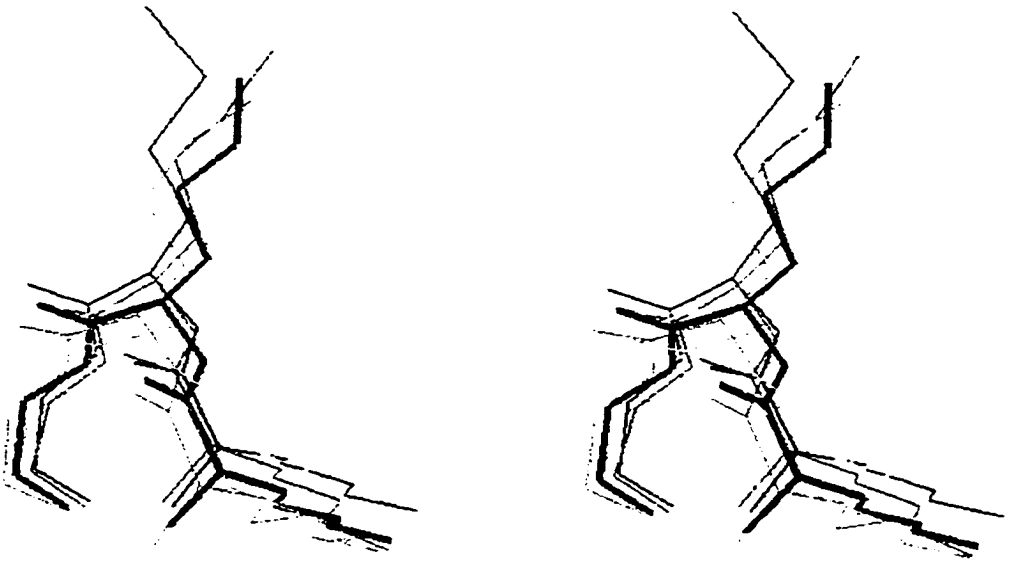


Figure 3.10: Dynamics Conformations for Val28Met

Conformations of the three residues Lys27-Met28-Gly29 are shown. The figure is constructed identically to Fig. 3.9.

(Schultz and Schirmer, 1979). Table 3.7 shows that despite the side chain replacement the backbone dihedrals remain largely the same. Figure 3.9 shows the backbone going through a fairly large range of movement; at 2ps the ϕ angle is -143°, an inaccessible angle for proline. The preferred conformation is essentially that found in the wild type, implying that conformation is not dependent on the amino acid, a fact which is supported by the large variety of residues found at this position in the sequence.

3.3.2.2 VAL28MET

The replacement of valine with methionine leads to no significant shifts in either ϕ or ψ values (Table 3.7). The side chain dihedral χ_3 fluctuates between -90° and 180° (Table 3.8); however, since this side chain projects into the solvent the degree of variability is not due to any significant shifts in the rest of the protein structure.

3.3.2.3 ILE35MET

Table 3.7 indicates that after 4ps in this simulation the value for ϕ changes from -75° to -140°. Both χ_1 and χ_2 remain near 180°, however χ_3 switches from -90° to 180° after 8ps (Table 3.8). This change, as seen in Figure 3.11, involves a reorientation of the C-S-C atoms away from the hydrophobic pocket. This residue is one of the most constrained of the chosen side chains, and movement of proximal side chains to

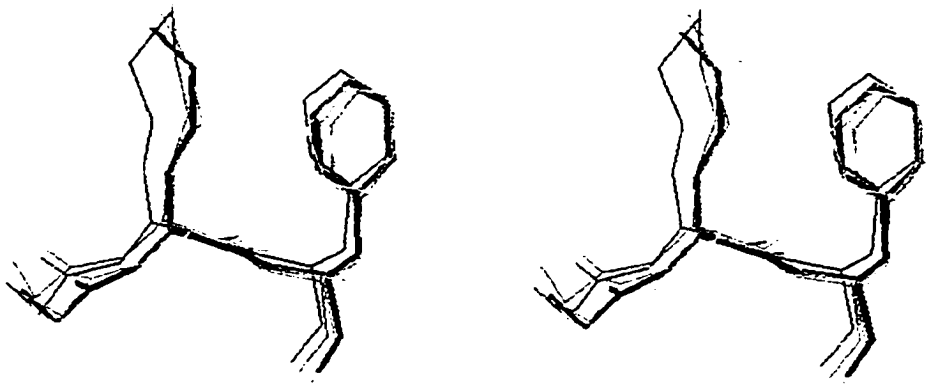


Figure 3.11: Dynamics Conformations for Ile35Met

Conformations of the three residues Gly34-Met35-Phe36 are shown. The figure is constructed identically to Fig. 3.9.

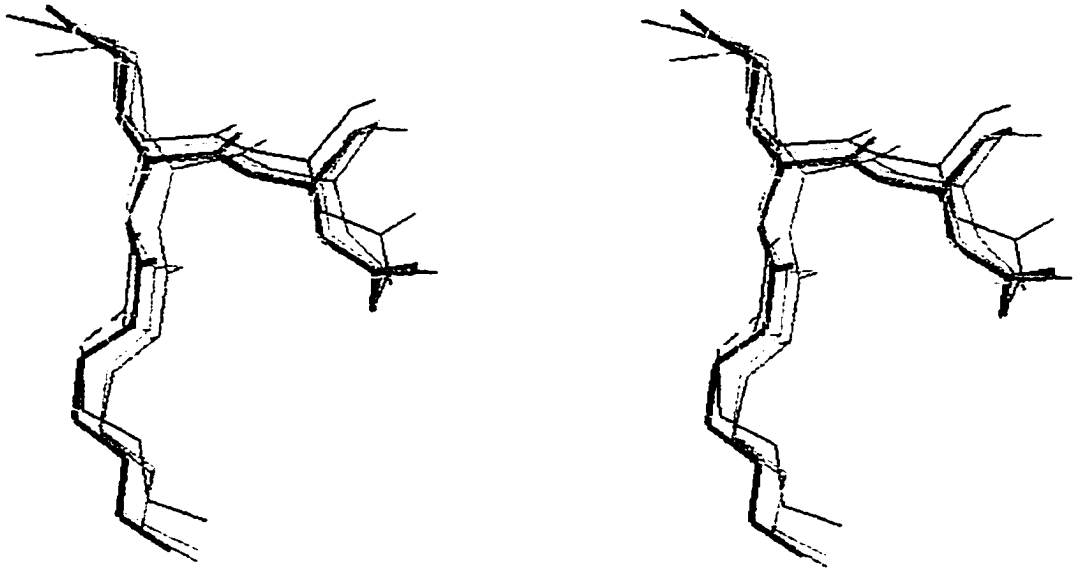


Figure 3.12: Dynamics Conformations for Lys55Met

Conformations of the three residues Lys54-Met55-Asn56 are shown. The figure is constructed identically to Fig. 3.9.

compensate for this replacement appears to be responsible for the larger deviations in local superposition values (Table 3.9).

3.3.2.4 Lys55Met

The replacement of this surface residue with methionine does not result in significant alterations of either ϕ or ψ dihedral values (Table 3.7). The slight variability of χ_3 is not exceptional either; as was observed in the Val28Met mutant this residue's flexibility would not be expected to have significant repercussions on the rest of the protein side chain conformations. It is noticeable, however, in the RMS superposition values in Table 3.9 that this mutant shows consistently high deviations from the wild type relative to the other mutants which merits a more detailed analysis.

The side-chain amino group of Lys55 forms a hydrogen bond to the carbonyl group of Tyr74, and may thus contribute to the structural integrity of the bottom Ω loop of cytochrome *c*. Replacement with methionine necessarily removes this interaction, and may result in the structural deviations observed. In the wild type simulations (and in all other simulations with lysine at position 55) the hydrogen bond is maintained, and as a result the distance between α-carbons for residues 55 and 74 is relatively constant. As can be seen in the graph in Figure 3.13, this distance increases significantly in the Met55 simulation. The structural movements which are associated with this distance increase are shown in Figure 3.14, which superimposes one of the average wild type structures with

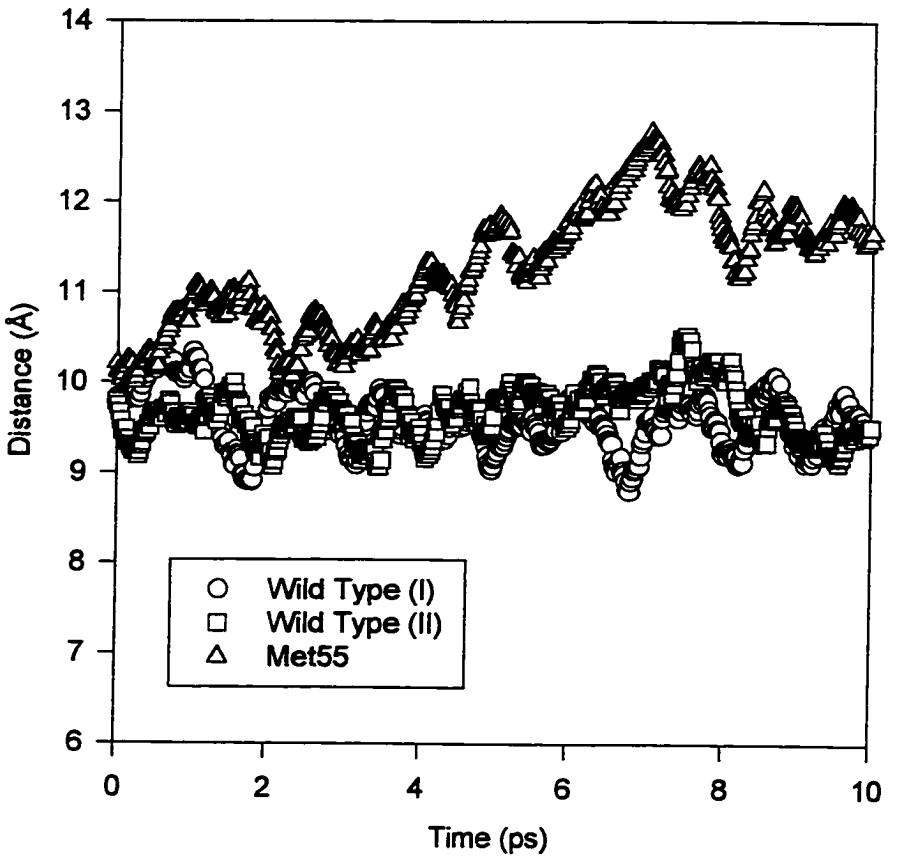


Figure 3.13: Interatomic Distances for Residues 55 and 74

The distance between α - carbons for the residues at position 55 and 74 are plotted over the course of three simulations. The two wild type simulations have lysine residues at position 55 which is replaced by methionine in the Met55 simulation.

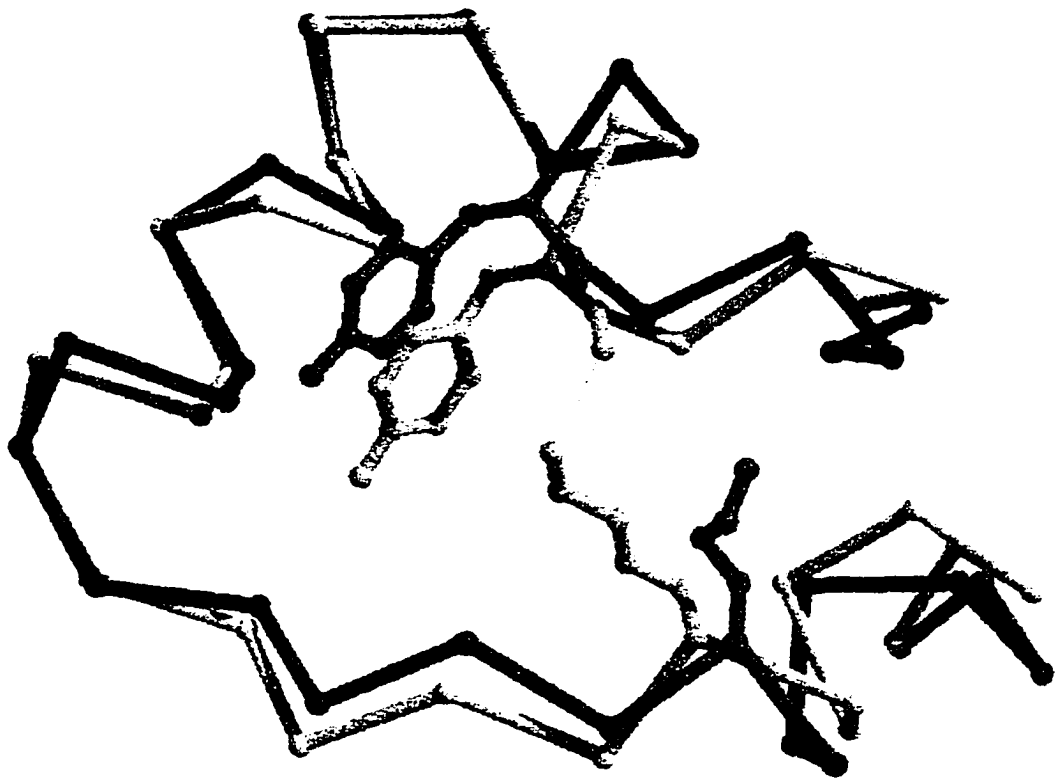


Figure 3.14: Superposition of Two Dynamics Simulations:

Lys55Met and Wild Type

Superposition of average structures from 2ps to 10ps for the dynamics runs of Lys55Met (dark grey) and one Wild Type (light grey) is shown. All residues from 48-79 are shown as alpha carbons with the exception of residues 55 and 74. The hydrogen bond between Lys55 and Tyr74 is shown as a dotted line.

the average Lys55Met structure for the protein from residues 48 to 79. It is apparent that the loss of this interaction leads to significant secondary effects in the simulation, particularly to the backbone conformation in the alpha helix from 49 to 55 and in those residues near Tyr74.

3.3.2.5 Leu68Met and Met64Leu/Leu68Met

These two simulations were carried out to examine the potential effects of both mutations on the structure. Table 3.7 shows that values of ϕ and ψ do not vary significantly, however Table 3.8 shows the side chain values do vary between the two runs. The reason can be directly attributed to the presence of the leucine at position 64, as seen in Figures 3.15 to 3.18. Figures 3.15 and 3.16 show the Met68 conformations at time intervals for the single mutation model Leu68Met and the double mutant Met64Leu/Leu68Met respectively. These two figures are superimposed in Figure 3.17 showing the significant difference between the two different sets of side chain conformations. The leucine which is present at position 64 in the double mutant is visible in Figure 3.18, the presence of the γ-methyl group is located in close proximity and creates a steric conflict requiring the alternate configuration. Despite the differences in side chain movement, the RMS deviations between the two simulations do not differ significantly from each other (Table 3.9) showing that the replacements do not result in more global conformational rearrangements.

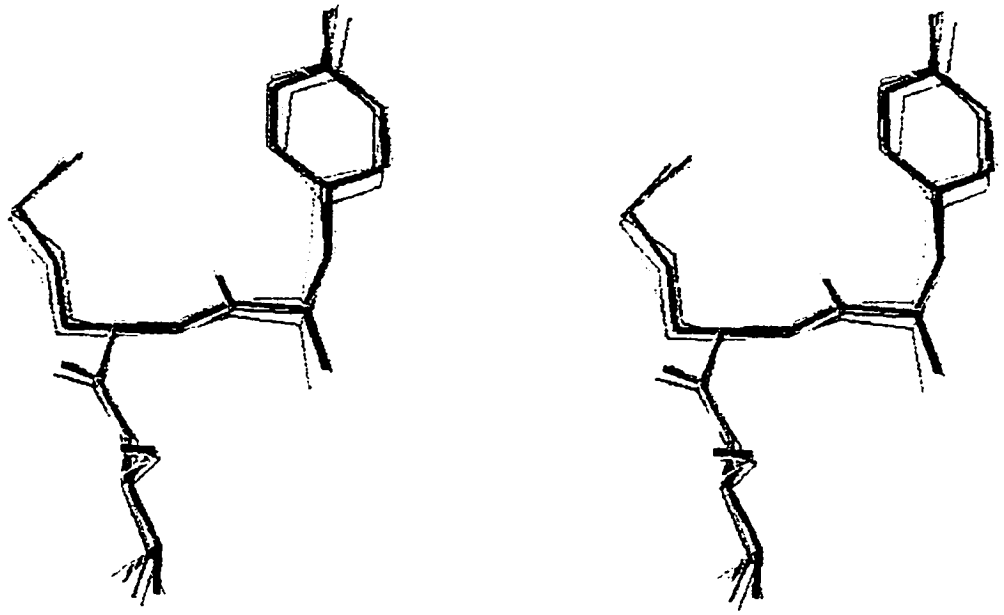


Figure 3.15: Dynamics Conformations for Leu68Met

Conformations of the three residues Tyr67-Met68-Thr69 in the single mutant Leu68Met are shown. The figure is constructed identically to Fig. 3.9.

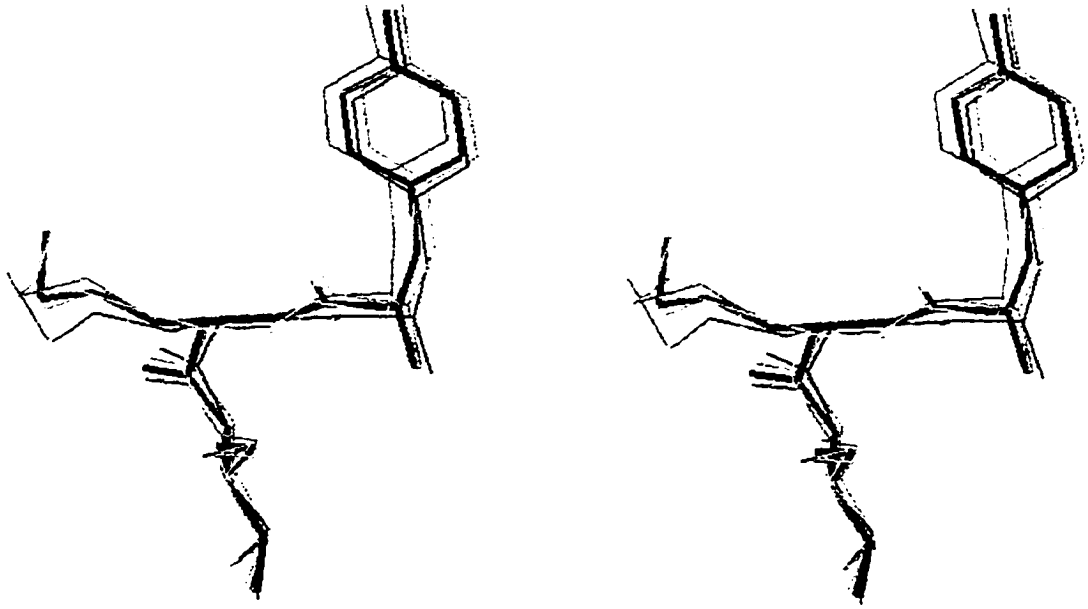


Figure 3.16: Dynamics Conformations for Met64Leu/Leu68Met

Conformations of the three residues Tyr67-Met68-Thr69 in the double mutant Met64Leu/Leu68Met are shown. The figure is constructed identically to Fig. 3.9.

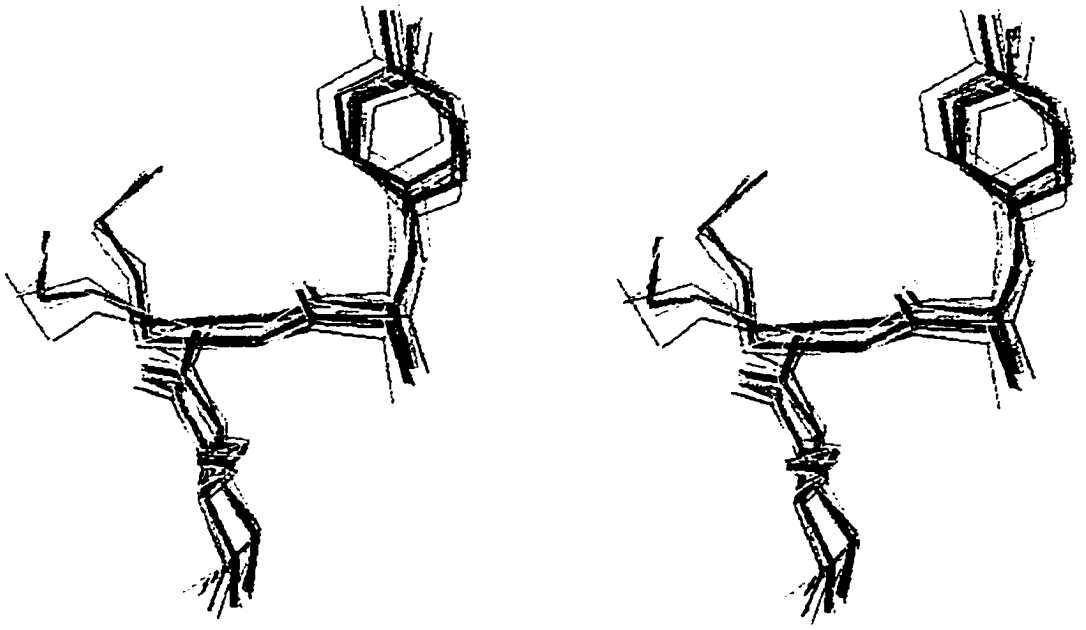


Figure 3.17: Dynamics Conformations for Both Leu68Met and Met64Leu/Leu68Met

Conformations of the three residues Tyr67-Met68-Thr69 in both the single mutant Leu68Met and the double mutant Met64Leu/Leu68Met are shown. The figure is constructed identically to Fig. 3.9.

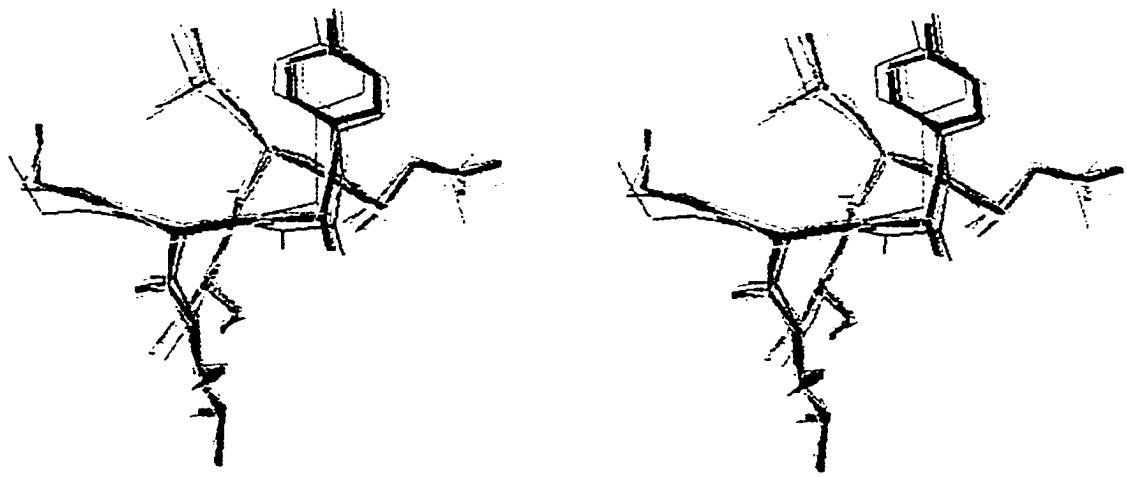


Figure 3.18: Dynamics Conformations for Met64Leu/Leu68Met

Conformations of the six residues Asn63-Leu64-Ser65 and Tyr67-Met68-Thr69 in the double mutant Met64Leu/Leu68Met are shown. The two residues appear on the same side of an α–helix approximately one turn apart. The figure is constructed identically to Fig. 3.9.

3.3.2.6 Ile75Met

Again, as in most of the other simulations, no significant backbone torsion angles were changed. The side chain dihedral χ_3 shifts from 90 to -160 after 4ps; despite this movement no significant secondary movements are observed for the interior of the protein. The sulphur atom is found near the internally bound water molecule Wat 166, which may have potential ramifications on function.

A comparison between the average structure and the X-ray crystal structure for Ile75Met is shown in Figure 3.20. The side chain conformations are noticeably different, mostly attributable to the difference in χ_1 which is -70° in the model (following that of Ile75 in the wild type) and 53° in the crystal structure. The sulphur hydrogen bonds to Wat 166 (shown in Fig. 3.20) and results in a rearrangement of the hydrogen bonding pattern for the water with Asn52; this hydrogen bond is not seen in the dynamics simulation owing to the differences in geometry.

It is possible that given a longer run time (>10ps) the methionine would have shifted to the conformation observed in the crystal structure; however the differences between the modelled and determined conformations illustrate the importance of the starting structure on the simulation and the high potential for error in the side chain conformations.

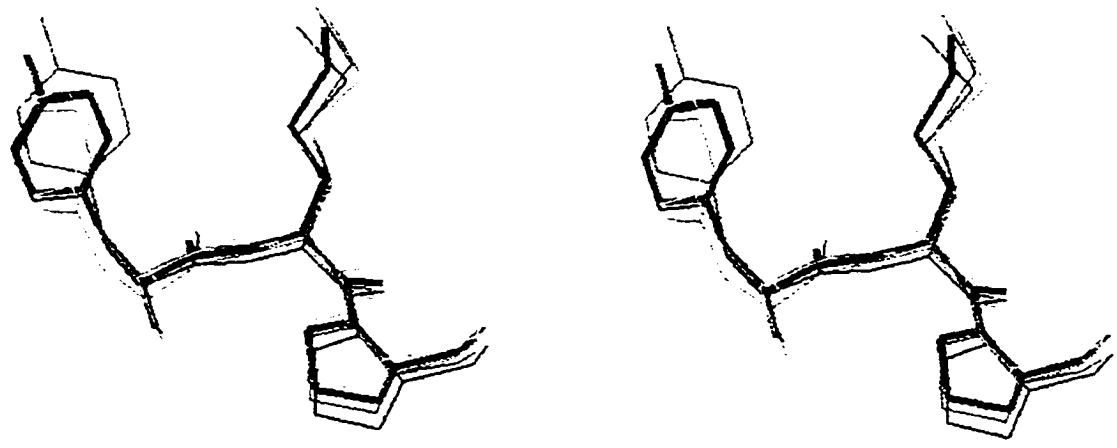


Figure 3.19: Dynamics Conformations for Ile75Met

Conformations of the three residues Tyr74-Met75-Pro76 in the simulation for Ile75Met are shown. The figure is constructed identically to Fig. 3.9.

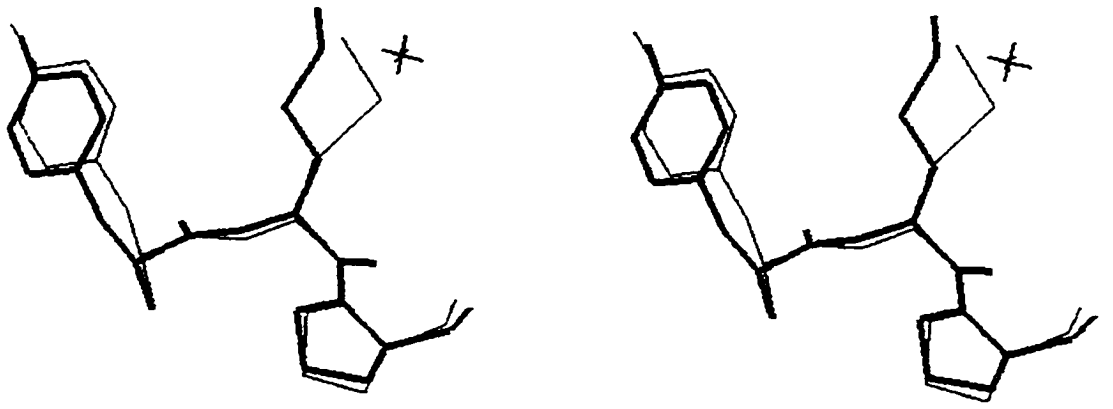


Figure 3.20: Comparison of Ile75Met Model to Ile75Met Crystal Structure

The average structure from Figure 3.19 (thick lines) is superimposed on the same three residues in the crystal structure for Ile75Met (thin lines). The star is the location of the internal water molecule Wat166. In the crystal structure it forms a hydrogen bond to the sulphur of Met75. The Ile75Met crystal structure PDB code is 1IRV (Rafferty et al.)

Residue Position/Mutant	Backbone Dihedral Angles (ϕ,ψ)	
	Mutant Simulation	Wild Type Simulation
Pro25Met	-83°, 146°	-75°, 141°
Val28Met	-60°, -61°	-84°, -60°
Ile35Met/Leu64Met 0-4ps 4-10ps	 -75°, -52° -140°, -52°	-78°, -50°
Lys55Met	-59°, -50°	-64°, -57°
Leu68Met	-73°, -34°	-71°, -15°
Met64Leu/Leu68Met	-65°, -23°	-71°, -15°
Ile75Met	-114°, 90°	-116°, 97°

Table 3.7: Backbone Dihedrals for Dynamics Simulations

Average values for the dihedral values ϕ and ψ are listed for the wild type simulations and the mutant simulations for the altered residues. Averages range from 0 to 10ps unless indicated.

Residue/Mutant	χ_1	χ_2	χ_3
Pro25Met	65	172	139
Val28Met	-171	176	-90 (0-2ps, 4-6ps) 180 (2-4ps, 6-10ps)
Ile35Met/Met64Leu	180	180	-90 (0-8ps) 180 (8-10ps)
Lys55Met	-170	67	180 (0-4ps) 90 (4-10ps)
Leu68Met	-76	-65	111
Met64Leu/Leu68Met	-146	-175	-101
Ile75Met	-70	-173	80 (0-4ps) -165 (4-10ps)
Ile75Met (Cryst.)	53	149	72

Table 3.8: Sidechain Dihedrals for Dynamics Simulations

Average values for the dihedral values χ_1 to χ_3 are listed for the inserted methionine sidechains. Averages range from 0 to 10ps unless indicated. The angle $\chi_1 = 0$ when Cγ is cis to the backbone N. Ile75Met (Cryst.) lists the dihedral values found in the crystal structure.

Mutant	$RMS_{BB\text{-}total}$	$RMS_{CA\text{-}total}$	$RMS_{BB\text{-}21}$	$RMS_{BB\text{-}21\text{-}Std}$	$RMS_{BB\text{-}10Å}$	$RMS_{BB\text{-}10Å\text{-}Std}$
Pro25Met	0.767 1.066	0.757 0.924	0.530 0.264	0.545	0.471 0.471	0.497
Val28Met	0.844 1.050	0.757 1.072	0.834 0.790	0.586	0.681 0.851	0.829
Ile35Met/ Met64Leu	0.887 0.903	0.880 0.881	0.751 0.758	0.515	0.815 0.913	0.674
Lys55Met	0.963 1.087	0.961 1.082	0.903 0.839	0.734	1.126 1.058	0.847
Leu68Met	0.854 0.970	0.839 0.961	0.721 0.405	0.695	0.707 0.732	0.778
Leu68Met/ Met64Leu	0.789 0.863	0.776 0.828	0.740 0.561	0.695	0.662 0.738	0775
Ile75Met	0.880 0.849	0.844 0.856	0.646 0.662	0.597	0.793 0.724	0.815
Standard	0.938	0.924				

Table 3.9: Superposition Values for Dynamics Simulations

Superposition values listed are derived from the average structures from 2ps to 10ps. Two values are listed, compared to the two standard runs which have no modified residues. $RMS_{BB\text{-}total}$ and $RMS_{CA\text{-}total}$ refer to the entire protein backbone and α-carbon trace respectively. The standards for these values (RMS deviations between controls) are identical in all cases, and are listed at the bottom of the columns. $RMS_{BB\text{-}21}$ refers to backbone superposition values for the mutant to the two standards for the 21 amino acids in the linear sequence surrounding the inserted methionine (i.e. residues 15-35 for Pro25Met) while $RMS_{BB\text{-}21\text{-}Std}$ is the RMS deviation in the controls. Similarly, $RMS_{BB\text{-}10Å}$ and $RMS_{BB\text{-}10Å\text{-}Std}$ list deviations for all residues found within a 10Å radius of the inserted methionines and the associated standards.

CHAPTER 4 : DISCUSSION

In assessing the effects of mutation on the different aspects of structure and function, we are faced with the task of trying to separate the different contributions of each amino acid to the whole. While clearly structural and functional roles can not be completely isolated from each other, the effects of either one or the other may be more significant, so it may still be instructive to consider each one separately.

4.1 STRUCTURAL EFFECTS OF MUTATIONS

For the amino acids at positions 13 and 90 in *S. cerevisiae* iso-1 cytochrome *c*, it is apparent that Arg13 has a more significant role in function than in structure, while Asp90 appears to have a more structural role; this is in agreement with previous work on the same mutants (Huang et al., 1994). For Arg13 the physical assays in Tables 3.1 and 3.2 show clearly that the electrostatics have been significantly altered, even in the conservative substitution Arg13Lys, while the structural integrity of the heme crevice is maintained. This directly contradicts one of the proposed functions for this residue, that of forming a stabilising salt-bridge to Asp90; an aspect which is also not supported by the crystal structure data (Osheroff et al, 1980; Berghuis and Brayer, 1992). The absence of a

significant change in the reduction potential with either Arg13Lys or Arg13Cys also suggests that one of the proposed roles, that of being involved in oxidation state linked conformational changes, may also not be tenable (Berghuis and Brayer, 1992). Any significant contribution to the conformational differences between the oxidised or reduced forms would be expected to manifest itself as a change in the reduction potential, suggesting that the side-chain rearrangements observed for this surface residue in the crystal structures are related to other factors not associated with the oxidation state.

The residue Asp90, in contrast, does appear to be critical for maintenance of the heme crevice, although the manner in which it does this is not clear. The significant drop in auto-oxidation half-time for Asp90Cys might also be expected to be reflected in a drop in the pK_{695} transition value; however, as mentioned in section 3.1.1.1, it may be possible that the titration restores the negative charge by deprotonating the thiol group. The absence of any significant changes in the electrostatic properties of this mutant as measured by the HPLC and ATP affinity chromatographies was surprising, and suggests that the overall changes in electrostatic field are either slight or undetectable by these techniques, despite the sensitivity of the HPLC to the conservative replacement of the spatially proximal Arg13 with lysine or cysteine; it is even conceivable that the thiol group could be deprotonated at pH 7.0 owing to the number of basic residues in its vicinity. The aspartic acid does not appear to form any specific interactions in the crystal structure except with solvent, and since it does not form an electrostatic interaction with Arg13 some other explanation needs to be found to rationalise its conservation and effect on

stability. Asp90 is surrounded by a number of positively charged residues in *S. cerevisiae* iso-1 cytochrome *c* including Lys5, Arg13, Lys86, Lys87, and Lys89. In the absence of any other reason it may be that the presence of the negative charge at position 90 compensates for potential repulsive effects between these different groups enhances stability, and that HPLC does not show a difference in retention time either because the specific region of cytochrome *c* which interacts with the column is removed from this location or because the thiol group is deprotonated.

In order to fully discuss the structural effects of the mutations at positions 82 and 85, it is necessary to reconcile the physical data obtained here with the structural data which is known from the crystal structure solutions. Successful combination of this information will hopefully yield a coherent model for the contributions of these residues to the whole. It is clear from the reduction potentials that none of these mutants has a significantly altered environment for the heme. The reduction potential of the heme is sensitive to modification of either solvent exposure or the polarity of proximal groups (Louie et al., 1988; Louie and Brayer, 1989), and the lack of reduction potential differences for the two mutants Leu85Cys and Phe82Tyr/Leu85Ala is somewhat surprising since they both introduce newly bound water molecules near the heme (Lo et al., 1995a and 1995b). This maintenance of reduction potential is ascribed to the location of the new water molecules, as they are situated far enough from the porphyrin ring (5.4Å) and in an orientation which will have minimal effects on the heme environment (Lo et al., 1995a). In the double mutant Phe82Tyr/Leu85Ala and in the single mutant

Leu85Ala the hydroxyl group is also remote enough to not significantly change the heme environment polarity or have a large impact on the reduction potential. For the other three mutants (Leu85Ala, Leu85Phe, Leu85Met) there is no significant change in the heme environment or in apparent solvent access to the heme crevice, so it is not surprising that the reduction potential remains unchanged in these cases (Lo et al., 1995a and 1995b).

Despite the similarities in heme reduction potentials, there is clearly a difference in structural integrity of the heme crevice as determined by the alkaline transition values. Again, this does not appear to reflect direct solvent access changes, as Leu85Cys (which introduces a water molecule, see Table 1.4) shows no significant depression of the transition midpoint, while Leu85Ala (which does not introduce water) does. Another interesting factor is that the salt bridge between Arg13 and Asp90, mentioned previously for its proposed stabilising role, is formed in the Phe82Tyr/Leu85Ala and Leu85Ala mutant crystal structures (Lo et al., 1995a) but yet these two have the lowest alkaline transition midpoints; thus, like the data for the mutations at positions 13 and 90, this suggests that the interaction between these two residues is not stabilising, certainly not in this titration. The drastic decrease in both mutants with Ala85 present suggests that there may be an aspect specific to the size of residue 85 which is responsible for the integrity of the heme crevice. The striking difference in alkaline transition midpoints between Leu85Ala and Leu85Cys shows the difference that a single sulphur atom makes in this position. While it might be the presence of a heavy (non-hydrogen) atom in the γ-position

that is sufficient for the difference in stability, it may also be that the sulphur atom is involved in stabilising the Leu85Cys structure through a favourable electronic interaction with the edge of the Phe82 ring (Lo et al., 1995b; Reid et al., 1985). The midpoint for the Phe82Tyr/Leu85Ala alkaline transition is not significantly higher than that of Leu85Ala; their similarity, however, indicates that the destabilisations apparent in the Phe82Tyr and Leu85Ala mutants are not additive in the double mutant. This is likely because the movement of the tyrosine residue towards the surface observed in the Phe82Tyr mutant structure, is no longer present in the double mutant, as the hydroxyl group can occupy the space opened up by the loss of the large side chain at position 85 (Lo et al., 1995a). The more moderate drop in the Phe82Tyr transition pH is possibly due to the tyrosine moving to a more solvent exposed position, which may be part of the transition rearrangement, or similarly it may be due to the difference between solvation of phenylalanine or tyrosine in the alkaline form. Neither of these factors are apparent in the double mutant, since the side chain is not shifted towards the surface and the hydroxyl group will already be solvated, this may be why the tyrosine does not contribute to further transition midpoint depression in the double mutant. The auto-oxidation times are not significantly different in any of the six mutants showing that, like the reduction potentials, there is no apparent increase in the solvent accessibility of the heme crevice occurring at pH 7.0.

The electrostatic aspects of these six mutants, as measured by the cation exchange and ATP affinity chromatography, indicate that there were no significant modifications of charge distribution. The crystal structures show distinct rearrangements of Arg13 in the

three mutants Phe82Tyr, Leu85Ala, and Phe82Tyr/Leu85Ala (Lo et al., 1995a); however, these do not lead to differences which are detectable by these chromatographic methods. This neither supports nor disproves the hypothesis that one of the roles residue 85 fulfils is the prevention of the Arg13 - Asp90 interaction (Lo et al., 1995a) but it does show that the mutations have no detectable electrostatic ramifications.

Taken together, these results indicate a strong structural role for Leu85. In terms of stability, size appears to be more of a factor than hydrophobicity, since the introduction of the relatively polar cysteine (with an associated water molecule) has no effect on the pK_{695} transition while the Leu85Ala mutant, which maintains the hydrophobic character of this region, shows the most marked decrease. This gives us our first indicator of the conservation of residue 85 as a large hydrophobic residue.

4.2 Biological Effects of Mutations

As mentioned previously, variants of Arg13 show more functional sensitivity in the biological assays than Asp90. This is exemplified by the conservative substitution Arg13Lys having as significant an impact on both the succinate oxidase and cytochrome *c* oxidase assays as the non-conservative substitution Asp90Cys. While it is unlikely that the reasons for this similarity are functionally related, it does suggest a relatively important role for Arg13 in physiological partner reactions; this is indicated further by the even lower activity for the Arg13Cys variant.

The behaviour of Arg13Lys can be rationalised in terms of some of the structural data. While the stability or heme environment of the mutant is unaltered, chromatography results showed that it does have a significant change in surface charge distribution. Arginine and lysine are non-identical residues in both charge distribution, as arginine has a delocalized positive charge in the guanido group while lysine's positive charge is localised to the amino group, and in length, as arginine's side chain is one atom longer than that of lysine. Since residue 13 is proposed to be part of the interface for interaction with both cytochrome *c* oxidase and the reductase complexes (Pettigrew and Moore, 1987) it is reasonable to suppose that slight modifications of either side chain conformations or electrostatic character would have repercussions in binding geometries of protein complexes which could lead to the slight decrease in electron transport efficiency.

Such an electrostatic effect was not noted in the Asp90Cys chromatography results, which may also help to explain its only modest decrease in electron transport activity. The column used for cation exchange chromatography will, like the cytochrome oxidase and reductase binding sites, present a negatively charged interaction surface. If the removal of a negative charge does not lead to enhanced binding (longer retention times) with the column, then this provides a precedent to believe that the interaction with physiological partners might also be unaffected. Cytochrome *c*'s interactive face is fairly consistent between physiological partners, and lies on the face of cytochrome *c* on the other side of Arg13 from Asp90; in addition to Arg13 (Lys13 in horse), it includes a number of basic residues including Lys72, Lys86 and Lys87. The presence of these

various charged groups may minimise the effects of the negative charge change in Asp90Cys on the intermolecular interactions. Alternately, it may be that the specific intermolecular associations (i.e. salt-bridges formed between protein partners) do not include Asp90, and in this sense the intermolecular association would not be affected by modification of this residue. Unfortunately, it is not possible with the current data to distinguish between these two possibilities.

The behaviour of Arg13Cys again indicates that Arg13 is an important modulator of electron transport activity. This substitution shows a more dramatic decrease in activity in both the succinate oxidase and the cytochrome oxidase assays than Arg13Lys which is reasonable considering the non-conservative nature of the replacement. The double mutant Arg13Cys/Asp90Cys shows a more than additive decrease in the activities, suggesting a synergistic effect for the two substitutions. This is particularly interesting given the only moderate effects of the single mutant Asp90Cys; unfortunately, no other comparisons can be drawn for this mutant owing to the lack of physical data, particularly the chromatographic data.

For the mutations at positions 82 and 85, the biological assays present more revealing information about the role of these two residues in cytochrome *c* function. Again using the structural data as a context for the analysis, the succinate oxidase assays (which are proposed to reflect cytochrome *c* / cytochrome reductase electron transfer rates (Wallace, 1984)) indicate that those substitutions which maintain the hydrophobicity

immediately around residue 85 (Leu85Ala, Leu85Phe, Phe82Tyr, Leu85Met) function as efficiently as the wild type. In contrast, the two substitutions which introduce water molecules to this region (Leu85Cys, Phe82Tyr/Leu85Ala) show the lowest electron transport activity. Since two of the hydrophobic substitutions are natural variants (either Met or Phe at position 85), this is not unexpected for these mutants. Residue 85 is close to (if not part of) the interaction site on cytochrome *c* for binding to various physiological partners (Pettigrew and Moore, 1987; Kornblatt et al., 1992; Guillemette et al., 1994), so the presence of water might have two different modes by which it could affect the electron transfer complex. First, the complex may be able to form with the water present, although with a geometry which differs enough from the native complex to affect electron transfer rates; secondly, the water may have to be excluded in order to form the complex, in which case the limiting step might be the formation of the electron transfer complex and not the electron transfer itself. Since in the succinate oxidase assay the cytochrome *c* must dissociate from the reductase and then reassociate with the oxidase; either of these might be feasible explanations for the results.

With two exceptions, the cytochrome oxidase assay results are largely comparable with the succinate oxidase assay. For each mutant, all of the values for electron transfer relative to wild type are respectively lower in this assay, with the exception of Leu85Cys which remains at essentially the same level and the double mutant Phe82Tyr/Leu85Ala which shows a marked increase. This suggests that the electron transfer complex between cytochrome *c* and cytochrome *c* oxidase is more sensitive to changes in these residues,

which is consistent with the oxidase and the reductase binding in similar but non-identical fashions (Pettigrew and Moore, 1987). In this case, the difference between Leu85Cys and Phe82Tyr/Leu85Ala activities may indicate a difference in how the mutations affect electron transport.

One of the largest differences between the succinate oxidase assay and the particular cytochrome *c* oxidase assay used is that in the first cytochrome *c* must dissociate and reassociate with two different electron transfer partners, while in the second cytochrome *c* remains bound to the oxidase and is reduced by TMPD without having to dissociate (Pettigrew and Moore, 1987). The normal electron transfer rate for the double mutant in the oxidase assay suggests that the intermolecular complex formed is as efficient as the wild type complex. This would also suggest a number of aspects concerning the intermolecular complex formed. The difference between Leu85Ala and Phe82Tyr/Leu85Ala rates in conjunction with the pK_{695} data suggests that the complex formation may stabilise the double mutant but not Leu85Ala. One possible explanation concerns Arg13, which in Leu85Ala moves to fill the space left by the leucine to alanine change in order to maintain the hydrophobic character of the space which is created. In the cytochrome *c* / cytochrome b_5 complex the potential effect of Arg13 movement increasing the solvent exposure of residue 85 has been suggested, and a similar situation might occur with other electron transport partners (Guillemette et al., 1994). If Arg13 is involved in an intermolecular contact with the oxidase, then it may not be able to fulfil both roles at the same time, or at least not in the standard geometry in Leu85Ala. In the

double mutant, it may only be required for the intermolecular contact and not to maintain the hydrophobicity of the cavity normally filled by Leu85, thus the complex formed might not have to be distorted. Similarly, the strikingly high electron transfer rate for the double mutant may reflect the point already raised, that for this protein the rate limiting step may be the formation of the complex and not the transfer itself. The low activity for the Leu85Cys cytochrome *c* may reflect the bound water, or it may also result from the electronic interaction of the sulphur with Phe82 (Lo et al., 1995b). If Phe82 forms part of the electron transfer route, then any changes to the distribution of the π-electron cloud may result in diminished efficiency of transfer. Certainly the results suggest that the two mutants which introduce water do not necessarily modulate electron transfer activity in an identical fashion to each other, and that other factors may contribute to the variability of biological activity.

4.3 ATP AND CYTOCHROME C

The ATP binding site provides a satisfactory model of the binding of ATP and horse ferricytochrome *c* for a number of reasons. First, the involvement of Arg91 in the binding of the γ-phosphate is well characterised (Corthésy and Wallace, 1986) and also the derivatisation of the lysine residues 72, 86, and 87 provides a reasonable location for the adenine moiety (McIntosh et al., 1995). Since these were used in the initial docking of the nucleotide, however, other reasons need to be found which further support this location.

Species	69	70	71	83	84	85	86	87	88	91	Time (min)
Horse	E	N	P	A	G	I	K	K	K	R	24
Rabbit	E	N	P	A	G	I	K	K	K	R	30
Cow	E	N	P	A	G	I	K	K	K	R	12.5
Dog	E	N	P	A	G	I	K	K	T	R	22.5
Tuna	E	N	P	A	G	I	K	K	K	R	40
Chicken	E	N	P	A	G	I	K	K	K	R	16.5
Pigeon	E	N	P	A	G	I	K	K	K	R	31
S. cerevisiae	T	N	P	G	G	L	K	K	E	R	564
C. krusei	E	N	P	A	G	L	K	K	A	R	42.5
Human	E	N	P	V	G	I	K	K	K	R	29

Table 4.1: ATP Binding Site Comparison

Proximal residues to the binding site for ATP are listed for those species whose retention time on an ATP affinity column has been assayed (Craig, 1993). Retention times are for the elution of oxidised cytochrome *c* from an AGATP column by isocratic elution with 35mM potassium phosphate pH 7.0. The *S. cerevisiae* cytochrome listed is iso-1 cytochrome *c*.

Since the arginine at position 91 is invariant across all known cytochrome *c* sequences, it is reasonable to expect that other residues involved in the binding would also be conserved. Table 4.1 lists those residues found proximal to the ATP (Tables 3.5 and 3.6) for a number of different mitochondrial cytochromes *c* along with their respective elution times from an ATP affinity column (Craig, 1993). It is apparent that the general composition of this binding site is maintained for these examples, with the apparent exception of the *S. cerevisiae* iso-1 sequence. Two distinctions in particular are Glu88 and Thr69, both of which show very specific interactions with the triphosphate and ribose groups respectively in this model. It should still be possible to form a hydrogen bond between Thr69 and the ribose moiety, and this may even reduce the possible electrostatic repulsions between this residue and the proximal phosphate groups. In addition, *S.cerevisiae* iso-1 cytochrome *c* residues which are not shown in this table include lysines at positions -2, which is found in the N-terminal extension, and at 89, both of these are found nearby. While it is clear that the binding of ATP to *S. cerevisiae* iso-1 cytochrome *c* might not be identical to that to horse, the high retention times suggest it does bind the nucleotide effectively. The precise modulation of the binding of ATP is also not solely dependent on these proximal residues, as is apparent from the differing elution times for tuna and dog cytochromes *c* despite the identical residues in this area.

One other aspect which supports this binding location is that while it would be expected for the *syn* conformation to be the dominant form for 8-N_3-ATP, the relatively high proportion of adducts formed to lysine 86 and 87 suggests that binding of the *anti*

conformation is preferred, which would resemble the natural ATP conformation. This is encouraging because it suggests that the analogue is behaving similarly to ATP in complex with cytochrome *c*, thus strengthening the validity of comparisons drawn between ATP binding and the covalently modified adducts of cytochrome *c*.

In the quest to develop a mechanistic scheme that explains the physiological relevance of this ATP binding site to the regulatory role proposed for cytochrome *c*, the model provides some useful insights. Firstly, the movement of the protein side chains upon binding is not reflected in associated motions of the protein backbone, suggesting that the mechanism does not proceed through internal conformational rearrangement. Also, since the nucleotide falls just outside the proposed binding sites for physiological partners (Pettigrew and Moore, 1987) it is unlikely that it functions through steric or competitive inhibition. The explanation likely lies in the electrostatic differences, either through side chain positions or through manipulation of the dipole moment of the protein as previously suggested (Craig and Wallace, 1991). ATP, in addition to co-ordinating the proximal lysine residues with the triphosphate group, will contribute four negative charges to the complex, significantly altering the charge characteristics of the protein surface. Previous work has shown that modification of cytochrome *c* with 8-N_3-ATP causes both dissociation of the cytochrome from the inner mitochondrial membrane as well as interfering with electron transfer activity (Craig and Wallace, 1995). Both of these modes of ATP regulation of oxidative phosphorylation would be consistent with this model.

4.4 Methionine Mutants of *S. cerevisiae* Iso-1 Cytochrome *c*

In general, the substitution of the various residues for methionine throughout the *S. cerevisiae* iso-1 cytochrome *c* sequence leads to no apparent gross steric conflicts. Additionally, in all of the dynamics simulations the hydrogen bonding patterns remained largely consistent, with the exception of the mutant Lys55Met which lost the hydrogen bond associated with the amine nitrogen group, as discussed in section 3.3.2.4. Some structural fluctuations were observed in all of the simulations in the N-terminal region (residues -5 to 1); however, since this area typically shows a relatively disordered conformation in crystal structures, this was not attributed to any of the methionine insertions (Louie and Brayer, 1990). In order for these modelling studies to be applied rationally, it is necessary to know some of the other qualities that these mutants possess in terms of structure and function in order to derive consistent explanations for their behaviour; these assays have been performed in the work of Woods et al. (1996) and this provides the data for the following comparisons.

In terms of reduction potentials, all of the mutants fall into the range between 268-288mV with the exception of Ile75Met which has a substantially lowered potential of 247 mV. While the dynamics simulation does not show a specific explanation for this, the crystal structure shows a hydrogen bond is formed between the sulphur and the mechanistically relevant Wat 166, as mentioned in section 3.3.2.6. The new hydrogen bonding pattern which develops may stabilise the oxidised form of the protein, explaining

the drop in reduction potential (Rafferty et al., unpublished). Alternatively, the extra length of the methionine may distort the bottom loop, increasing solvent access to the protein interior. The lowered reduction potential also explains this mutant's lowered activity (48%) in the succinate oxidase biological assay, since diminishing the reduction potential leads to concomitant reduction of the thermodynamic driving force for the electron transfer between the reductase and cytochrome *c* (Marcus and Sutin, 1985).

The mutant Leu68Met shows a decrease in the alkaline transition midpoint (pK_{695}=7.6) but the modelling provides no clear explanation here to explain that effect. The invariant Leu68 packs against the heme group close to Met80, it is this methionine that is replaced with another iron ligand during this transition. Since Met80 must be displaced during the transition, any structural alterations proximal to it may influence this process; however, nothing substantial was observed during the dynamics. It may also be that the contribution of this residue to the alkaline transition is not manifested in the native structure, but in the rearrangement mechanism or in the alkaline form; however, such a proposal cannot be supported nor disproved by this work.

The conformational perturbance observed in the Lys55Met simulation is apparently reflected in the structural studies as well. For this mutant, the pK_{695} is diminished to 7.8 indicating a substantial decrease in heme crevice stability which could potentially be explained by the loss of the interaction between Lys55 and Tyr74. Although this residue is a methionine in some cytochrome *c* sequences, particularly in rice and other plants, the

crystal structure of the rice form also has an additional hydrogen bond from a threonine at position 63 which yeast lacks and which is predicted to add stability (Louie and Brayer, 1990). This is a good example of the fact that even though the replacement was not outside the natural variation of amino acids for this position, the context is still important in considerations of structure and function.

For Val28Met, the modelling provides no direct clues which might apply to the protein's characteristics, however it may provide some explanation. Val28Met shows decreased activity in both succinate oxidase and cytochrome oxidase assays (63% and 85% respectively) with no changes in other characteristics. It could be that there is a steric explanation for this effect, as the methionine residue projects above the heme edge in the electron port where it could interfere with the interaction with physiological partners. Additionally, the sulphur group occupies a position in the heme plane approximately 5.5 Å from the nearest heme-aromatic atoms, which places it in the optimal sulphur-aromatic interaction position (Reid et al., 1985). This association may have repercussions in terms of electron transport, although the lack of significant change in reduction potential argues that the electron cloud distribution of the heme is not notably perturbed.

For the last two mutants Pro25Met and Ile35Met no significant changes in activity or structural features were noted. Pro25 is a surface residue and it is apparent that the backbone restrictions for proline are not a conformational imperative, both from this work

and the variability of the amino acid residue. Ile75Met, although it is buried, is also apparently able to accommodate the methionine without significant perturbation of the fundamental properties of cytochrome *c*. This is in contrast to Ile75Met which, while being a similar replacement leads to the hydrogen bond rearrangement and reduction potential drop mentioned earlier.

All in all, the insertions of the various methionines indicates that replacements in the numerous locations are tolerated reasonably well. The largest caveat for this work is that the replacement of amino acids should involve non-critical residues for semisynthetic work, and that context is of primary importance even for evolutionarily variable locations.

CHAPTER 5 : CONCLUSIONS

5.1 PATTERNS OF CONSERVATION AND FUNCTION

It is clear for all of the aspects which have been addressed in this thesis that the conservation of specific residues in a protein sequence is necessary when that residue performs a particular function, and that the nature of that conservation can be governed by a number of factors. The concept of a conservative substitution is context dependent, because the very nature of conservation itself is a fluid and ambiguous one. If we imagine amino acids as existing in subsets of similarity, based on aspects of size, polarity, or other qualities, then even for the twenty residues with their relatively limited physical properties, we end up with the need for fairly extensive descriptions of their relationships to each other (Taylor, 1986). In those cases where a single amino acid needs to fulfil more than one function it will be the most conserved; on the other hand, where the maintenance of only a single function is necessary larger degrees of variation are possible. For the specific residues dealt with in this thesis, we can now attempt to rationalise a coherent picture for the whole protein in terms of three basic concepts: structural, electrostatic and functional integrity; although these are not necessarily distinct aspects, as we shall see.

5.1.1 Structural Integrity

A number of the mutants examined in this study lead to alterations in structural integrity, as measured by the appropriate tests. To begin with, the acidic residue Asp90 in *S. cerevisiae* iso-1 cytochrome *c* has a largely structural role. While the manner in which Asp90 fulfils this role is not clear from this work, the conservation of this residue as primarily aspartic acid or glutamic acid suggests that this effect is mediated through the acidic nature of the residue, yet replacement with cysteine does not significantly affect chromatographic properties sensitive to electrostatic alterations. This implies that the sphere of influence is highly localised, but when we combine this with the knowledge that it does not form a specific electrostatic interaction with Arg13 it becomes necessary to rationalise its conservation in another way. It may be that its electrostatic effects are distributed among some or all of the nearby charged residues from both ends of the polypeptide chain, including lysines 5, 86,87, and 89 as well as Arg13, allowing the two ends of the polypeptide chain to maintain their close proximity in space while minimising electrostatic repulsion. While more work would need to be done to confirm this, there is no clear explanation yet as to why modification of this residue leads to heme crevice destabilisation.

The other residue which has a clear structural role in *S. cerevisiae* iso-1 cytochrome *c* based on this work is Leu85. Modification of this residue to alanine in particular leads to a dramatic decrease in the alkaline transition, suggesting that one of the

reasons it is maintained as a large residue is a matter of structural integrity. This is not a question of hydrophobicity, but may be more related to the presence of a heavy atom attached to the β-carbon, since the substitution of cysteine here, with its associated polarity and a newly introduced water molecule, has little or no impact on heme crevice stability.

5.1.2 Electrostatic Integrity

It is clear from other work that cytochrome *c* does rely heavily on its electrostatic qualities (Osheroff et al., 1980; Ferguson-Miller et al., 1978) and it is clear that Arg13 is an integral part of this reliance. Even the substitution of a lysine here is not absolutely conservative, and differences are noted in both chromatographic and biological aspects. For this residue the non-conservative substitution of a cysteine for the arginine does lead to significant impairment of function and much larger electrostatic rearrangements, far more significant than the similar non-conservative replacement of aspartic acid for cysteine.

Another aspect of electrostatic conservation concerns the ATP binding site, where the introduction of a different charge arrangement into an electrostatically defined binding site forms the basis for a potential regulatory role. The elucidation of the role Arg91 plays in ATP binding is a good example of needing to apply an appropriate test to resolve the contribution of an invariant residue in an associated aspect of cytochrome *c* function, since

specific chemical modification (Wallace and Rose, 1983) or substitution of this amino acid with norleucine yielded no change in the basic character of the protein in either structure or electron transfer (C.J.A. Wallace, unpublished data). Clearly in this case the concept of conservation was dependent on the situation in which the protein was examined, and reveals the necessity of examining all facets of a protein in order to develop a coherent and holistic model.

5.1.3 Functional integrity

Residue 13 has an important and fundamental role in cytochrome *c* function revealed by both previous work and this study. It forms an integral part of the intermolecular complex (Pettigrew and Moore, 1987; Kornblatt et al., 1992; Huang et al., 1994), and modification affects both electrostatics and electron transfer rate with physiological partners significantly, even mutation to the chemically similar amino acid lysine. This is an example of the intrinsic trouble in trying to separate different aspects of amino acid structure and function, since clearly we are dealing with intimately connected properties of electrostatics and function for this residue.

Residue 85, previously mentioned in terms of structural integrity based on its size, shows a different aspect in terms of functional integrity. In this case the hydrophobicity appears to be a more significant modulator of activity, and this may provide the final piece in the puzzle of residue 85 's conservation as a bulky hydrophobic group: to maintain both

the aspects of structural stability (size) and activity (hydrophobicity). Only the large hydrophobic residues Ile, Leu, Phe, and Met fill both of these roles adequately.

5.2 Future Directions

Further exploration of the residues at position 85 would help define the most important aspects of this residue's contribution to the structure. With the exceptions of glycine and alanine, the only other natural amino acid smaller than cysteine is serine. Two potentially useful mutations would be to either serine or threonine, since both would introduce another small polar residue into this region, and the differences between these substitutions and those of Leu85Ala and Leu85Cys might help differentiate between polar, small side chain, and branched β-carbon aspects. Further study of the residues at positions 13 and 90 could also yield more information on their specific role, particularly residue 90. Introduction of either an uncharged (i.e. alanine or valine) and/or polar residue (i.e. serine, threonine or asparagine) here might more clearly define what its structural role is.

With regards to the ATP binding, the differences suggested for the *S. cerevisiae* iso-1 cytochrome *c* binding site(s) are worth investigating. The extremely high retention times in ATP affinity columns relative to the other species are intriguing, as are the sequence differences in the proximal residues. Compared to horse cytochrome *c*, little information is known about specific binding affinities, the number of binding sites, or other aspects which might be useful in defining the parameters for this association. The

advantage of pursuing this work in yeast is the availability of site-directed mutagenesis technology which would be an extremely powerful tool.

The methionine scan residues also present a number of opportunities to be followed up. One aspect involves the Lys55 contribution to structure and function, and whether or not it is electrostatic or structural (or both). Some work has already been undertaken to address this question through alkylation of the methionine residues, the result of this modification is that the sulphurs become positively charged through being bonded to three atoms and a new alkyl group is added to the side chain ($-CH_2C(O)NH_2$). For Lys55Met this potentially restores the positive charge without the restoration of the appropriate hydrogen bond, allowing a potential differentiation between electrostatics and hydrogen bond contributions. Also, there are intriguing results coming from the same reaction for Met35, another one of the buried methionine scan residues. It appears as though after religation (formation of homoserine at position 35) activity drops sharply in the succinate oxidase assay. Surprisingly, alkylation of this methionine (which buries both a large polar group and a positive charge) results in unaffected activity. Preliminary analysis suggests that this may involve the hydrophobic pocket at the back of the heme and a hydrophilic pocket located near the propionate groups of the heme and Arg38. While homoserine may form a hydrophilic passage connecting the two pockets, the acyl group may be able to be accommodated in the interior of the protein maintaining aspects of both pockets. More work needs to be done, however, to completely characterise this phenomenon.

References

Becker, D.M. and L. Guarente (1992) "Protocol for High-Efficiency Yeast Transformation" in Guide to Electroporation and Electrofusion (Academic Press, New York)

Berghuis, A.M. and G. Brayer (1992) "Oxidation State-dependent Conformational Changes in Cytochrome *c*" J. Mol. Biol. **223**:959-976

Berghuis, A.M., J.G.Guillemette, G. McLendon, F. Sherman, M. Smith and G.D. Brayer (1994) "The Role of a Conserved Internal Water Molecule and its Associated Hydrogen Bond Network in Cytochrome *c*" J. Mol. Biol. **236**:786-794

Broach, J.R., J.N. Strathern and J.B. Hicks (1979) "Transformation in Yeast: Development of a Hybrid Cloning Vector and Isolation of the CAN1 Gene" Gene **8**:121-133

Burrows, A.L., L-H. Guo, H.A.Hill, G. McLendon and F. Sherman (1991) "Direct Electrochemistry of Proteins" Eur. J. Biochem. **202**:543-549

Bushnell, G.W., G.V. Louie, G.D. Brayer (1990) "High Resolution Structure of Horse Heart Cytochrome *c*" J. Mol. Biol., **214**:585-595

Corradin, G. and H.A. Harbury (1974) "Reconstitution of Hrose-Heart Cytochrome *c*. Reformation of the Peptide Bond Linking Residues 65 and 66" Biophy. Res. Comm. **61**:4100-4106

Corthésy, B.E. and C.J.A. Wallace (1986) "The Oxidation-state-dependent ATP-binding Site of Cytochrome *c*. A Possible Physiological Significance" Biochem J **236**:359-364

Corthésy, B.E. and C.J.A. Wallace (1988) "The Oxidation-state-dependent ATP-binding Site of Cytochrome *c*. Implication of an Essential Arginine Residue and the Effect of Occupancy on the Oxidation-reduction Potential" Biochem J. **252**:349-355

Craig, D.B. and C.J.A. Wallace (1991) "The Specificity and K_d at Physiological Ionic Strength of an ATP-Binding Site on Cytochrome *c* Suit it to a Regulatory Role" Biochem. J. **279**:781-786

Craig, D.B. (1993) "A Regulatory Role for ATP Binding to Cytochrome *c*" PhD Thesis, Dalhousie University, Halifax, N.S.

Craig, D.B. and C.J.A. Wallace (1993) "ATP Binding to Cytochrome *c* Diminishes Electron Flow in the Mitochoindrial Respiratory Pathway" Prot. Sci. **2**:966-976

Craig, D.B. and C.J.A.Wallace (1995) "Studies of 8-Azido-ATP Adducts Reveal Two Mechanisms by which ATP Binding to Cytochrome *c* could Inhibit Respiration" Biochemistry **34**:2686-2693

Creighton, T.E. (1993) Proteins: Structure and Molecular Principles (Freeman and Co., New York)

Crick, F. (1958) "On Protein Synthesis" Symposia of the Society for Experimental Biology, **12**:138-163

Cutler, R.C., G.J. Pielak, A.G. Mauk and M. Smith (1987) "Replacement of Cysteine 107 of *Saccharomyces cerevisiae* iso-1-cytochrome *c* with Threonine: Improved Stability of the Mutant Protein" Prot. Eng. **1**:95-99

Darley-Usmar, V.M., R.A. Capaldi, S. Takamiya, F. Millett, M.T. Wilson, F. Malatesta and P. Sarti (1987) "Reconstitution and Molecular Analysis of the Respiratory Chain" in Mitochondria: A Practical Approach (IRL Press, Oxford)

Dente, L., G. Cesareni, and R. Cortese (1983) "pEMBL: a New Family of Single Stranded Plasmids" Nuc. Acid Res. **11**:1645-1655

Dickerson, R.E. (1972) "The Structure and History of an Ancient Protein" Science **226**:189-202

Ferguson-Miller, S., D.L. Brautigan, and E. Margoliash (1978) "Definition of Cytochrome c Binding Domains by Chemical Modification" J. Biol. Chem. **253**:149-159

Finkelstein, A.V., A.M. Guton and A.Y. Badretinov (1993) "Why are the Same Protein Folds Used to Perform Different Functions" FEBS Letts., **325**:23-28

Guillemette, J.G., P.D. Barker, L.D. Eltis, T.P. Lo, M. Smith, G.D. Brayer, and A.G. Mauk (1994) "Analysis of the Bimolecular reaction of Ferricytochrome *c* by Ferrocytochrome b_5 through Mutagenesis and Molecular Modelling" Biochimie **76**:592-604

Gunner, M.R. and B. Honig (1991) "Electrostatic Control of Midpoint Potentials in the Cytochrome Subunit of the *Rhodopseudomonas viridis* Reaction Center" Proc. Natl. Acad. Sci. USA **88**:9151-9155

Hawkins, B.K., S. Hilgen-Willis, G. Pielak and J.H. Dawson (1994) "Novel Axial Ligand Interchange in Cytochrome *c*: Incorporation of a Histidine at Position 82 Leads to Displacement of Wild-Type Methionine-80 Ligand" J. Am. Chem. Soc. **116**:3111-3112

Hazzard, J.T., G. McLendon, M.A. Cusanovich, G. Das, F. Sherman and G. Tollin (1988) "Effects of Amino Acid Replacements in Yeast Iso-1-cytochrome *c* on Heme Accessibility and Intra Complex Electron Transfer in Complexes with Cytochrome *c* Peroxidase" Biochemistry **27**:4445-4451

Huang, Y., S. Beeser, J.G. Guillemette, R. Storms and J. Kornblatt (1994) "Mutation of Iso-1-cytochrome *c* at Positions 13 and 90" Eur. J. Biochem. **223**:155-160

Inglis, S.C., J.G. Guillemette, J.A. Johnson and M. Smith (1991) "Analysis of the Invariant Phe82 Residue of Yeast Iso-1 Cytochrome *c* by Site-Directed Mutagenesis Using a Phagemid Yeast Shuttle Vector" Protein Eng. **4**:569-574

Kadenbach, B. (1986) "Regulation of Respiration and ATP Synthesis in Higher Organisms: Hypothesis" J. Bioenerg. Biomembr. **18**:39-54

Kornblatt, J.A., J. Theodorakis, G. Hui Bon Hoa, and E. Margoliash (1992) "Cytochrome *c* and Cytochrome *c* Oxidase Interactions: the Effects of Ionic Strength and Hydrostatic Pressure Studied with Site-Specific Modifications of Cytochrome *c*" Biochem. Cell. Biol. **70**:539-547

Liang, N., A.G. Mauk, G.J. Pielak, J.A. Johnson, M. Smith and B. Hoffman (1988) "Regulation of Interprotein Electron Transfer by Residue 82 of Yeast Cytochrome *c*" Science **240**:311-313

Lo, T.P., J.G. Guillemette, G.V. Louie, M. Smith and G.D. Brayer (1995a) "Structural Studies of the Roles of Residues 82 and 85 at the Interactive Face of Cytochrome *c*" Biochemistry **34**:163-171

Lo, T.P., M.E.P. Murphy, J.G. Guillemette, M. Smith and G.D. Brayer (1995b) "Replacements in a Conserved Leucine Cluster in the Hydrophobic Heme Pocket of Cytochrome *c*" Protein Science **4**:198-208

Louie, G.V. and G.B. Brayer (1990) "High Resolution Refinement of Yeast Iso-1-cytochrome *c* and Comparison with Other Eucaryotic Cytochromes *c*" J. Mol. Biol. **214**:527-555

Louie, G.V., W.L.B. Hutcheon and G.D. Brayer (1988) "Yeast Iso-1-Cytochrome *c*: A 2.8 Å Resolution Three-Dimensional Structure Determination" J. Mol. Biol., **199**:295-314

Lum, V.R., G.D. Brayer, G.V. Louie, M. Smith, and A.G. Mauk (1987) "Computer Modelling of Yeast Iso-1 Cytochrome *c* - Yeast Cytochrome *c* Peroxidase Complexes" Protein Struct. Fold. and Design **2**:143-150

Lundblad, R.L. and C.M. Noyes (1984) Chemical Reagents for Protein Modification (Volume 1) (CRC Press, Boca Raton)

Marcus, R.A. and N. Sutin (1985) "Electron Transfers in Chemistry and Biology" Biochem. Biophys. Acta **811**:265-322

McIntosh, D.B., J.C. Parrish and C.J.A. Wallace (1996) "Definition of a Nucleotide Binding Site on Cytochrome *c* by Photoaffinity Labelling" J. Biol. Chem **271**:18379-18386

Moore, G.R. and G.W. Pettigrew (1990) Cytochromes *c*: Evolutionary, Structural, and Physicochemical aspects (Springer-Verlag, New York)

Murphy, M.E.P., B.T. Nall and G.D. Brayer (1992) "Structure Determination of Yeast Iso-2 cytochrome *c* and a Composite Mutant Protein" J. Mol. Biol., **227**:160-176

Northrup, S.H., M. Pear, J.D. Morgan, and J.A. McCammon (1981) "Molecular Dynamics of Ferrocytochrome *c*" J. Mol. Biol **153**:1087-1109

Ochi, H., Y. Hata, N. Tanaka and M. Kakudo (1983) "Structure of Rice Ferricytochrome *c* at 2.0 Å Resolution" J. Mol. Biol. **166**:407-418

Oscheroff, N., D. Borden, W.H. Koppenol and E. Margoliash (1980) "Electrostatic Interactions in Cytochrome *c*" J. Biol. Chem. **255**:1689-1697

Pearce, L.L., A.K. Gärtner, M. Smith and A.G. Mauk (1989) "Mutation-Induced Perturbation of the Cytochrome c Alkaline Transition" Biochemistry **28**:3152-3156

Pelletier, H. and J. Kraut (1992) "Crystal Structure of the Complex Between Electron Transfer Partners, Cytochrome *c* Peroxidase and Cytochrome *c*" Science **248**:1748-1755

Pettigrew, G.W. and G.R. Moore (1987) Cytochromes *c*: Biological Aspects (Springer Verlag, New York)

Pielak, G.J., A.G. Mauk and M. Smith (1985) "Site-directed Mutagenesis of Cytochrome *c* Shows that an Invariant Phe is not Essential for Function" Nature **313**:152-154

Rafferty, S.P., J.G. Guillemette, A.M. Berghuis and M. Smith, G.D. Brayer, A.G. Mauk (to be published) "Mechanistic and Structural Contributions of Critical Surface and Internal Residues to Cytcohrome *c* Electron Transport Reactivity" (PDB reference code 1IRV)

Rafferty, S.P., J.G. Guillemette, M. Smith and A.G. Mauk (1996) "Azide Binding and Active Site Dynamics of Position-82 Variants of Cytochrome *c*" Inorg. Chim. Acta **242**:171-177

Rafferty, S.P., Pearce, L.L., Barker, P.D., J.G. Guillemette, C.M. Kay, M. Smith and A.G. Mauk (1990) "Electrochemical, Kinetic, and Circular Dichroic Consequences of Mutations at Position 82 of Yeast Iso-1-cytochrome *c*" Biochemistry **29**:9365-9369

Reid, K.S.C., P.F. Lindley, and J.M. Thornton (1985) "Sulphur-aromatic Interactions in Proteins" FEBS Letts. 190:209-213

Rieder, R., and H.R. Bosshard (1980) "Comparison of the Binding Sites on Cytochrome *c* for Cytochrome Oxidase, Cytochrome bc_1 and Cytochrome c_1." J. Biol. Chem. **255**:4732-4739

Rost, B. and C. Sander (1994) "Combining evolutionary Information and Neural Networks to Predict Protein Secondary Structure" Proteins **20**:216-226

Saenger, W. (1984) Principles of Nucleic Acid Structure (Springer Verlag, New York)

Salemme, F.R., J. Kraut and M.D. Karmen (1973) "Structural Basis for Function in Cytochromes *c*" J. Biol. Chem. **248**:7701-7716

Sano, S. (1972) "Chemical Synthesis of the Cytochrome *c* Molecule" in Structure and Function of Oxidation-Reduction Enzymes (A. Akeson and A. Ehrenberg, eds.) (Pergamon, Oxford) p. 69-83

Schultz, G.E. and R.H. Schirmer (1979) Principles of Protein Structure (Springer-Verlag, New York)

Smith, M. (1985) "In Vitro Mutagenesis" Ann. Rev. Genet. **19**:423-462

Smith, M.B., J. Stonehuerner, A.J. Ahmed, N. Staudenmeyer, and F. Millet (1980) "Use of Specific Triflouroacetylation of Lysine Residues in Cytochrome *c* to study the reaction with Cytochrome b_5, Cytochrome c_1 and Cytochrome *c* Oxidase" Biochim. Biophys. Acta **592**:303-313

Takano, T. and R. Dickerson (1981a) "Conformation Change of Cytochrome *c*: I. Ferrocytochrome *c* Structure Refined at 1.5 Å Resolution" J. Mol. Biol. **153**:79-94

Takano, T. and R. Dickerson (1981b) "Conformation Change of Cytochrome *c*: II. Ferricytochrome *c* Refinement at 1.8 Å and Comparison with the Ferrocytochrome Structure" J. Mol. Biol. **153**:95-115

Tanaka, N., T. Yamane, T. Tsukihara, T. Ashida and M. Kakudo (1975) "The Crystal Structure of Bonito (Katsuo) Ferrocytochrome *c* at 2.3 Å Resolution" J. Biochem., **77**:147-162

Taylor, W.R. (1986) "The Classification of Amino Acid Conservation" J. Theor. Biol. **119**:205-218

Wallace, C.J.A. and K. Rose (1983) "The Semisynthesis of Analogues of Cytochrome c" Biochem J. **215**:651-658

Wallace, C.J.A. (1984) "The Effect of Complete or Specific Partial Acetimidylation on the Biological Properties of Cytochrome *c* and Cytochrome *c*-T" Biochem J. **217**:595-599

Wallace, C.J.A, G. Corradin, F. Marchiori and G. Borin (1986) "Cytochrome *c* Chimarae from Natural and Synthetic Fragments: Significance of the Biological Properties" Biopolymers 25:2121-2132

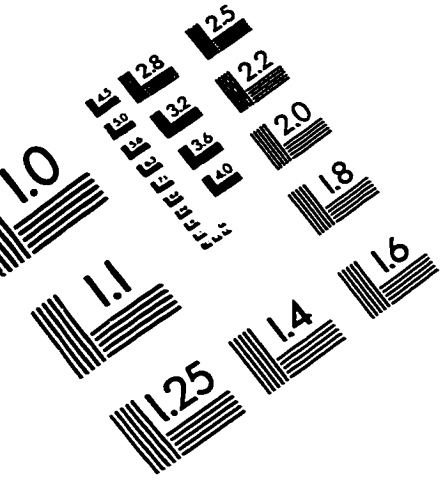
Wallace, C.J.A., J.G. Guillemette, Y. Hibiya and M. Smith (1991) "Enhancing Protein Engineering Capabilities by Combining Mutagenesis and Semisynthesis" J. Biol. Chem **266**:21355-21357

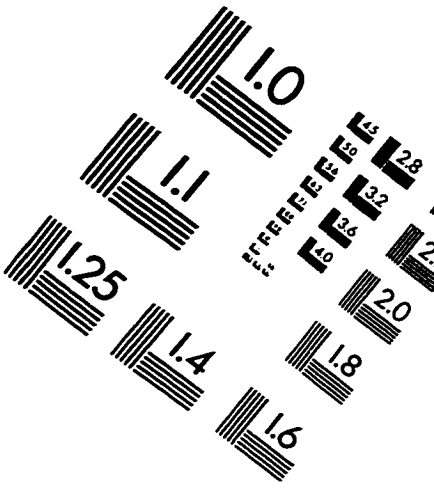
Wallace, C.J.A. (1993) "Understanding Cytochrome *c* Function: Engineering Protein Structure by Semisynthesis" Faseb J. **7**:505-515

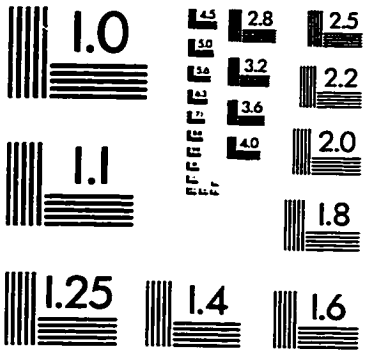
Wong, C.F., C. Zeng, J. Shen, J.A. McCammon, and P. Wolynes (1993) "Cytochrome *c*: A Molecular Proving Ground for Computer Simulations" J. Phys. Chem. **97**:3100-3110

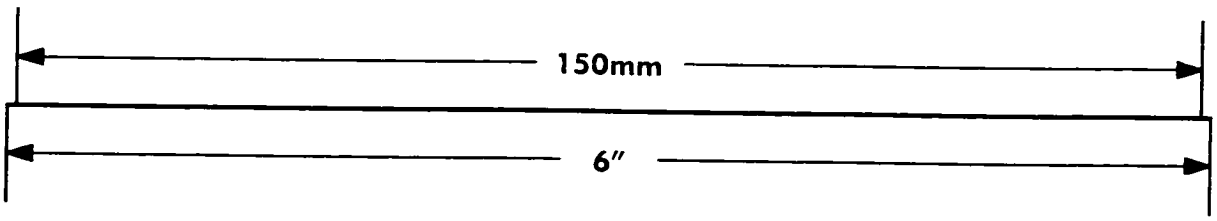
Woods, A., J.G. Guillemette, J.C. Parrish, M. Smith, and C.J.A. Wallace (1996) "Synergy in Protein Engineering" J. Biol. Chem. **271**:32008-32015

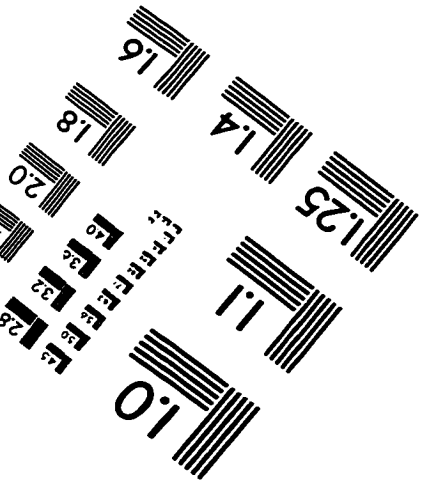
IMAGE EVALUATION TEST TARGET (QA-3)

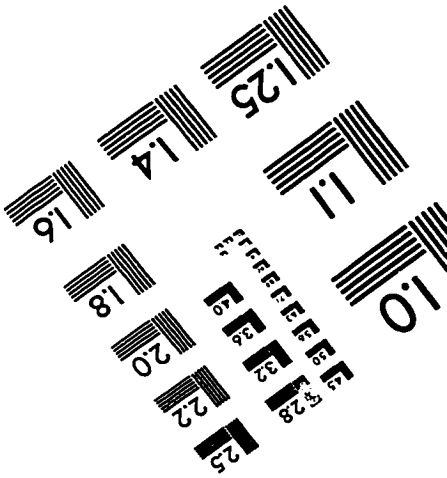












APPLIED IMAGE, Inc
1653 East Main Street
Rochester, NY 14609 USA
Phone: 716/482-0300
Fax: 716/288-5989