

**STRUCTURAL STUDIES OF THE TWO
IRON RUBREDOXIN OF
*PSEUDOMONAS OLEOVORANS***

Thesis submitted for the degree of
Doctor of Philosophy
at the University of Leicester

by

Ashlee Perry BSc
Department of Biochemistry
University of Leicester

October ²⁰⁰¹ 2000

UMI Number: U485872

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



UMI U485872

Published by ProQuest LLC 2013. Copyright in the Dissertation held by the Author.
Microform Edition © ProQuest LLC.

All rights reserved. This work is protected against
unauthorized copying under Title 17, United States Code.



ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

Abstract

Sequence alignment studies and biochemical investigations have suggested that the two iron rubredoxin of *Pseudomonas oleovorans* consists of two-rubredoxin type domains connected by a ~60 amino acid linker region. The rubredoxin gene (*alk G*) contained in the rubredoxin expression vector was targeted by site-directed mutagenesis to produce independent expression vectors for the N and C-terminal domains. The C-terminal vector directed the high-level production of the C-terminal domain and a simple procedure was used to purify the recombinant domain in the Fe-form. Attempts to directly express the N-terminal domain of the 2Fe-rubredoxin from a mini-gene were unsuccessful. However, the apo-form of the N-terminal domain was isolated through cleavage of an engineered 2Fe-rubredoxin in which a factor Xa proteolysis site had been introduced into the putative interdomain linker. The respective proteins could readily be converted to the apoprotein, cadmium or iron forms following precipitation with trichloroacetic acid and resolubilisation in the presence or absence of cadmium chloride or ferrous ammonium sulphate, respectively. Both domains have been characterised by UV-visible, circular dichroism and 1D NMR spectroscopies. The spectra collected for the C-terminal domain are characteristic of typical rubredoxins. The spectra collected for the N-terminal domain are less characteristic of typical rubredoxins, likely reflecting poor structural stability, which is consistent with the properties of the full-length di-domain rubredoxin. In steady-state assays, both N and C-terminal domain rubredoxins were redox-active and able to transfer electrons from reduced rubredoxin reductase to cytochrome *c*. The solution structure of the C-terminal domain was determined using 2D NMR techniques. The structure is typical of rubredoxins and reveals that the linker region is structureless, which is consistent with limited proteolysis studies carried out on the C-terminal domain. The structure of the full-length rubredoxin and the C-terminal domain were investigated by small angle X-ray scattering methods. These analyses confirmed the finding that the linker region in the C-terminal domain is structureless, and demonstrated that the folded N and C-terminal domains of the full-length rubredoxin are associated, with their metal binding sites in close proximity.

Acknowledgements

I would like to thank my supervisor Prof. Nigel Scrutton. Nigel has provided guidance and support during my PhD, which has proved invaluable. I would also like to thank Prof. Lu-yun Lian, who has also provided guidance especially with respect to the NMR work, her time and effort are greatly appreciated.

I would also like to thank my collaborator on the small angle X-ray scattering work, Dr. Günter Grossman, from the CCLRC Daresbury laboratory. Along with assisting in the collection and processing of data, Günter also provided technical and theoretical insight. I would like to thank Dr. Mike Sutcliffe, who was a member of my committee.

I would like to thank past and present members of lab 124, especially Richard who was there with me from the start. Special mentions go to Cameron, Matthew, François, Winston, Spencer and Phil.

Abbreviations

The following abbreviations have been used in this thesis, sometimes without previous definition.

A_{280}	Absorbance at 280 nm (similarly for other wavelengths)
ARIA	Ambiguous restraints for iterative assignment
bp	Base pairs
CC	Cytochrome <i>c</i>
CCY	Yeast iso-1-cytochrome <i>c</i>
CCH	Horse heart cytochrome <i>c</i>
CCP	Cytochrome <i>c</i> peroxidase
CD	Circular dichroism
COSY	Correlated spectroscopy
DNA	Deoxyribonucleic acid
DQF	Double quantum filtered
ET	Electron transfer
FAD	Flavin adenine dinucleotide
HOMO	Highest occupied molecular orbital
HSQC	Heteronuclear single quantum coherence
IPTG	Isopropyl β -D-thiogalactoside
MADH	Methylamine dehydrogenase
NAD	Nicotinamide adenine dinucleotide ((H)-reduced)
NADP	Nicotinamide adenine dinucleotide phosphate
NMR	Nuclear magnetic resonance
NOE	Nuclear Overhauser effect
NOESY	Nuclear Overhauser effect spectroscopy
LUMO	Lowest unoccupied molecular orbital
OD ₆₀₀	Optical density at 600 nm
PAGE	Polyacrylamide gel electrophoresis
RMSD	Root mean standard deviation
RR	Rubredoxin reductase

SAXS	Small angle X-ray scattering
SDS	Sodium dodecyl sulphate
TCA	Trichloroacetic acid
TOCSY	Total correlated spectroscopy
TTQ	Tryptophan tryptophylquinone
UV	Ultraviolet

Mutant enzymes are represented using one letter code for the wild-type amino acid; number of the amino acid in the primary sequence; one letter code for the amino acid in the mutant protein e.g. W191F = tryptophan 191 mutated to phenylalanine.

Contents

ABSTRACT.....	I
ACKNOWLEDGEMENTS	II
ABBREVIATIONS.....	III
CHAPTER 1: INTRODUCTION.....	1
1.1 INTRODUCTION.....	2
1.2 BIOLOGICAL ELECTRON TRANSFER THEORY	2
1.2.1 <i>The nuclear factor</i>	3
1.2.2 <i>The electronic factor</i>	6
1.3 INTERPROTEIN ELECTRON TRANSFER	10
1.3.1 <i>The methylamine dehydrogenase-amicyanin-cytochrome c_{551i} system</i>	11
1.3.2 <i>Cytochrome c peroxidase-cytochrome c system</i>	18
1.3.3 <i>Cytochrome f-plastocyanin system</i>	26
1.4 IDEAL SYSTEMS FOR STUDYING BIOLOGICAL ELECTRON TRANSFER	30
1.5 THE TWO IRON RUBREDOXIN OF <i>PSEUDOMONAS OLEOVORANS</i>	31
1.5.1 <i>Alkane hydroxylase system</i>	31
1.5.2 <i>Rubredoxin</i>	33
1.5.3 <i>Rubredoxin reductase</i>	39
1.5.4 <i>Alkane hydroxylase</i>	40
1.6 AIMS OF THIS THESIS.....	44
CHAPTER 2: MATERIALS AND METHODS.....	45
2.1 MATERIALS	46
2.1.1 <i>Chemicals</i>	46
2.1.2 <i>Equipment</i>	46
2.1.3 <i>Molecular weight markers</i>	46
2.1.4 <i>Bacterial strains and media</i>	47
2.1.5 <i>DNA modifying enzymes</i>	47
2.1.7 <i>Chromatographic media and membranes</i>	47
2.1.8 <i>Primers</i>	48
2.2 METHODS.....	48
2.2.1 <i>DNA purification</i>	48
2.2.2 <i>Agarose gel electrophoresis</i>	49
2.2.3 <i>Polyacrylamide gel electrophoresis</i>	49
2.2.4 <i>Preparation of competent cells using the rubidium chloride method</i>	49
2.2.5 <i>Site-directed mutagenesis</i>	50

2.2.6	<i>Restriction digest</i>	51
2.2.7	<i>DNA ligation</i>	51
2.2.8	<i>Phenol/chloroform extraction</i>	51
2.2.9	<i>Determination of protein concentration</i>	51
2.2.10	<i>Transformation of competent cells</i>	52
2.2.11	<i>SDS polyacrylamide gel electrophoresis</i>	52
2.2.12	<i>DNA sequencing</i>	53
2.2.13	<i>N-terminal sequencing</i>	53
2.2.14	<i>Electrospray mass spectrometry</i>	53
2.2.15	<i>Preparation of Fe-, Cd and apo- forms of the rubredoxin</i>	53
2.2.16	<i>Steady-state assay of rubredoxin</i>	54
2.2.17	<i>NMR</i>	54
2.2.18	<i>Limited proteolysis</i>	55
2.2.19	<i>X-ray scattering data collection</i>	55
CHAPTER 3: ISOLATION OF THE INDIVIDUAL DOMAINS		56
3.1	INTRODUCTION.....	57
3.2	MOLECULAR BIOLOGY.....	58
3.3	PURIFICATION.....	61
3.3.1	<i>Protein Expression</i>	61
3.3.2	<i>Protein Purification</i>	61
3.4	PROTEIN CHARACTERISATION.....	66
3.4.1	<i>UV-Visible Spectroscopy</i>	66
3.4.2	<i>Far-UV CD Characterisation</i>	70
3.4.3	<i>NMR Spectroscopy of the Individual Domains</i>	72
3.4.5	<i>Steady-State Characterisation</i>	76
3.4.6	<i>Limited Proteolysis</i>	77
3.5	DISCUSSION.....	79
CHAPTER 4: NMR ANALYSIS OF THE C-TERMINAL DOMAIN		81
4.1	INTRODUCTION.....	82
4.2	ASSIGNMENT STRATEGY.....	83
4.3	SPIN SYSTEM ASSIGNMENTS.....	84
4.4	SEQUENTIAL SEQUENCE ASSIGNMENTS.....	85
4.5	SECONDARY STRUCTURE IDENTIFICATION.....	90
4.6	STRUCTURAL RESTRAINTS.....	92
4.6.1	<i>NOE connectivities</i>	92
4.6.2	<i>Hydrogen bonds</i>	92
4.6.3	<i>Metal binding sites</i>	93
4.7	STRUCTURE GENERATION.....	93

4.8	STRUCTURE DESCRIPTION AND ANALYSIS	98
4.9	DISCUSSION.....	102
CHAPTER 5: SAXS ANALYSIS		104
5.1	INTRODUCTION.....	105
5.2	SAXS.....	105
5.2.1	<i>SAXS theory and applications</i>	105
5.2.2	<i>Molecular shape computation</i>	107
5.3	SAXS ANALYSIS OF THE C-TERMINAL DOMAIN RUBREDOXIN.....	109
5.4	SAXS ANALYSIS OF THE DI-DOMAIN RUBREDOXIN	112
5.4.1	<i>2Fe di-domain rubredoxin</i>	112
5.4.2	<i>1Fe di-domain rubredoxin</i>	113
5.5	DISCUSSION.....	116
CHAPTER 6: DISCUSSION		119
6.1	DISCUSSION.....	120
6.2	FUTURE WORK.....	123
REFERENCES.....		125

Chapter 1

Introduction

1.1 Introduction

Long-range interprotein electron transfer (ET) is fundamental to respiration, photosynthesis and the redox reactions of intermediary metabolism, and is therefore one of the most important and ubiquitous reactions in biological systems. It involves donor and acceptor proteins with redox centres that are separated by relatively long distances (10-30 Å), and the ET reactions themselves may involve several steps i.e. binding of proteins, protein rearrangements and chemical transformations, such as proton transfer, before the actual ET step. Therefore, the study of biological ET reactions requires a multidisciplinary approach. This is illustrated by the variety of techniques used in the study of ET, which include, site-directed mutagenesis, X-ray crystallography, computer simulations, NMR, and a variety of sophisticated kinetic measurements.

ET reactions are important to many fundamental biological processes and have been an area of intense research activity over the years. The level of knowledge has increased over the last decade. Advances in recombinant DNA technology have made the production of recombinant proteins, and production of specific mutants, relatively straightforward procedures. This has led to a dramatic increase in the number of protein structures solved, among which are the crystal structures of several ET complexes.

1.2 Biological electron transfer theory

ET reactions occur when there is an overlap between the orbitals (wavefunctions) of the electrons of donor and acceptor molecules. In biological ET reactions redox centres are embedded in the proteins (preventing direct contact between the centres), which act as an insulating medium. This insulating medium forms a high potential energy barrier between the redox centres (typically 10-30 Å apart), thus at physiological conditions ET is possible only by the quantum tunnelling of electron wavefunctions.

Chemical reactions are traditionally modelled using transition state theory. Reactants move along a potential energy surface, the activation energy is overcome (through thermal vibrations) and the transition state is achieved. This transition

state almost inevitably decays to form the products, and such a reaction would be described as adiabatic. This is in contrast to biological ET reactions where the probability of ET occurring between the distant (10-30 Å), weakly coupled donor and acceptor, at such a transition state is small. Thus, these reactions are described as being non-adiabatic (or diabatic), and a mathematical description of the ET rate (k_{et}) is described by "Fermi's Golden Rule" (Equation 1.1):

$$k_{\text{ET}} = \frac{2\pi}{\hbar} H_{\text{AB}}^2 (FC)$$

Equation 1.1

Where \hbar is Planck's constant divided by 2π , and H_{AB}^2 is the electronic factor representing the electronic coupling between the reactant and the product wavefunctions, giving the probability of ET once the productive nuclear transient state has formed. *FC*, the Franck-Condon factor, is the nuclear factor that determines the total density of the productive nuclear transient states, which are able to transfer electrons. This equation is essentially composed of two parts, the electronic factor and the nuclear factor, and serves to separate the dependence of the ET rate into electronic and nuclear terms.

1.2.1 The nuclear factor

Franck and Condon suggested that when an electron jumps from one quantum level to another in a molecule of more than one atom, it does so in a time too short for interatomic distances or velocities to change (Condon, 1926; Franck, 1925). Thus, immediately after ET, the nuclear configuration of the initial state is still preserved. This does not correspond to the equilibrium nuclear configuration for the final state but rather to a vibration-excited configuration of the final state. The consequence of this is that all ET reactions require reorganisation energy to overcome the cost of distorting the equilibrium geometry of the reactant state into the equilibrium geometry of the product state, in the absence of ET. This reorganisation of the protein scaffold is a prerequisite for ET. A requirement for nuclear reorganisation is confirmed by the demonstration of 'overvoltage' (the extra

electrode voltage above the equilibrium needed to observe electrolytic current) in electrolysis (Gurney, 1931).

As proteins are multi-atom systems the principle of nuclear reorganisation needs to be applied. Also, there should be no external energy input since biological ET reactions are closed systems, where energy must be conserved. It is useful to illustrate the energy terms that comprise the *FC* factor using potential energy diagrams (Figure 1.1). The energy of the reactants and surrounding medium, a function of all nuclei whose motion is coupled to that of the electron, is represented as a single harmonic potential energy surface. Similarly, the energy of the products and surrounding medium is also represented as a harmonic potential, but with its equilibrium position and energy minima moved as a result of the ET reaction. The energy conservation principle requires that ET must occur at nuclear configurations where the total potential energy of the reactants and the surrounding medium is equal to that of the products and surrounding medium. Therefore, the only position on the graph where both nuclear geometry and the potential energy of the product and reactant correspond is at the intersection point of both curves. Thus, for ET to occur sufficient activation energy (ΔG^*) must be provided via thermal nuclear vibration; this is the energy difference between the reactant energy minima and the intersection point.

Marcus (1956) derived an equation (Equation 1.2) that relates the activation free energy (ΔG^*) to the driving force of the reaction (ΔG), and the reorganisational energy (λ). This is called the "Marcus energy gap law", which states:

$$\Delta G^* = \frac{(\Delta G + \lambda)^2}{4\lambda}$$

Equation 1.2

The effect of the Marcus energy gap law on the rate of ET is illustrated in Figure 1.1. As $-\Delta G$ increases (with $-\Delta G < \lambda$), the activation energy decreases and the rate increases; this is called the "normal region". The activation energy decreases until $-\Delta G = \lambda$, at which point the activation energy is zero and the rate is at its maximum; this region is called the "activationless region". As $-\Delta G$ increases

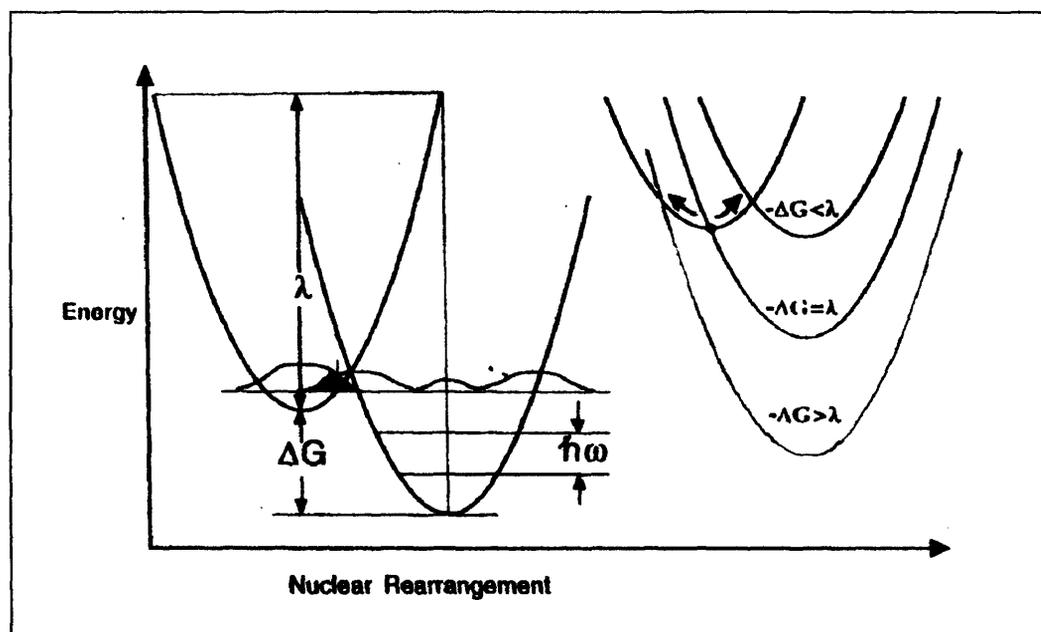


Figure 1.1. Potential energy diagram. The nuclear motions of the reactants (left) and the products (right) are approximated as a single harmonic potential energy surface. The equilibrium geometry corresponds to the bottom of the potential wells. The product potential has the same shape, but is shifted with respect to both potential energy and nuclear geometry. The potential energy is shifted by the free energy of the reaction (ΔG), which corresponds to the difference between the reactants and products energy minima. The nuclear geometry is shifted due to the change in equilibrium nuclear geometry after the ET reaction. The reorganisation energy (λ), corresponds to the energy that must be added to the reactants to move from the equilibrium reactant nuclear geometry to the equilibrium product nuclear geometry, but remaining on the reactant potential surface (without ET occurring). In the quantised view of the simple harmonic oscillator potentials only certain energy levels are permitted, shown as horizontal lines spaced out by the quantum energy of the oscillator, $\hbar\omega$. In the quantised view nuclear tunnelling from reactant to product depends upon the overlap of the harmonic oscillator wavefunctions, illustrated as the shaded region. The right hand side of the diagram illustrates three different situations. When $-\Delta G = \lambda$, there is no activation energy, and the reaction rate is highest. When $-\Delta G < \lambda$ (the normal region) and when $-\Delta G > \lambda$ (the inverted region) the activation energy is non-zero and the reaction rate is slower. Taken from Moser, 1996.

further ($-\Delta G > \lambda$) the rate falls again, giving rise to the "inverted region". This has been demonstrated experimentally for a number of systems (Millar *et al.*, 1984; Moser *et al.*, 1992).

When this equation is applied to the Arrhenius equation, combined with Fermi's golden rule (Equation 1.1) and the equation for the liquid phase collision frequency, the "classical equation" for the rate of ET is derived (Equation 1.3):

$$k_{\text{ET}} = \frac{4\pi^2 H_{\text{AB}}^2}{h(4\pi\lambda RT)^{1/2}} e^{[-(\Delta G + \lambda)^2 / 4\lambda RT]}$$

Equation 1.3

Where R is the gas constant and T is the absolute temperature. The classical equation is accurate at physiological conditions, but at low temperatures it breaks down. This was first observed by Chance & Nishimura (1960), in their measurement of ET reaction rates in the bacterium *Chromatium vinosum* at 77 K (Chance & Nishimura, 1960). From equation 1.3 it is expected that the nuclear vibrations are frozen at the ground state at 77 K and the nuclei would not be able to reach the intersection point by thermal vibrations. Therefore, the ET observed by Chance & Nishimura (1960) can only be explained by quantum mechanical tunnelling of the nuclear wavefunction from the reactant to potential energy curve. At low temperatures ET can be explained in terms of the nuclear wavefunctions of the reactant state tunnelling through the potential barrier to overlap with the nuclear wavefunctions of the products, as illustrated in Figure 1.1. In this situation the classical interpretation of the FC factor is not appropriate and it must be expressed quantum mechanically (for review see Moser & Dutton, 1996).

1.2.2 The electronic factor

The electronic factor (H_{AB}), determines the probability of ET once the transition state has been achieved, and is a function of the coupling of reactant and product wavefunctions. The degree of orbital overlap is dependent on the distance between the redox centres and the nature of the intervening medium that separates

SPECIAL NOTE

**ITEM SCANNED AS SUPPLIED
PAGINATION IS AS SEEN**

further ($-\Delta G > \lambda$) the rate falls again, giving rise to the "inverted region". This has been demonstrated experimentally for a number of systems (Millar *et al.*, 1984; Moser *et al.*, 1992).

When this equation is applied to the Arrhenius equation, combined with Fermi's golden rule (Equation 1.1) and the equation for the liquid phase collision frequency, the "classical equation" for the rate of ET is derived (Equation 1.3):

$$k_{\text{ET}} = \frac{4\pi^2 H_{\text{AB}}^2}{h(4\pi\lambda RT)^{1/2}} e^{[-(\Delta G + \lambda)^2 / 4\lambda RT]}$$

Equation 1.3

Where R is the gas constant and T is the absolute temperature. The classical equation is accurate at physiological conditions, but at low temperatures it breaks down. This was first observed by Chance & Nishimura (1960), in their measurement of ET reaction rates in the bacterium *Chromatium vinosum* at 77 K (Chance & Nishimura, 1960). From equation 1.3 it is expected that the nuclear vibrations are frozen at the ground state at 77 K and the nuclei would not be able to reach the intersection point by thermal vibrations. Therefore, the ET observed by Chance & Nishimura (1960) can only be explained by quantum mechanical tunnelling of the nuclear wavefunction from the reactant to potential energy curve. At low temperatures ET can be explained in terms of the nuclear wavefunctions of the reactant state tunnelling through the potential barrier to overlap with the nuclear wavefunctions of the products, as illustrated in Figure 1.1. In this situation the classical interpretation of the FC factor is not appropriate and it must be expressed quantum mechanically (for review see Moser & Dutton, 1996).

1.2.2 The electronic factor

The electronic factor (H_{AB}), determines the probability of ET once the transition state has been achieved, and is a function of the coupling of reactant and product wavefunctions. The degree of orbital overlap is dependent on the distance between the redox centres and the nature of the intervening medium that separates

them. In biological ET the redox centres are situated in proteins, which insulate them, thus ET can only occur by electron tunnelling. The rate of tunnelling decays exponentially with distance, as predicted by Gamow's tunnelling equation (Equation 1.4) (Gamow, 1928):

$$H_{AB}^2 = H_0^2 e^{(-\beta r)}$$

Equation 1.4

Where H_0 is the electronic coupling between the two redox centres in van der Waals contact (equivalent to the maximum electronic coupling), β is the exponential coefficient of electronic coupling decay, and r the 'edge-to-edge' distance between donor and acceptor. The ET transfer rate is predicted to depend on distance (r) as described by Equation 1.5, which is a combination of the classical Marcus theory with Gamow's tunnelling equation for the electronic factor:

$$k_{ET} = k_0 e^{-\beta(r-r_0)} e^{-(\Delta G + \lambda)^2 / 4\lambda RT}$$

Equation 1.5

In equation 1.5, k_0 is the characteristic vibration of the nuclei, which is assigned a value of 10^{13} s^{-1} (Marcus & Sutin, 1985) and r_0 represents the van der Waals distance (3 Å). From these equations it can be seen that the decay factor β is exceptionally important in the ET reaction. This is confirmed by the work of Moser *et al.* (1992).

Studies of ET in photosynthetic reaction centres of *Rhodobacter sphaeroides*, *Rhodobacter viridis*, ruthenated cytochrome *c* and ruthenated myoglobin have indicated that the protein medium between the redox centres behave homogeneously (with respect to β) irrespective of the intervening protein structure (Moser *et al.*, 1992). Plots of the log of the optimal ET rate (where $\Delta G = \lambda$) against the edge-to-edge distance of donor and acceptor redox centres were analysed. A linear relationship between the two parameters was obtained which indicates that the ET rate decays exponentially with a single common decay factor, determined from the slope of the plot as 1.4 \AA^{-1} . This indicates proteins present a

uniform electronic barrier to electron tunnelling, and the ET rate, with respect to the electronic factor, is dependent on distance, irrespective of the protein structure.

However, other experiments have demonstrated an anisotropic coupling decay. Using similar analysis as that performed by Moser *et al.* (1992) the maximal ET rate in cytochrome *c*, ruthenium modified at four different amino acids (each one giving a different edge-to-edge distance) was investigated (Wuttke *et al.*, 1992). A plot of the log of the optimal ET rate against edge-to-edge distance did not produce a linear relationship. This suggests that in this system the intervening protein structure provides a heterogeneous electronic barrier to electron tunnelling, thereby highlighting a role for both structure and distance in determining the electronic factor. Studies of other ruthenated proteins have confirmed these findings, and it is now a widely held view that the simple homogeneous model can not be universally applied to all biological ET systems.

In order to better understand the heterogeneous model the concept of "super exchange" has been used (Beratan *et al.*, 1991). According to super exchange the electronic coupling of the wavefunctions between the donor and acceptor is mediated by the HOMO (highest occupied molecular orbital) and LUMO (lowest unoccupied molecular orbital) of the protein bridging atoms. This approach assumes that ET follows pathways of protein structure and that these pathways can be divided into individual steps. For each step the electronic coupling, H_{AB} , decays by a factor of ϵ . This electronic decay factor can be defined as the probability of ET through a connectivity, and the total decay across the pathway is sum of the individual decay factors.

With this in mind an algorithm has been devised which enables total decay factors to be calculated, and can also be used to identify potential ET pathways (Beratan *et al.*, 1991; Beratan *et al.*, 1992; Betts *et al.*, 1992). In this algorithm three types of connectivities are identified, (i) covalent connections (ϵ_C), (ii) hydrogen bond connections (ϵ_H), and (iii) "through space jumps" (ϵ_S). Semi-empirical expressions for the decay functions have been determined by calibration with ET rates from ruthenated proteins (Beratan *et al.*, 1991).

The pathway model allows the electrons to choose a pathway which both maximises the coupling and minimises the distance, thereby optimising the

electron coupling factor, H_{AB} . The model highlights the importance of hydrogen bonds in biological ET, which has been supported by experimental evidence (Therien *et al.*, 1990). Additionally, the model indicates that β -sheet regions, with a higher content of hydrogen bonds, will exhibit a greater degree of coupling than α -helix regions, which is also supported experimentally (Langen *et al.*, 1995). The pathway algorithm has been used to identify ET pathways in ruthenated proteins, and these pathways can be converted into the 'effective' tunnelling distance. A plot of the effective tunnelling distance against $\log k_{ET}$ produces a linear plot, thus indicating a dependence between ET and the effective tunnelling distance (Langen *et al.*, 1995; Wuttke *et al.*, 1992). The pathway model has been modified to accommodate interprotein ET (Aquino *et al.*, 1995), and also the effects of multiple pathways (Curry *et al.*, 1995).

Experimental data has demonstrated that the two models (homogeneous and heterogeneous) are not mutually exclusive. There are systems where the data fits a homogeneous model with the ET rate decaying exponentially with distance, but the β value differs from system to system (Langen *et al.*, 1995; Moser *et al.*, 1992). However, there are also systems where the data fits the heterogeneous model, where ET rate does not fall exponentially with distance. Additionally, both homogeneous and heterogeneous models can co-exist in the same system (Evenson & Karplus, 1993). Studies of tuna cytochrome *c* demonstrated that in one regime the effective coupling changes exponentially with distance, while in another regime the distance dependence is more complex and sensitive to intervening protein structure.

Finally, a recent study has questioned the whole idea of super exchange, and raised some interesting issues with respect to the biology of ET (Page *et al.*, 1999). This work, carried out with several natural multi-redox centre oxidoreductases, indicates that the intervening protein medium is not important as long as the redox centres are certain distance apart. For these oxidoreductases it is argued that the 14 Å (or less) spacing of redox centres imparts highly robust protection to the protein, with respect to naturally occurring mutations. The implications of this are that the control of ET reactions may require much more than simple mutagenic replacement of residues in the ET pathway. For individual

classes of ET proteins, the specific structure of the class as a whole may be the predominant mechanism controlling the ET rate by dictating the distance between the redox centres. Thus, to truly control ET rates may require control of the structure as a whole, which is beyond the scope of current knowledge. However, whether this theory holds true for other ET system remains to be seen, but it is clear that there is still much to learn with regard to biological ET reactions.

1.3 Interprotein electron transfer

ET reactions occur between redox centres housed in different protein molecules. Therefore, a prerequisite for ET is the formation of a productive encounter complex between the redox proteins. The specific binding of protein molecules is seen widely in biology, but for ET complexes there are special considerations. While some biological protein interactions are highly specific, and others more transient, the reactions between redox partners is somewhere in-between. The interactions must be specific enough so that the electron flow is not misdirected by chance protein-protein encounters, but weak enough to allow the productive ET complex to disassociate after ET. Thus, there is a fine balance between specificity and strength of binding. Couple with this the fact that some proteins, such as rubredoxin and amicyanin, shuttle electrons between two separate and distinct types of protein, the control these proteins confer over the directionality and specificity of ET reactions is impressive (Husain & Davidson, 1985; Peterson & Coon, 1968).

There is some evidence that redox linked conformational changes may aid the process of dissociation (Drepper *et al.*, 1996). However, while this may be the case in some systems the majority of protein structures are not significantly altered by the redox state for this to be the case (Day *et al.*, 1992). Therefore, protein recognition must be governed by weak forces resulting in a complex that can disassociate. Since protein-protein interactions between redox partners are weak, resulting in mostly transient complexes, structural determination of intact ET complexes by X-ray crystallography is extremely difficult. However, several physiological ET complexes have been crystallised and their structures determined, such as the complex of methylamine dehydrogenase (MADH) and amicyanin

(Chen *et al.*, 1992), and the complex of cytochrome *c* and cytochrome *c* peroxidase (Pelletier & Kraut, 1992). These structures have been extremely useful in both confirming previous experimental data, and providing impetus for further experiments.

The exact nature by which proteins form productive encounter complexes is not fully understood. Electrostatic interactions are known to be essential in biological ET as many ET rates show an ionic strength dependence. Furthermore, the structures of several redox proteins reveal that surface charges are asymmetrically positioned, and that the respective redox partners contain high concentrations of complementary surface charge (Roberts *et al.*, 1991; Salemmé, 1976). On the basis of these observations it has been postulated that long-range electrostatic interactions "steer" the redox partners into the productive encounter complex during diffusional collisions (Northrup *et al.*, 1988; Roberts *et al.*, 1991). However, it has also been postulated that proteins employ the principle of "reduction in dimensionality", whereby proteins initially associate in random fashion via weak electrostatic interactions. This loosely associated system may then rotationally realign (or diffuse) into the optimum configuration for ET. Thus, the study of the dynamic nature of protein-protein interactions, which lead to ET, requires a multi-disciplinary approach.

The systems discussed in the following Sections (1.3.1 to 1.3.3) represent the most thoroughly characterised interprotein ET systems known to date. Several themes have been identified in these systems, which seem universally important for interprotein ET reactions. Namely, the importance of electrostatic interaction, and the importance of protein dynamics for effective ET. These systems also give an indication of the experimental procedures used, and their limitations.

1.3.1 The methylamine dehydrogenase-amicyanin-cytochrome c_{551i} system

The Methylamine Dehydrogenase-Amicyanin-Cytochrome c_{551i} system is one of only a few systems where the inter-protein ET complex structure is known. MADH, amicyanin and cytochrome c_{551i} are periplasmic proteins, which are

induced when the bacterium *Paracoccus denitrificans* is grown on methylamine as the sole carbon source (Husain & Davidson, 1985). MADH catalyses the oxidation of methylamine to formaldehyde and ammonia and donates electrons to amicyanin. The cytochrome c_{551i} is thought to be the physiological electron acceptor for amicyanin and acts as an intermediate for transferring the electrons to a membrane bound cytochrome oxidase (Husain & Davidson, 1986).

MADH is a soluble periplasmic enzyme with an $\alpha_2\beta_2$ structure. The subunits have molecular masses of 47 kDa (H subunit) and 15 kDa (L subunit) (Husain & Davidson, 1987). For some time it was thought that the redox cofactor for MADH was pyrroloquinoline quinone (PQQ). Although structurally similar to PQQ the actual cofactor for MADH is tryptophan tryptophylquinone (TTQ) (Chen *et al.*, 1991; McIntire *et al.*, 1991). TTQ is formed by the post-translational modification of two tryptophan residues from the L subunit (Trp 57 and Trp 107). One of the tryptophan residues exists as an orthoquinone (indole-6, 7-dione) and is cross-linked via its 4 position to position 2 of the second tryptophan (see Figure 1.2). Amicyanin is a 11.5 kDa type one blue copper protein (sometimes referred to as a cupredoxin), and cytochrome c_{551i} is a 17.5 kDa protein containing one heme group.

X-ray crystal structures have been solved for the binary complex formed between MADH and amicyanin, and the ternary complex between MADH, amicyanin and cytochrome c_{551i} (Chen *et al.*, 1992; Chen *et al.*, 1994). It should be noted that the crystals were grown in high salt conditions, which needs to be taken into account when analysing these structures.

The complex of MADH with amicyanin comprises a heterohexamer of the type $H_2L_2A_2$. Although each amicyanin molecule is in contact with both the H and the L subunit (of a single HL dimer of the MADH molecule), the major intermolecular interactions occur between amicyanin and the L subunit. Approximately 12 non-polar or neutral side chains on the MADH molecule are involved in the interface with each amicyanin, three on the H and nine on the L subunit (including the Trp 107 moiety of the TTQ cofactor). This is consistent with chemical cross-linking experiments that indicate amicyanin is in contact with both the MADH subunits, although these studies suggest that the primary site for interaction is the

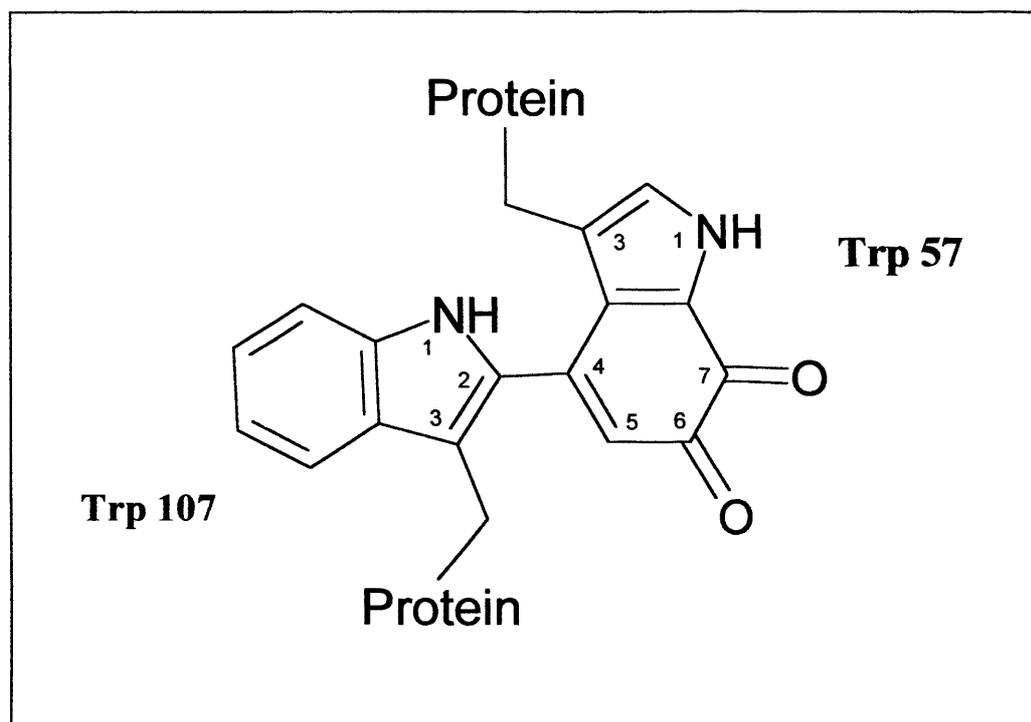


Figure 1.2. The structure of the tryptophan tryptophylquinone (TTQ) cofactor.

H subunit; these studies were also performed in high salt conditions (Kumar & Davidson, 1990).

The copper of amicyanin is located about 2.5 Å from the protein surface, co-ordinated in approximately tetrahedral fashion to the ligands His 53, Cys 92, His 95 and Met 98. Three of the ligands are buried within the protein, but the fourth, His 95, lies on the protein surface. On the surface, His 95 sits within the hydrophobic patch formed by seven hydrophobic residues, with the copper atom lying underneath this patch. On complex formation, each amicyanin loses approximately 600 Å² of accessible surface area. Of this area, 63 % is contained within the hydrophobic patch, indicating that hydrophobic interactions are the dominant form of interaction between amicyanin and MADH. The cross-linking experiments also suggest that the complex is stabilised by hydrophobic interactions. However, both the cross-linking and complex crystallisation experiments were performed in high salt conditions. Such conditions favour hydrophobic interactions and deter hydrophilic interactions, which makes the above results not unexpected.

The TTQ is orientated so that the o-quinol portion of Trp 57 is pointed away from the copper, while the Trp 107 portion is closer to the copper, with closest approach of 9.3 Å. The indole moiety of TTQ is exposed on the surface of the L subunit, facing the copper region of the amicyanin. The active site of MADH is centred on the o-quinol (of Trp 57 in the TTQ), and this is far enough away from the amicyanin binding site to allow substrate binding and product release without interfering with the ET to the copper atom. One of the copper ligands of amicyanin, His 95, is situated between the copper and TTQ group. The closest distance between His 95 and TTQ is 5.4 Å, therefore His 95 is well placed to mediate ET.

Electrostatic interactions have long been thought to play a crucial role in stabilising protein-protein interactions between soluble redox partners. Although the structural data does not concur with this, the conditions in which these results were obtained i.e. high salt concentrations, may have masked the electrostatic interactions involved in this system. Certainly, previous studies of the MADH and amicyanin system (chemical cross-linking (Kumar & Davidson, 1990), spectroscopic, redox potentials (Gray *et al.*, 1988) and kinetic efficiency (Davidson & Jones, 1991)) were sensitive to ionic strength in a manner that suggests an important role for electrostatic interactions.

The X-ray structure of the ternary complex shows that amicyanin is sandwiched between MADH and cytochrome c_{551i} (see Figure 1.3). The orientation of the MADH to amicyanin is identical to that seen in the binary complex and the association is very similar (the amicyanin loses 770 Å² of accessible surface area in the ternary complex compared to 600 Å² in the binary complex). The cytochrome molecule is adjacent to amicyanin, but is not in contact with the MADH. The iron of the heme group and the copper are about 24.8 Å apart. Within the complex, the L subunit, amicyanin and cytochrome c_{551i} sit in a row, with the C6 of Trp 57 separated from the iron atom by 40.1 Å. This arrangement of the proteins (linear, with no contact between cytochrome c_{551i} and MADH) is consistent with the fact that no ET takes place between MADH and cytochrome c_{551i} in the absence of amicyanin. This also supports the idea that a ternary complex is formed in solution. The amicyanin-cytochrome c_{551i} interface is considerably more polar than the MADH-amicyanin interface and also smaller, occupying 440 Å² of accessible

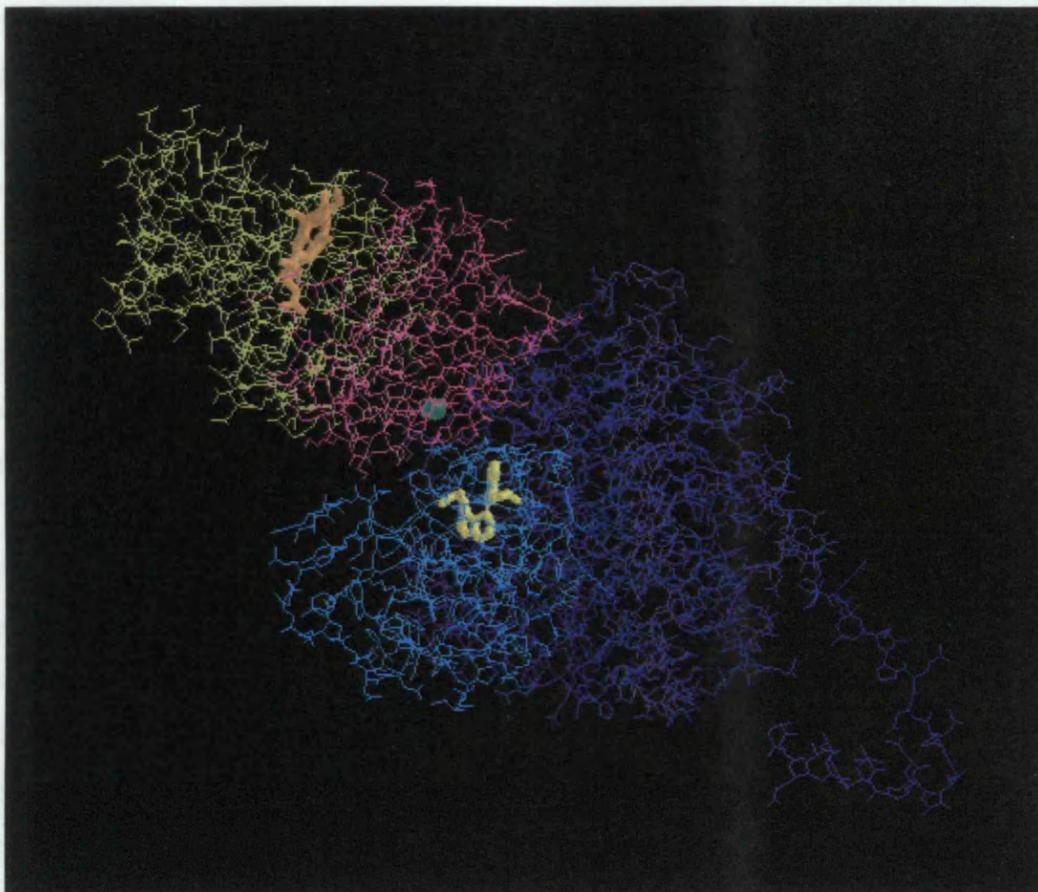


Figure 1.3. Structure of the ternary complex. Proteins are shown as wireframes and colours assigned as follows: yellow for cytochrome c_{551i} , purple for amicyanin, light blue and dark blue for the L and H subunits of MADH respectively. The redox centres are shown in thick wireframe (heme and TTQ) or spacefill (copper) and the colours assigned as follows: red for the heme of cytochrome c_{551i} , green for the copper of amicyanin and yellow for the TTQ of MADH. For clarity, only half of the complex is shown. Figure produced using PDB file 2MTA.

surface area on the amicyanin. The number of links in the amicyanin-cytochrome c_{551i} interface is greater than that between amicyanin and MADH.

Analysis of the mid-point potentials in this system provided some unusual results. Free amicyanin has an E_m value of 294 mV and cytochrome c_{551i} 190 mV (Gray *et al.*, 1986). These values make ET between these two proteins unfavourable, which is consistent with the observation that amicyanin is a poor reductant for cytochrome c_{551i} . In fact, no ET occurs between reduced amicyanin and cytochrome c_{551i} in the absence of MADH (Davidson & Jones, 1995; Davidson

& Jones, 1996). However, upon complex formation with MADH, there is a 73 mV shift in the E_m value of amicyanin (from 294 to 221 mV). This results in a significant narrowing of the difference between the E_m values of amicyanin and cytochrome c_{551i} from 104 to 31 mV, which ultimately makes ET favourable (Gray *et al.*, 1988). Further investigations demonstrated that the presence of MADH does not have a significant effect on the E_m value for cytochrome c_{551i} . Therefore, the enhancement of ET in this system is due to interactions between MADH and amicyanin. These studies emphasise that neither redox potentials nor kinetic data alone should be used as the sole basis for assigning sequences of ET proteins.

The temperature-dependence of the rate of ET was measured between dithionite-reduced MADH and oxidised amicyanin, and the data analysed by ET theory (Brooks & Davidson, 1994b). The ET reaction exhibited a large apparent reorganisational energy of approximately 225 kJ mol⁻¹ (2.3 eV) and coupling matrix element of 11.7 cm⁻¹. These values were confirmed by analysis of the free energy dependence of the ET reaction (Brooks & Davidson, 1994a). In these experiments the ET rates from the fully reduced and the semiquinone TTQ and the reverse reactions were measured. A plot of $\ln k_{ET}$ against ΔG was analysed by ET theory, using Equation 1.3, with H_{AB} value of 11.7 cm⁻¹ and a λ value of 2.2 eV. The fit of the data to this equation was excellent, thus indicating that the experimentally determined λ value was real. Therefore, it was postulated that the large reorganisational energy was due either to a reorientation of the two proteins once complexed, a change in the torsion angle between the ring systems of TTQ, or a combination of both. To investigate this, Phe 97 of amicyanin, which is thought to be involved in the interface with MADH, was mutated to Glu. Kinetic analysis of this mutant showed that the value of λ did not change significantly (Davidson *et al.*, 1998). This suggests little re-orientation of proteins with respect to each other at the protein-protein interface, and points towards a rapid but unfavourable re-orientation of the two TTQ rings being coupled to the ET reaction. It would be unlikely for such an event, which takes place within MADH, to be affected by the mutation to amicyanin.

The ET reaction from copper to heme, within the MADH-amicyanin-cytochrome c_{551i} complex, has also been analysed using transient kinetic and

thermodynamic methods (Davidson & Jones, 1996). This reaction exhibited a re-organisational energy of approximately 1.1 eV and coupling matrix element of 0.3 cm^{-1} . This implies that there is little re-arrangement occurring within the ternary complex to achieve the optimum orientation required for ET between copper and heme.

Analysis of kinetic data using ET theory has allowed determination of the distances between TTQ and the copper redox centres in the binary complex, and copper and heme in the ternary complex. The kinetically determined distance between TTQ and copper is 7.9-12.4 Å using β values of 1.4 to 0.7 Å^{-1} , which is comparable to the distance of 9.4 Å determined from the crystal structure (Brooks & Davidson, 1994b). The kinetically determined distance between the copper and the heme is 13-24 Å using β values of 1.4 to 0.7 Å^{-1} ; again the upper distance compares well with the distance of 24.8 Å determined from the crystal structure (Davidson & Jones, 1996). The β values used are upper and lower values observed in other protein ET reactions (Moser *et al.*, 1992; Onuchic *et al.*, 1992). The crystal structures were also analysed using pathway prediction software for ET reactions (Greenpath v0.97 and Pathways II), and these predictions correlate well with the kinetically determined distances (and those from the crystal structure). Additionally, the H_{AB} values in both cases seem consistent with the predicted ET pathways.

These results suggest that at the very least the crystal structures are good approximations of the real ET complexes. Further evidence for this comes from kinetic studies on the crystals of the ternary complex. These crystals are able to oxidise methylamine, and the electrons derived from this are passed onto the copper of amicyanin and then to the heme of the cytochrome c_{551i} (Merli *et al.*, 1996). It should be noted that the ET rates observed for the crystalline complex are slower than those observed in solution.

For the MADH system there is good correlation between kinetic and structural data, which along with its analysis by ET theory has provided insight into biological ET reactions. The studies have also demonstrated the positive aspects as well as the limitations of X-ray crystallography. Both the binary and ternary complexes have been crystallised and their structures solved, which is quite

an accomplishment considering that the ternary complex is over 180 kDa. However, there may be some masking of important interactions in these structures due to the conditions in which the crystals were grown. Additionally, due to the nature of the technique, the structures essentially represent a single 'snapshot' of the ET reaction. Therefore, although these structures may be similar to the working ET complex, they provide little information on the process between the initial binding reaction and ET. However, this should not detract from the significance of these studies. A challenge for the future is to see how the crystal structure might be manipulated to achieve the efficiency of electron transfer that is observed in solution.

1.3.2 Cytochrome *c* peroxidase-cytochrome *c* system

One of the most highly developed model systems for studying interprotein ET is the cytochrome *c* peroxidase-cytochrome *c* system (CCP:CC). This system has been studied extensively using kinetics, modelling and electrochemical methods, and structures are available for both the individual proteins and the complex. Some of the most important processes in biological systems, such as respiration and photosynthesis, are mediated by cytochromes. Cytochrome *c* (CC) is a 12.5 kDa water soluble protein containing a covalently bound heme, which is involved in respiration and facilitates the flow of electrons between two large multimeric membrane bound proteins, cytochrome reductase and cytochrome oxidase (Pettigrew & Moore, 1987). The crystal structures for these latter two proteins have recently been solved (Yoshikawa *et al.*, 1998; Zhang *et al.*, 1998). CC is involved in an additional physiological ET process with cytochrome *c* peroxidase (CCP). CCP is a 31 kDa water soluble yeast mitochondrial protein, containing a non-covalently bound heme, whose crystal structure has been solved at high resolution (1.7 Å) (Finzel *et al.*, 1984). CCP was originally earmarked as a simplified model for cytochrome oxidase, and although the biological function of CCP remains ambiguous, the complex formed between CCP and CC serves as a simple model for studying interprotein ET.

CCP catalyses the reduction of alkyl hydroperoxides, during which it undergoes a two electron oxidation by peroxide to form an intermediate containing

an oxyferryl heme [Fe (IV)] and a free radical Trp 191 side chain (Sivaraja *et al.*, 1989). This intermediate is reduced in two consecutive one-electron steps by two ferroCC molecules. As early as 1980, a theoretical model of the CCP:CC ET complex was constructed, based on known crystal structures using computer modelling techniques (Poulos & Kraut, 1980). The basis of the modelling was a ring of positively charged surface residues surrounding the CC heme crevice, and a spatially complementary ring of negatively charged surface residues on the CCP molecule. Therefore, in the model, the forces holding the complex together are primarily electrostatic, involving four salt bridges between the complementary regions mentioned above. The hemes are co-planar with a distance of 25 Å between the iron atoms, and a closest heme edge-to-heme edge distance of 17 Å. These distances correlate well with distances derived from NMR and fluorescence studies (Gupta & Yonetani, 1973; Leonard & Yonetani, 1974). The model is also consistent with kinetic and cross-linking experiments (Bisson & Capaldi, 1981; Mochan & Nicholls, 1971).

In 1992 the structure of the CC:CCP complex was solved by X-ray crystallography, as shown in Figure 1.4, which allows a direct comparison with the theoretical model (Pelletier & Kraut, 1992). The actual structures solved were 1:1 complexes between CCP and two different cytochromes, yeast iso-1-cytochrome *c* (CCY) and horse heart cytochrome *c* (CCH). In the crystal structures the hemes are not co-planar, but inclined at an angle of about 60° to each other. The distance between the iron atoms is 26 Å, which is in good agreement with the model distance. However, the nature of the interaction between the proteins is different. The CCP:CCY complex (crystals grown at 150 mM NaCl) is held together predominantly by hydrophobic and van der Waals interactions, with only one hydrogen bond involved. The CCP:CCH complex (crystals grown at 0 mM NaCl) has three intermolecular hydrogen bonds, two of which are salt bridges (one of which was predicted in the modelling work).

Some degree of structural validation between the CCP:CCY crystal structure and the solution structure was obtained by site-directed mutagenesis studies (Miller *et al.*, 1994). Mutations that neutralised the charges on charged surface residues of CCP were introduced, in and around the binding region

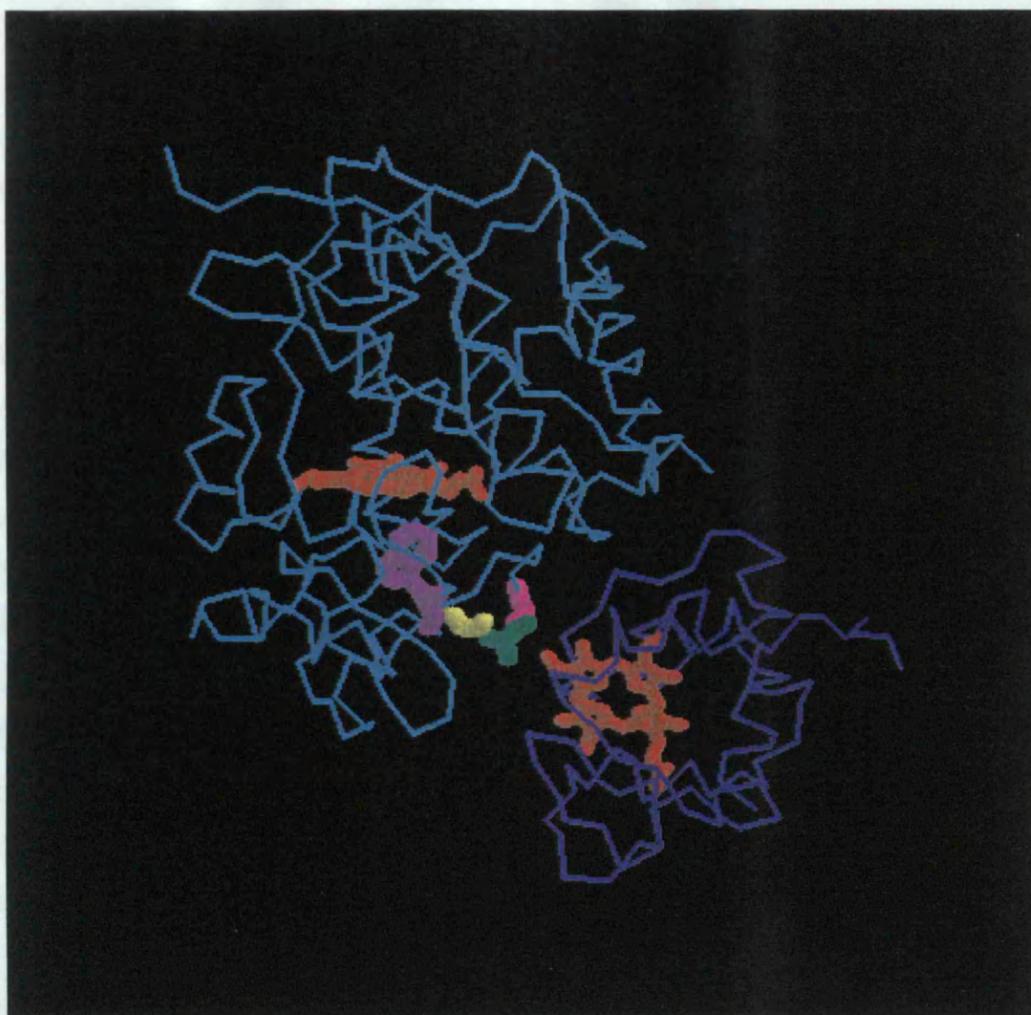


Figure 1.4. Structure of the CCP:CCY complex. Proteins are shown in backbone representations. Colours assigned as follows: cyan for CCP, blue for CCY. The heme groups of both proteins are shown in red. The proposed electron pathway from the heme methyl group of CC through the CCP residues Ala 194 (magenta), Ala 193 (green), Gly 192 (yellow) and Trp 191 (purple) is shown. Figure produced using PDB file 2PCC.

identified in the crystal structure (E32Q, D34N, E35Q, E290N and E291Q). Additionally, a large bulky residue was introduced into the centre of the binding region with the mutant A193F. Those proteins with mutations of residues directly associated with the binding region (Asp 34, Glu 290 and Ala 193) exhibited decreased ET rates by a factor of 2-4. Those proteins with mutations of residues outside the binding region (Asp 32 and Glu 291) had unaltered ET rates. The mutation of Glu 35, which is partially associated with the binding region, gave an ET rate intermediate compared to those seen for the other mutations. These results

indicate that the binding region observed in the crystal structure corresponds to the binding region seen in solution for the ET reactions. These results also suggest that electrostatic interactions are more important in this system than the crystal structures would suggest.

Using the crystal structure of CCP:CCY, an ET pathway was postulated. This extends from the CC methyl group of the heme, through the CCP residues Ala 194, Ala 193, Gly 192 and Trp 191 to the CCP heme (which is in van der Waals contact with the indole ring of Trp191) (see Figure 1.5). This correlates with data that suggest Trp 191 is involved in the ET reaction. Trp 191 had previously been identified as the site of a rare stable amino acid radical, in the fully oxidised CCP, by electron nuclear double resonance spectroscopy (Sivaraja *et al.*, 1989). Mutation of Trp 191 to phenylalanine resulted in a significant decrease (1000-5000 fold) in the ability of CCP to catalyse the H₂O₂-dependent oxidation of ferrocycochrome *c* (Mauro *et al.*, 1988). Additionally, laser flash photolysis studies using ruthenium labelled CC's demonstrated that ET occurs from the CC heme directly to the radical site of the CCP (namely Trp 191), and not to the oxyferryl heme [Fe (IV)] (Geren *et al.*, 1991; Hahm *et al.*, 1992).

There is a slight difference between the two complexes (CCP:CCY and CCP:CCH) with respect to the relative rotation of the molecules. This rotation results in the distance between Ala 194, of the proposed ET pathway, and the CC heme being smaller in CCP:CCY compared to CCP:CCH (see Figure 1.5). The difference in the complex structures is probably due to the different ionic strengths at which the crystals were grown. CCP:CCH was grown at a much lower ionic strength. At a higher ionic strength it is thought that several hydrogen bonds in the CCP:CCH complex will break allowing the complex to adopt a structure similar to CCP:CCY (unfortunately it has not been possible to grow crystals at the same ionic strength to test this). The existence of two binding geometries is consistent with the biphasic nature of the kinetics observed in studies of ruthenated CCH and CCP at different ionic strengths (biphasic kinetics were not observed in similar studies of CCY) (Geren *et al.*, 1991; Hahm *et al.*, 1992). Whether both of these geometries are ET active is unknown, but it is thought that the binding orientation responsible for the slow phase is not ET active, and the observed ET is conformationally gated by the rate of conversion to the ET active conformation (in

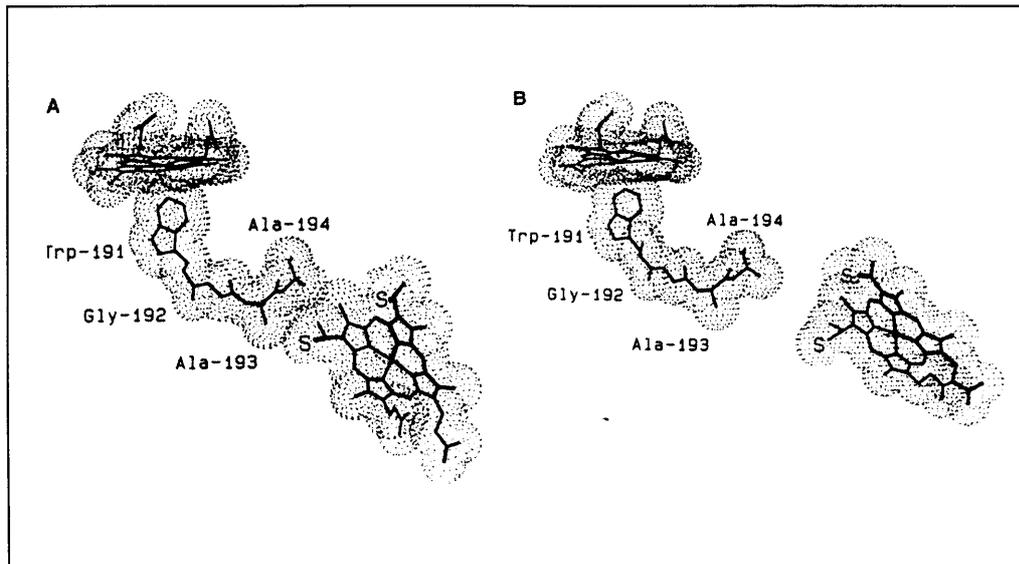


Figure 1.5. The proposed ET pathway in the CCP:CC complexes. The CC heme is shown on the right of each diagram. The cysteine atoms (S), covalently bonded to the CC hemes, are included. Panel A; CCP:CCY complex. Panel B; CCP:CCH complex (Taken from Pelletier and Kraut, 1992).

effect the CCP:CCH conformation going to a CCP:CCY type conformation) (Hahm *et al.*, 1992). The idea that the CCP:CCY complex is ET active is supported by a recent thermodynamic study which indicates that the crystal and solution structures for the CCP:CCY complex are the same, while those for the CCP:CCH complex are different (Wang & Pielak, 1999).

The biphasic nature of the ET reaction between CCP and CCH implies dynamics within the complex not revealed by the crystal structure (or the modelled structure), which gives a static picture. A dynamic view of the interaction was provided by the use of computer simulations using Brownian dynamics (Northrup *et al.*, 1988). The starting point for these simulations was crystal structures of CCP and CC; CCP was chosen as the central protein and CC was positioned randomly at a distance of 65 Å from CCP. The geometric criteria between CCP and CC allowing successful ET was set at a heme edge-to-heme edge distance of 20 Å, with the heme planes parallel to within 60°. These simulations indicate that successful ET criteria were achieved in a large number of electrostatically stable orientations, and not a single specific orientation, as suggested by the crystal structure. This ensemble of orientations form in three distinct regions around CCP,

all of which are approximately in the heme plane (see Figure 1.6, panel A). These regions coincide with the electrostatically attractive regions observed in that plane (see Figure 1.6, panel B). The region most highly populated was that around Asp 34 of CCP, as observed in the theoretical model (Poulos & Kraut, 1980), and which forms a salt bridge in the crystal structure (Pelletier & Kraut, 1992). The second region is around Asp 148, and the third less populated region is between the two dominant areas around Asp 217. Analysis of the contacts between CCP and CC in 243 docked complexes revealed an even distribution of ionic contacts, and a multitude of complexes with slightly different orientations were possible. This suggests that charge complementarity does not lock the protein complex into a unique orientation as seen in the crystal structure, and highlights the importance of protein dynamics in complex formation and electron transfer. Experimental support for the findings of Northrup *et al.* (1988) was provided by fluorescence quenching experiments (Zhang *et al.*, 1990). These experiments indicate that the complex exists in a number of conformations, which can be trapped at low temperature, but equilibrate rapidly at room temperature.

Several other studies have also indicated the importance of conformational realignments within the encounter complex. The rate of solvent exchange (in D₂O) of the backbone amide protons was measured for both the complex and the individual proteins. The rate of solvent exchange is an indicator of solvent exposure; therefore amides protected by involvement in the protein-protein interface can be identified. The NMR studies demonstrated that all the amides that would be protected in the crystal structures are protected, and several residues lying outside the interface are also protected. This implies that the interaction is not described by a single conformation (Jeng *et al.*, 1994; Yi *et al.*, 1994). Other studies supporting conformational changes within the complex are kinetic analysis using Zn-substituted CCP and ionic strength dependence effects (Hazzard *et al.*, 1988; Nocek *et al.*, 1990).

The observation of Northrup *et al.* (1988) that three regions can accommodate CC has been interpreted as evidence for multiple binding sites for CC on CCP. This is interesting, as the binding stoichiometry of the complex is open to debate. A binding stoichiometry of 1:1 is supported by spectrophotometric analysis (Erman & Vitello, 1980), a sedimentation equilibrium study (Dowe *et al.*,

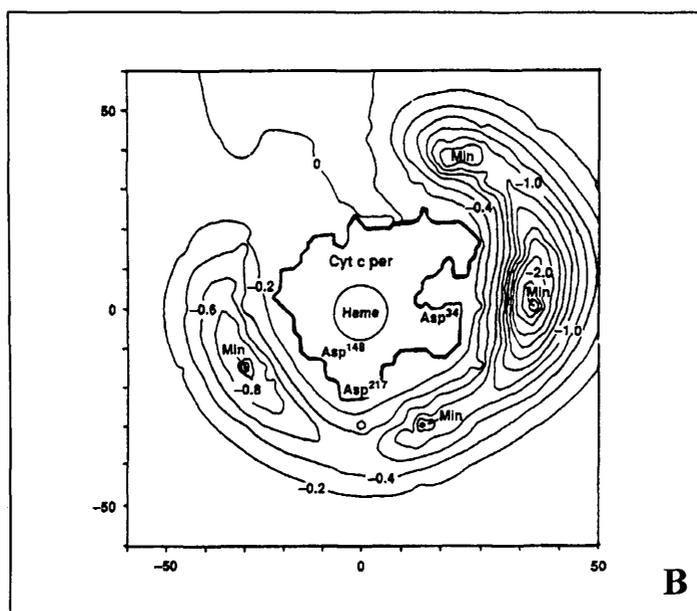
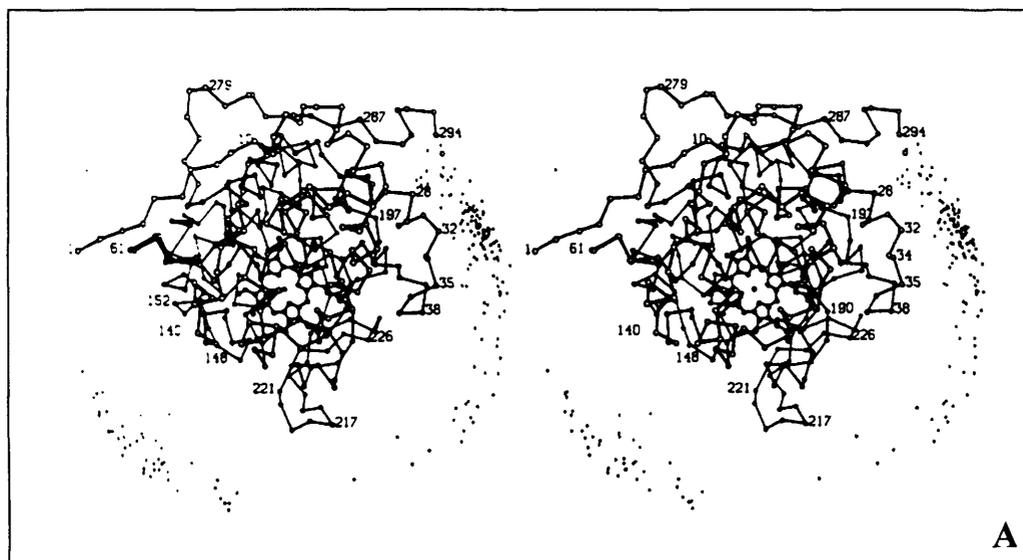


Figure 1.6. Brownian simulations and electrostatic potentials. Panel A; Brownian simulation of CCP:CC association. CCP is represented as an α -carbon skeleton and the heme atoms are shown. Points represent the centre of mass of CC when criteria for ET are successfully met. Panel B; Boltzmann-averaged total electrostatic potential energy of interaction between CCP and CC, in units of $k_B T$, as a function of the centre of mass of CC. The cross section of the potential energy taken in the heme plane of CCP is in the same orientation as the molecules shown in Panel A (Taken from Northrup *et al.*, 1988).

1984), kinetic analysis (Kim *et al.*, 1990) and the crystal structures (Pelletier & Kraut, 1992). However, there is also data that suggests CCP has two separate CC binding sites at low ionic strengths, giving binding stoichiometry of 2:1. This data includes steady-state kinetic analysis (Kang *et al.*, 1977), potentiometric studies (Mauk *et al.*, 1994) and studies of ruthenium-labelled CC (Geren *et al.*, 1991; Hahm *et al.*, 1992). Kinetics performed on Zn-substituted CC (ZnCC), which allows ET to be initiated by photo-excitation of the Zn-heme to form the excited triplet state, indicated two non-overlapping binding sites; one with high affinity and low ET coupling, and the other with low affinity and high ET coupling (Stemp & Hoffman, 1993; Zhou & Hoffman, 1993). This has been confirmed by the use of redox inert Cu-substituted CC (CuCC) as a competitive inhibitor, which allowed the ET rates of the high and low affinity binding sites to be probed (Zhou *et al.*, 1995). The ET rate for the low affinity site is approximately 1000-fold greater than the high affinity site. Finally, a study using a reverse Stern-Volmer quenching technique, which utilised the photo-inducible ZnCC, demonstrated that the 2:1 stoichiometry seen at low ionic strengths endures in physiological conditions (Zhou & Hoffman, 1994). Therefore, it seems likely that CCP contains a second binding site for CC at low ionic strengths, although whether it is physiologically significant is open to debate.

The CCP:CC system, like the MADH system, represents one of the best structurally characterised systems to date. The X-ray crystal structures for the CCP:CCY and CCP:CCH complexes have been solved and show good correlation with data from solution studies i.e. site-directed mutagenesis studies. As with the MADH system, the crystal structures do not indicate the dynamic nature of the ET complex, which has been demonstrated through the use of Brownian dynamics, NMR and kinetic analysis. Additionally, with the MADH system there may be some masking of electrostatic interaction in the crystal structure. However, in the the CCP:CC system there is no uncertainty owing to the conditions used for crystal growth, as the CCP:CCY and CCP:CCH crystals were grown under different conditions but have very similar structures.

1.3.3 Cytochrome *f*-plastocyanin system

The Cytochrome *f*-Plastocyanin system is not as well characterised as the systems described previously (see Sections 1.3.1 and 1.3.2). There is no crystal structure for the protein complex. However, a structure has been produced using a novel NMR based approach (Ubbink *et al.*, 1998). Additionally, information from a variety of techniques has provided valuable insight into the ET reaction.

Cytochrome *f*, at 31 kDa is the largest of the four polypeptides that make up the chloroplast cytochrome *b₆f* complex (the others being cytochrome *b₆*, the [2Fe-2S] Rieske protein, and the cofactor-free subunit IV), and acts as the electron acceptor for the complex (Gray, 1992). The complex contains one copy of each of the polypeptides and is an integral membrane protein of the thylakoid membrane of chloroplasts (Willey *et al.*, 1984). The structure of a truncated cytochrome *f* from turnip, a member of the *Cruciferae* family, has been solved by X-ray crystallography (Martinez *et al.*, 1994). Many studies have utilised cytochrome *f* from the *Cruciferae* family, as these cytochrome *f*'s are found to be soluble and monomeric. This is due to cleavage of the 33 amino acids, which form the transmembrane and carboxy-terminal regions of the protein. The soluble domain is organised into two domains, one large heme-binding domain and one small domain, and is predominantly of a β -sheet structure. The heme binding is unusual, with the amino group of the amino-terminal residue (Tyr 1) involved as a ligand for the heme iron; ligation of this type was previously unprecedented.

Plastocyanin is a type I blue copper protein (10.5 kDa) containing one copper atom per molecule. It serves as an electron carrier between the membrane bound cytochrome *b₆f* and P-700 in the chloroplasts of higher plants as part of the photosynthetic ET chain. The structure of the protein has been determined by X-ray crystallography (Guss *et al.*, 1992) and NMR methods (Moore *et al.*, 1988), indicating a β -barrel structure. The copper atom is co-ordinated to His 37, Cys 84, His 87 and Met 92 in a hydrophobic region. Studies with small molecules, such as $[\text{Cr}(\text{NH}_3)_6]^{3+}$, $[\text{Fe}(\text{CN})_6]^{3-}$ and $[\text{Co}(\text{NH}_3)_6]^{3+}$, have identified two reaction sites on plastocyanin; one close to the copper ligand His 87 (described as the adjacent site) in the hydrophobic region and the other close to Tyr 83 (described as the remote site) in the acidic region (Armstrong *et al.*, 1986; Pladziewicz & Brenner, 1987).

These sites are 6 Å and 10-12 Å from the copper atom, respectively. The residues Tyr 83 and His 87, and the acidic and hydrophobic regions they are associated with are conserved in plastocyanins from different species.

Interestingly a theoretical study using the *Pathways* method of Beratan and Onuchic, and laser flash photolysis techniques have demonstrated that the electronic coupling to both sites is approximately equal (Betts *et al.*, 1992; Qin & Kostic, 1996; Ullmann & Kostic, 1995). This is probably due to the high β -sheet content of plastocyanin, which provides effective electronic coupling between the remote site and the copper atom (Langen *et al.*, 1995). Also, a study of the metal-ligand covalency in plastocyanin has demonstrated that the electronic structure of the copper atom favours ET through its cysteinyl ligand, which is part of the proposed ET pathway for the remote site (Lowery *et al.*, 1993). These factors contribute to making both sites equally "attractive", for ET, as each other, notwithstanding the fact that one site's ET pathway is 6 Å longer.

Therefore, many of the studies of plastocyanin have centred on whether the protein utilises both sites for ET reactions, or favours either one. Mutation of the surface-exposed Tyr 83 to phenylalanine or leucine, results in a substantial reduction of ET rates (He *et al.*, 1991), indicating that Tyr 83 is involved in both the ET pathway, and binding to the cytochrome *f*. Chemical cross-linking identified a binding site for plastocyanin on cytochrome *f*, with cross-linking observed between Asp 44 of plastocyanin and Lys 187 of cytochrome *f* (Morand *et al.*, 1989). Asp 44 is a component of the acidic region, associated with the remote reaction site for plastocyanin; Lys 187 is associated with a basic region on cytochrome *f*. Also, the reaction between plastocyanin and cytochrome *f* is sensitive to ionic strength (BeokuBetts *et al.*, 1985; Meyer *et al.*, 1993; Qin & Kostic, 1992). Thus, there is evidence that plastocyanin and cytochrome *f* interact via the attraction of oppositely charged regions on the surface of the proteins, indicating that plastocyanin utilises the remote site, around Tyr 83, for the ET reaction with cytochrome *f*.

However, the structure of the complex, solved using a novel NMR based method, contradicts this. Using chemical shift changes and intermolecular pseudocontact shifts in the NMR spectrum of plastocyanin as input, rigid-body

molecular dynamics were used to generate the complex structure (Ubbink *et al.*, 1998). In this NMR spectra were recorded for free plastocyanin, cytochrome *f*(II)-plastocyanin and cytochrome *f*(III)-plastocyanin (with the plastocyanin being ^{15}N labelled). This enabled the identification of the plastocyanin residues that are involved in the protein-protein interface, through simple chemical changes between the free and bound plastocyanin. While pseudocontact shifts (shifts due to the paramagnetic Fe (III)) were used to determine the orientation and distance of a number of plastocyanin nuclei relative to the haem. The structure of both proteins were treated as rigid-bodies, with the position of the atoms fixed within the molecule. The position of cytochrome *f* was fixed and the plastocyanin allowed to move, relative to the cytochrome *f*, in order to search for an orientation that best fitted the restraints. In the complex both the hydrophobic and acidic patch of plastocyanin are in contact with the cytochrome *f* surface (see Figure 1.7). A very short ET pathway is present from the heme to the copper atom, as the copper ligand His 87 makes van der Waals contact with the cytochrome *f* residues Tyr 1 and Phe 4; the former is a ligand for the heme iron. Therefore, this work strongly suggests that ET proceeds via His 87, which is also in agreement with modelling studies (Pearson *et al.*, 1996; Ullmann *et al.*, 1997).

The importance of protein dynamics has been demonstrated by several experiments. Cytochrome *f* and plastocyanin were non-invasively cross-linked so that the exposed heme edge is in contact with the remote site (Qin & Kostic, 1993). The result of the cross-linking was the loss of ET; this was determined by following the ^1H NMR line widths of the heme methyl group. This result is interpreted as a need for conformational fluctuations within the complex, which are impeded by cross-linking. A study of the ionic strength dependence of the ET between plastocyanin and various *c*-type cytochromes has shown that as the ionic strength increases from 5-40 mM the observed rate constant also increases; above 40 mM, the rate constant drops with increased ionic strength (Meyer *et al.*, 1993). This indicates that the most stable complex is not optimally configured for ET, although how relevant this data is, is open to discussion since the physiological ionic strength is approximately 150 mM. Also, the photoinduced ET between Zn-Cytochrome *c* and plastocyanin has been shown to be conformationally gated by a viscosity study (Zhou & Kostic, 1992). Therefore, this work suggests that initially

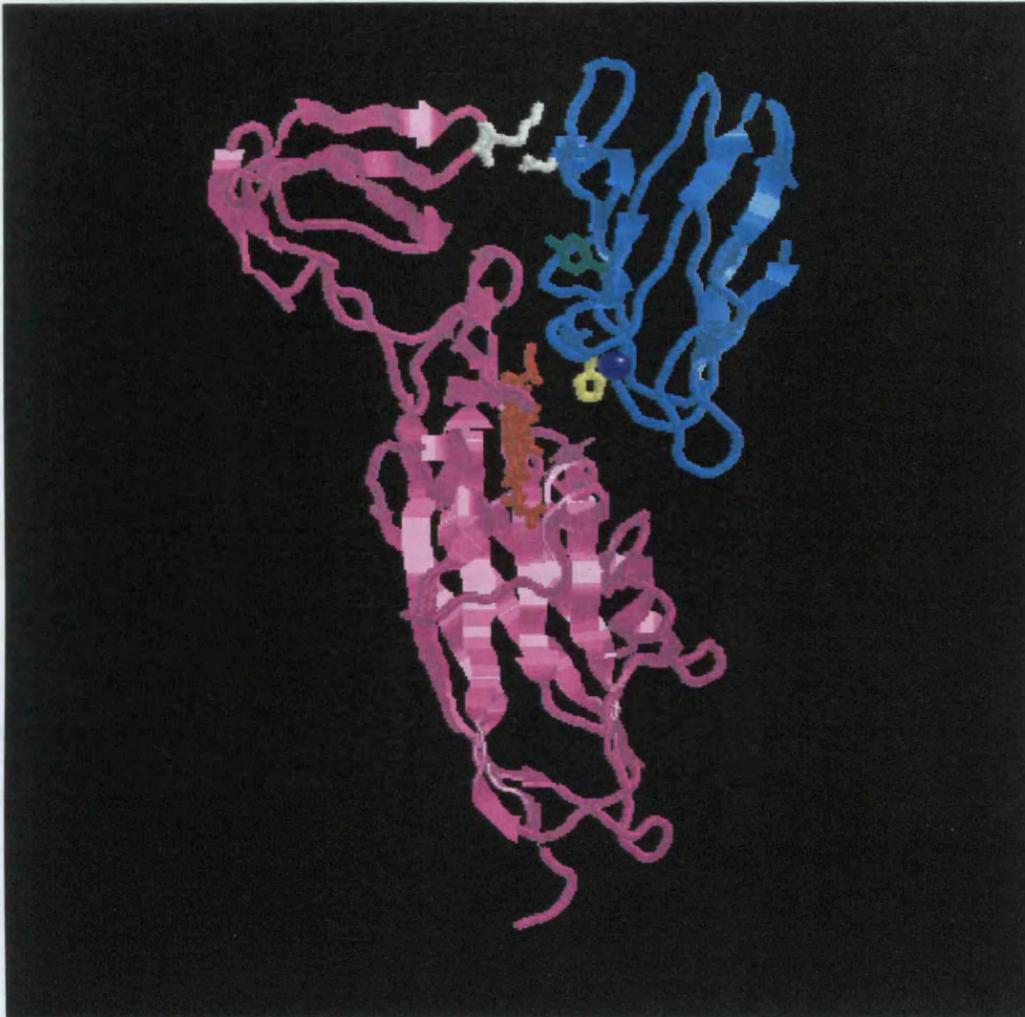


Figure 1.7. Structure of the cytochrome *f*-plastocyanin complex. The cytochrome *f* molecule is shown in magenta with the small domain at the top, and the large heme-binding domain at the bottom. The heme is shown in red. The plastocyanin is shown in cyan. The copper is shown in blue, and the side chains of Tyr 83 and His 87 are shown in green and yellow, respectively. The side chains of Lys 187 (cytochrome *f*) and Asp 44 (plastocyanin), identified in the cross-linking experiment of Morand *et al.* (1989), are shown in white. Figure produced using PDB file 2PCF.

the proteins come together via electrostatic interactions, and a set of complexes, with multiple orientations, coexists. This is followed by the formation of the functional ET complex in which hydrogen bonds, van der Waals contacts, hydrophobic interactions and electrostatic interactions contribute to form a single-orientation complex. The NMR complex structure is thought to represent this

single-orientation complex as the structure has a RMSD of 1.0 Å, relative to the cytochrome *f*.

The interaction between plastocyanin and P-700 has been probed by site-directed mutagenesis, indicating that plastocyanin interacts with P-700 via the adjacent site (Haehnel *et al.*, 1994; Nordling *et al.*, 1991). These studies also highlighted the importance of Tyr 83; mutation of Tyr 83 to leucine resulted in a 2-fold increase in ET rate. Therefore, the fact that the ET reaction is hindered by Tyr 83 relative to leucine, and that Tyr 83 is conserved in plastocyanins, suggests an advantageous property is imparted by Tyr 83 in a reaction distinct from that with P-700. Thus, Tyr 83 may play a prominent role in the initial complex formation between plastocyanin and cytochrome *f*, this would account for the findings of He *et al.*, (1991).

The study of the plastocyanin and cytochrome *f* system has been aided by the availability of a complex structure. This structure has allowed the identification of the site of interaction of plastocyanin on cytochrome *f*. Additionally, the importance of protein dynamics and electrostatic interactions has also been demonstrated through a variety of techniques.

1.4 Ideal systems for studying biological electron transfer

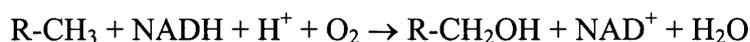
The model systems discussed previously (Section 1.3) have contributed greatly to our understanding of biological ET reactions. However, they are not without their problems and limitations. In fact no appropriate redox system has been developed that lends itself to a comprehensive and rigorous analysis of interprotein ET. With this in mind Scrutton and co-workers have begun to develop the soluble redox partners rubredoxin reductase-rubredoxin, of the alkane hydroxylase system, as a model system for investigation interprotein ET (Lee *et al.*, 1998; Lee *et al.*, 1996; Lee *et al.*, 1997).

This system is physiological and the individual components have all been cloned, overexpressed and thoroughly characterised (Lee *et al.*, 1998; Lee *et al.*, 1997; Shanklin *et al.*, 1997). The ET reaction between rubredoxin reductase and rubredoxin has been characterised kinetically (Lee *et al.*, 1998). As the redox active iron of the rubredoxin can be substituted with different metals, this will

allow the investigation of driving force dependence. However, a major drawback of this system is the lack of structural information for either the rubredoxin or rubredoxin reductase (or the alkane hydroxylase). Detailed structural information is considered to be essential in the study of interprotein ET. Structures can be used to study distance dependence and for ET pathway analysis. Additionally, structural information gives a reliable starting point for the design of mutants. The following section is an overview of the current knowledge of the alkane hydroxylase system, and its components.

1.5 The two iron rubredoxin of *Pseudomonas oleovorans*

The two iron rubredoxin of *Pseudomonas oleovorans* is part of the alkane hydroxylase system. The rubredoxin and rubredoxin reductase form a physiological ET complex that donates electrons to a membrane-bound alkane hydroxylase. This redox chain is responsible for the hydroxylation of aliphatic hydrocarbons, functionalised hydrocarbons, and various aromatic compounds



Rubredoxin reductase is a NADH-dependent flavoprotein that transfers reducing equivalents from NADH to the novel 2Fe rubredoxin. The hydroxylase component receives electrons from reduced rubredoxin and is a member of the non-haem diiron family.

1.5.1 Alkane hydroxylase system

During the research of petroleum microbiologists a *Pseudomonad* able to grow on the short chain alkane n-hexane as the sole carbon source was isolated from soil samples (Baptist *et al.*, 1963); this was notable because many organisms find n-hexane toxic. The *Pseudomonad* was identified as *Pseudomonas oleovorans*, previously isolated from industrial cutting oil (Lee & Chandler, 1941). Further investigations demonstrated that cell-free extracts were able to perform

alkane oxidation of octane to n-octanol and octanoic acid, and that pyridine nucleotides were required as co-factors (Gholson *et al.*, 1963).

The system enabling *Pseudomonas oleovorans* to utilise alkanes was initially termed the ω -hydroxylation system from its ability to produce the ω -hydroxyl forms of octanoate, decanoate and laurate (Kusunose *et al.*, 1964); it is now commonly referred to as the alkane hydroxylase system. Although fundamentally the alkane hydroxylase system consists of three components (soluble rubredoxin and rubredoxin reductase and a membrane bound alkane hydroxylase; Peterson *et al.*, 1966), it is part of a complex system of proteins encoded by genes on the transmissible extrachromosomal OCT-plasmid (Chakrabarty *et al.*, 1973). The OCT-plasmid carries the *alk* regulon, which consists of two gene clusters, *alk* BFGHJKL and *alk* ST (Fennewald *et al.*, 1979) (Eggink *et al.*, 1988); the regulon and a graphical representation of the gene products is shown in Figure 1.8.

The two gene clusters are sufficient to enable the host organism (*E. coli* or *Pseudomonas*) to utilise alkanes as the sole carbon source (Eggink *et al.*, 1987). The small gene cluster, *alk* ST, encodes two polypeptides, AlkS and AlkT. The AlkS protein has a regulatory role and is necessary for the activation and expression of the *Alk* BFGHJKL gene cluster (Eggink *et al.*, 1988), while the AlkT protein was found to be rubredoxin reductase (Eggink *et al.*, 1990). The *alk* BFGHJKL gene cluster encodes seven proteins. The genes *alk* F and G encode structurally related rubredoxins of approximately 15 and 19 kDa, respectively (Kok *et al.*, 1989a), with the 19 kDa rubredoxin being the active rubredoxin identified by Benson *et al.* (1971). Deletion experiments indicate that the 15 kDa rubredoxin is not active in transferring electrons within the alkane hydroxylase ET system, and it is thought that the *alk* F gene has arisen through a gene duplication event. The *alk* B gene encodes the membrane bound alkane hydroxylase (Kok *et al.*, 1989b), and the genes *alk* H, *alk* J, and *alk* K were found to encode aldehyde dehydrogenase, alcohol dehydrogenase, and acyl-CoA synthetase respectively (Kok *et al.*, 1989a; van Beilen *et al.*, 1992a). The function of the gene product *alk* L remains undetermined.

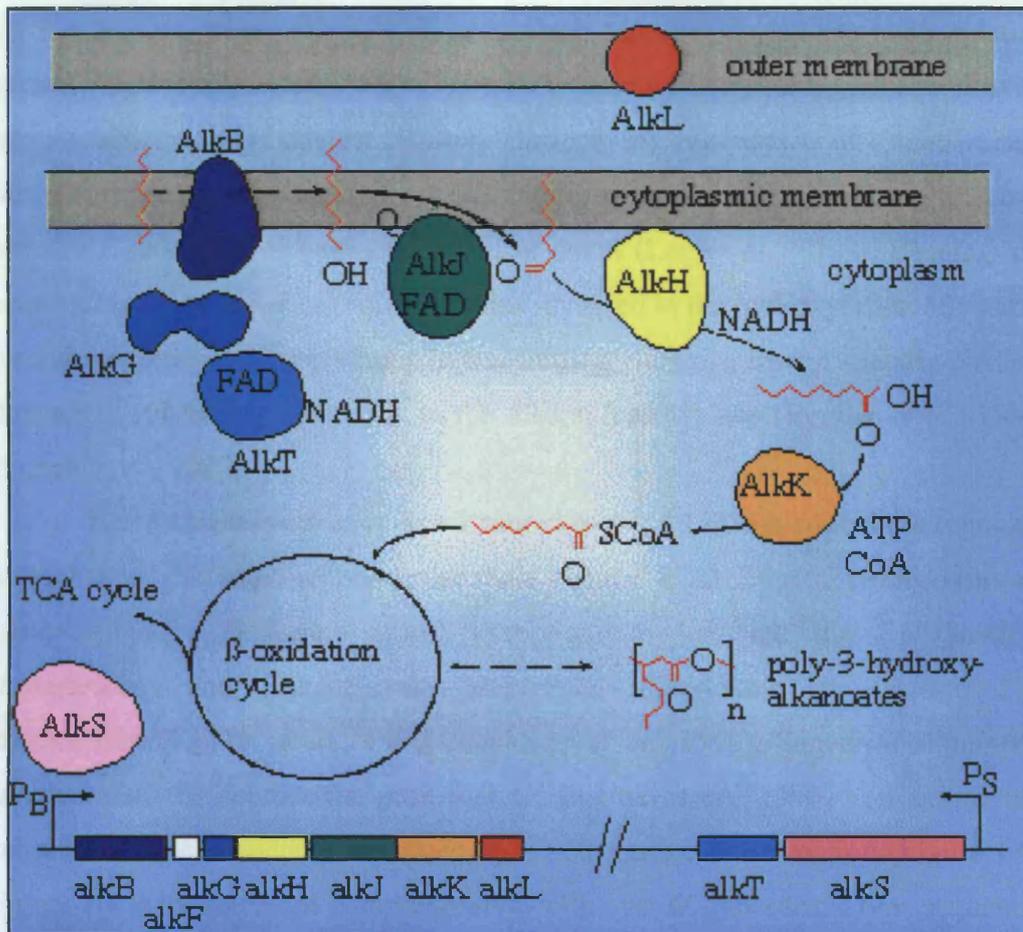


Figure 1.8. Genetics of the alkane hydroxylase system, and gene products.

1.5.2 Rubredoxin

The rubredoxins are a family of non-heme iron proteins whose active sites comprise an iron atom tetrahedrally coordinated to four cysteine residues (Sieker *et al.*, 1994). The first rubredoxin gene isolated was that from *Clostridium pasteurianum* (Lovenberg & Sobel, 1965). Recently a eukaryotic rubredoxin has been identified and characterized from the unicellular algae *Guillardia theta*, and found to be associated with photosystem II (Wastl *et al.*, 2000). The metabolic roles of rubredoxin in general are poorly understood. However, roles as a cofactor of a terminal oxidase in *Desulfovibrio gigas* and in hydrogen oxidation in *Azotobacter vinelandii* have been described (Chen & Mortenson, 1992; Gomes *et al.*, 1997). Rubredoxin is also used as an electron donor in the superoxide reductase electron transfer chain in anaerobic organisms (Jenney *et al.*, 1999). In

aerobic bacteria rubredoxin acts as an electron carrier in the alkane hydroxylation system (Geissdorfer *et al.*, 1995; Peterson & Coon, 1968). Although rubredoxins are presumed to function as electron carriers, the rubredoxin of *Pseudomonas oleovorans* is the only member of this family of proteins for which the electron transfer reactions have been delineated in detail (Lee *et al.*, 1998). Cloning and sequencing of the genes encoding proteins involved in the hydroxylation of alkanes established the role of rubredoxin in transferring electrons from a specific NADH-dependent rubredoxin reductase to the alkane hydroxylase (Eggink *et al.*, 1990; Eggink *et al.*, 1987).

The *Pseudomonas oleovorans* rubredoxin at 19 kDa is over three times the size of other rubredoxins (molecular mass about 6 kDa). Typical rubredoxins are comprised of ~55 amino acids; exceptions to this are the *Desulfovibrio desulfuricans* and *Pseudomonas oleovorans* rubredoxins, at 45 and 173, respectively (Eggink *et al.*, 1990; Stenkamp *et al.*, 1990). Sequence alignments indicate that the rubredoxin gene (*alk G*; Eggink *et al.*, 1990) consists of two rubredoxin type sequences separated by a ~60 residue linker region (Figure 1.9). Figure 1.9 is by no means comprehensive, with over 60 rubredoxin type sequences in the gene bank. However, the sequences shown are those of rubredoxins that have been characterised. The gene *alk G* is most likely the result of a gene duplication event. From the sequence alignment, six strictly conserved residues are apparent (residue positions 6, 9, 13, 39, 40, 42; using the top number scale shown in Figure 1.9), and there are several other positions that are well conserved. The iron binding motif of Cys-X-X-Cys-Gly is conserved in all the rubredoxins, with the exception of one motif in the *Guillardia theta* rubredoxin, and both in the N-terminal domain of the *Pseudomonas oleovorans* rubredoxin.

Rubredoxins have been studied extensively from a structural viewpoint. High resolution crystal structures are known for rubredoxins from *Desulfovibrio sulfuricans* (Stenkamp *et al.*, 1990), *Desulfovibrio gigas* (Frey *et al.*, 1987), *Desulfovibrio vulgaris* (Adman *et al.*, 1991), *Clostridium pasterianum* (Dauter *et al.*, 1996), and *Pyrococcus furiosus* (Day *et al.*, 1992). The solution structure for the *Pyrococcus furiosus* and *Clostridium pasterianum* rubredoxins have been determined by NMR methods (Bertini *et al.*, 1998; Blake *et al.*, 1992b). For the *Pyrococcus furiosus* rubredoxin the NMR structure was determined for the zinc

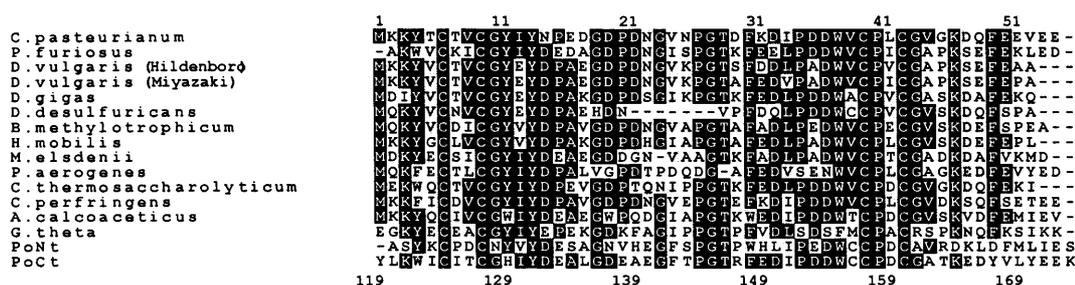


Figure 1.9. Sequence alignments of rubredoxins from different organisms. The top six sequences are from rubredoxins whose structures have been solved by X-ray crystallography. The sources of the data are: *Clostridium pasteurianum* (Dauter *et al.*, 1996); *Pyrococcus furiosus* (Bau *et al.*, 1998); *Desulfovibrio vulgaris* (strain Hildenborough)(Adman *et al.*, 1991); *Desulfovibrio vulgaris* (strain Miyazaki)(Misaki *et al.*, 1999); *Desulfovibrio gigas* (Frey *et al.*, 1987); *Desulfovibrio desulfuricans* (Stenkamp *et al.*, 1990); *Butyribacterium methylotrophicum* (Saeki *et al.*, 1989); *Heliobacillus mobilis* (Lee *et al.*, 1995); *Megasphaera elsdenii* (Bachmayer *et al.*, 1968b); *Peptococcus aerogenes* (Bachmayer *et al.*, 1968a); *Clostridium thermosaccharolyticum* (Meyer *et al.*, 1990); *Clostridium perfringens* (Seki *et al.*, 1989); *Acinetobacter calcoaceticus* (Geissdorfer *et al.*, 1995); *Guillardia theta* (Wastl *et al.*, 2000); PoNt and PoCt are the N- and C-terminal domain regions of the *Pseudomonas oleovorans* rubredoxin, respectively (Kok *et al.*, 1989a). Sequence alignment performed with ClustalW using the 'blosum' weight matrix, a gap penalty of 25 and a gap extension penalty of 0.05. Conserved amino acids in all sequences are shown in black, partially conserved amino acids are shaded grey. The numbering is as follows, the top numbers are for typical rubredoxins and PoNt, the bottom numbering is for PoCt.

substituted protein to avoid the relaxation problems associated with paramagnetic iron (Blake *et al.*, 1991), while the NMR structure for *Clostridium pasterianum* was solved for the native iron containing protein; however the paramagnetic iron makes such work more demanding (Banci *et al.*, 1994; Bertini *et al.*, 1998). While it is possible to solve NMR structures for paramagnetic proteins in the case of rubredoxins this is not always necessary. With rubredoxins it is relatively straightforward to substitute the iron with non-paramagnetic metals (Zn, Cd, and Hg; Blake *et al.*, 1994), and there is firm evidence that the overall structure of the protein is unaltered by this metal substitution (Blake *et al.*, 1992a; Blake *et al.*,

1994; Dauter *et al.*, 1996). Also, the rubredoxin structure is not sensitive to the oxidation state of the iron (Dauter *et al.*, 1996; Day *et al.*, 1992).

The secondary and tertiary structures of the rubredoxins are well conserved across the different organisms. The major structural characteristics are the tetrahedral coordination of four cysteine sulphur atoms ligating a single metal ion; a three stranded anti parallel β -sheet, typically consisting of the first 14 or 15 residues and residues 49-54; and a hydrophobic core usually containing six aromatic residues (positions 4, 11, 13, 30, 37, and 49) and an aliphatic residue at position 33 (Figure 1.10). The residues of the hydrophobic core are not totally conserved in the *Pseudomonas oleovorans* rubredoxin, with the N and C-terminals domain containing aspartate and histidine residues, respectively, in the place of one aromatic residue in other rubredoxins.

A major driving force for some of the structural work has been to investigate the structural contributions to thermostability. The structure of the rubredoxin from the hyperthermophilic bacteria *Pyrococcus furiosus*, which exhibits optimal growth at 100 °C, is unaffected by a 24 hour incubation at 95 °C (Blake *et al.*, 1991). In comparison the rubredoxins from mesophilic bacteria are less thermostable; *Clostridium pasteurianum* rubredoxin is rapidly denatured at 80 °C (Lovenberg & Sobel, 1965); *Desulfovibrio gigas* rubredoxin loses 50 % of its visible absorption after 2 hours at 80 °C (Papavassiliou & Hatchikian, 1985). Comparison of the structures suggested that a number of localised interactions in the β -sheet region conferred the extreme thermostability seen with the *Pyrococcus furiosus* rubredoxin. These include several hydrogen bonds and an extra salt-bridge, involving the unique residues Ala 2 and Glu 15, that is thought to inhibit "unzipping" of the β -sheet at high temperatures (Blake *et al.*, 1992b; Day *et al.*, 1992). However, introduction of the unique Ala 2 and Glu 15 residues by site-directed mutagenesis into the rubredoxin of *Clostridium pasteurianum* (Richie *et al.*, 1996), and the construction of *Pyrococcus furiosus*-*Clostridium pasteurianum* β -sheet "chimeras" (Eidsness *et al.*, 1997), have not significantly increased the thermostability of the *Clostridium pasteurianum* rubredoxin. It is thought that the thermostability of the *Pyrococcus furiosus* rubredoxin is due to a global alignment which optimises both main chain and side chain interactions between β -strands and

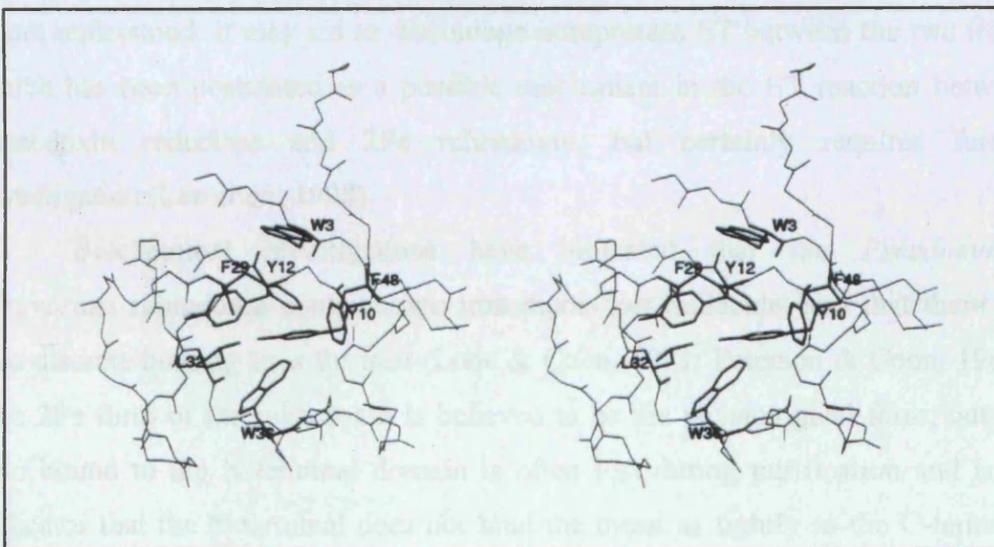
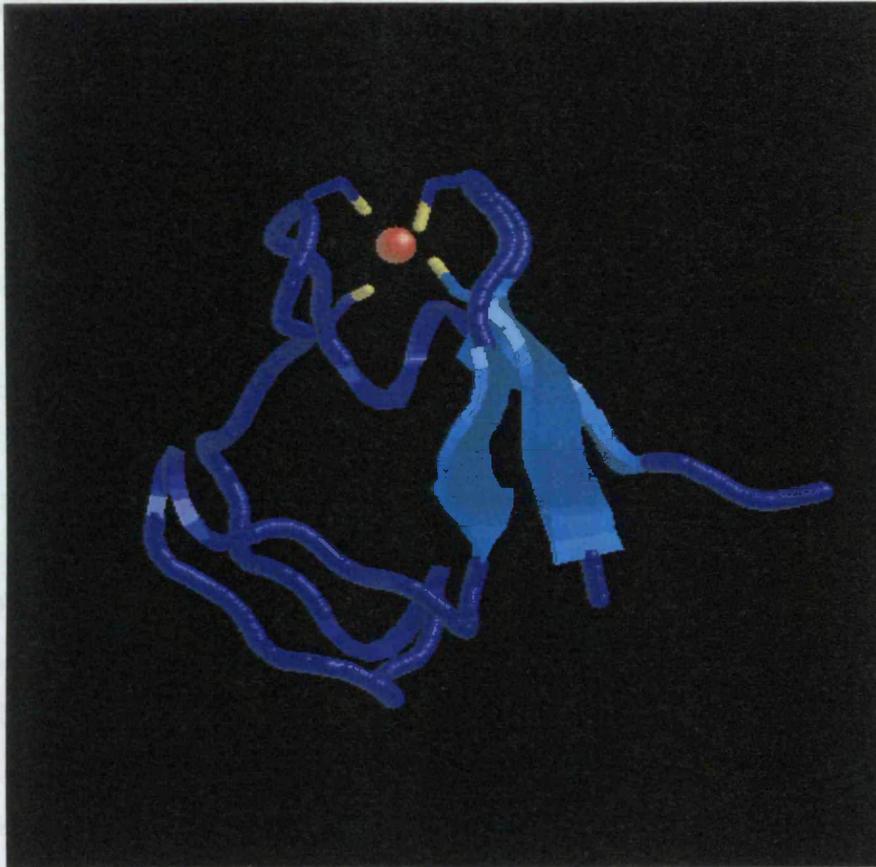


Figure 1.10. Structure of rubredoxin. Upper panel, the rubredoxin of *Pyrococcus furiosus*, the β -sheets are shown in cyan, the iron is shown in red, and the cysteine ligands are shown as wireframes with the sulphur atoms depicted in yellow (PDB code 1brf). Lower panel, the hydrophobic core of the *Pyrococcus furiosus* rubredoxin (taken from Day *et al.*, 1992).

core residues, which are more important than a few localised interactions within the β -sheet (Eidsness *et al.*, 1997).

Several studies have investigated the reduction potentials and electron self exchange (ese) constants of rubredoxins by site directed mutagenesis. The k_{ese} value is a measure of the exchange of the electron between the Fe (II) and Fe (III) oxidation states in a mixture of oxidised and reduced protein, and gives an indication of the reactivity of a molecule as an ET agent (NMR is used to determine this value; Gaillard *et al.*, 1996). Mutants of the *Clostridium pasterianum* rubredoxin in which the relatively well conserved residues at positions 8, 10, 11, and 41 were converted to a range of different residue types, all showed significantly altered reduction potentials (Kummerle *et al.*, 1997; Zeng *et al.*, 1996). However, the k_{ese} did not vary greatly except when a charged residue was placed at positions 8 or 10. Understanding the factors governing the reduction potential and k_{ese} is thought to be a key factor in understanding the ET reactions of rubredoxins (Kummerle *et al.*, 1997). Interestingly the N-terminal domain of the *Pseudomonas oleovorans* rubredoxin has charged residues at positions 8 and 10 (Asp and Asn, respectively). The exact contribution this makes to the ET reaction is not understood. It may aid or discourage intraprotein ET between the two irons, which has been postulated as a possible mechanism in the ET reaction between rubredoxin reductase and 2Fe rubredoxin, but certainly requires further investigation (Lee *et al.*, 1998).

Biochemical investigations have indicated that the *Pseudomonas oleovorans* rubredoxin contains two iron atoms per molecule, and that there are two discrete binding sites for iron (Lode & Coon, 1971; Peterson & Coon, 1968). The 2Fe form of the rubredoxin is believed to be the physiological form, but the iron bound to the N-terminal domain is often lost during purification and is an indicator that the N-terminal does not bind the metal as tightly as the C-terminal domain (Lee *et al.*, 1997). The 1Fe protein can be readily converted to the 2Fe form through a simple trichloroacetic acid precipitation technique. The oxidised form of the protein displays typical rubredoxin characteristics, the electronic absorption spectra has absorption maxima around 495, 378, and 280 nm (Lee *et al.*, 1997; Peterson & Coon, 1968). Further characterisation by CD and NMR

spectroscopies, including analysis of ^{113}Cd -substituted protein, were consistent with studies of other rubredoxins, and indicated that the Cd-substituted protein is isostructural with the iron forms (Blake *et al.*, 1994; Henehan *et al.*, 1993; Lee *et al.*, 1997). Cleavage of the 2Fe rubredoxin at residue Met-50 with cyanogen bromide yields a C-terminal peptide (residues 51 to 173) that can be folded around iron and cobalt, and the folded peptide can act as an electron carrier during steady-state hydroxylation reactions (Lode & Coon, 1971; May *et al.*, 1984). The corresponding N-terminal fragment (residues 1 to 50) can be transiently folded around iron, but is highly labile, and although it is active during steady-state hydroxylation reactions, this is at a much lower level compared to the C-terminal fragment (Lode & Coon, 1971). The sequence alignments indicate that this instability may be a result of the position of cleavage, which in other rubredoxins would occur within a β -sheet region (Eidsness *et al.*, 1997).

Detailed kinetic investigations of the ET reaction between rubredoxin and rubredoxin reductase have been performed (Lee *et al.*, 1998). This work indicated that both 1Fe and 2Fe forms of rubredoxin combine with rubredoxin reductase to form functional ET complexes, and that the reaction rates in each complex are comparable. The ET reaction is gated by a step preceding ET, and this adiabatic reaction is rate limiting for ET. The data also suggests that the iron containing N-terminal domain of rubredoxin plays little or no part in complex assembly, even though the N-terminal domain is reduced by rubredoxin reductase, either directly, or by rapid intraprotein ET from the C-terminal domain iron. This raises major questions over the role of the N-terminal domain of rubredoxin in both the recognition of its redox partners and ET reactions.

1.5.3 Rubredoxin reductase

Rubredoxin reductase (RR) is a flavoprotein of about 55kDa containing one mole of FAD per mole of enzyme (Ueda *et al.*, 1972). The gene encoding RR, *alk T*, has a coding capacity of ~41 kDa, which is considerably lower than the apparent molecular weight of the encoded protein (Eggink *et al.*, 1990). Sequence analysis reveals fingerprints for FAD binding sites near the N-terminus (positions 2-34) and the C-terminus (positions 265-275). Additionally, a fingerprint involved in the

binding of NAD(P)H is located in the middle of the sequence (positions 144-172). The sequence shows similarity with a number of oxidoreductases containing FAD and NAD(P)H nucleotides as cofactors, including the rubredoxin reductase from *Acinetobacter calcoaceticus* ADP1, glutathione reductases and dihydrolipomide dehydrogenases (Eggink *et al.*, 1990; Geissdorfer *et al.*, 1995; Geissdorfer *et al.*, 1999).

RR contains one mole of non-covalently bound FAD per mole of protein, and is able to accept two electron equivalents during titrations with NADH (Ueda & Coon, 1972). RR is able to transfer electrons to the rubredoxin of *Pseudomonas oleovorans*, and to the rubredoxins of the anaerobic bacteria *Peptostreptococcus elsdenii*, *Clostridium pasterianum* and *Desulfovibrio gigas*. It is also able to transfer electrons directly to ferricyanide and dichlorophenolindophenol. RR is, however, unable to transfer electrons to other non-haem iron proteins, namely ferredoxin, putaredoxin and adrenodoxin. RR also transfers electron indirectly to cytochrome *c* in the presence of rubredoxin. Measurements of the decrease in fluorescence of RR in the presence of varying amounts of rubredoxin demonstrated that a complex is formed between the rubredoxin and RR, with a ratio of 1:1 (Ueda & Coon, 1972).

The gene *alk T* has been cloned into the expression vector pKK223-3 and RR over-expressed in *E. coli* strain JM109. Purification of this protein and subsequent stopped-flow analysis has allowed characterisation of several electron transfer reactions within this system; NADH to the flavin of RR, and RR to 1Fe- and 2Fe-forms of di-domain rubredoxin (Lee *et al.*, 1998). This work indicated that RR is reduced by NADH in a simple one-step mechanism to an enzyme-NAD⁺ charge-transfer species. The data also suggest that the N-terminal domain of rubredoxin plays little or no part in complex assembly with RR, and that the ET reaction between these proteins is gated by a rate-limiting step prior to ET.

1.5.4 Alkane hydroxylase

The alkane hydroxylase (sometimes referred to as ω -hydroxylase) is a non-haem iron protein of about 41kDa (Mckenna & Coon, 1970). It catalyses the oxidation of a range of alkanes with one oxygen derived from molecular oxygen.

The second oxygen is reduced using electrons from the reduction of NADH via rubredoxin reductase and rubredoxin. The apo-alkane hydroxylase is catalytically inactive, but activity can be restored by the addition of ferrous ions (Ruettinger *et al.*, 1977). Immuno gold labelling and cell fractionation studies demonstrated that the enzyme is an integral membrane protein (Benson *et al.*, 1979; Lageveen, 1986), which is in agreement with experimental data indicating that the enzyme requires about 20 phospholipid molecules per alkane hydroxylase molecule for activity (Ruettinger *et al.*, 1974).

The gene encoding the alkane hydroxylase (*alk B*) has been sequenced. A hydrophilicity plot of the amino acid sequence shows nine hydrophobic sequences of which eight are long enough to span the membrane as α -helices (Kok *et al.*, 1989b). Studies of *alk B* fusions with alkaline phosphatase and β -galactosidase predicts a model for the alkane hydroxylase containing six transmembrane segments, with two hydrophilic loops and a large carboxyl-terminal domain located in the cytoplasm (van Beilen *et al.*, 1992b).

The alkane hydroxylase shows sequence similarity with the xylene monooxygenase of *Pseudomonas putida*. There is a conserved 8-histidine motif in both of these proteins, and in a number of eukaryotic fatty acid desaturases, indicating that these proteins may constitute a family of structurally related proteins (Shanklin, 1994). Indeed the work of van Beilen *et al.* (1992), with *alk B* fusion proteins supports a model in which these proteins share a common fold. Mössbauer spectra of purified alkane hydroxylase demonstrated that it contains a diiron cluster of the type found in soluble diiron proteins, such as methane monooxygenase, toluene-4-monooxygenase, stearyl-acyl carrier protein Δ^9 desaturase, ruberythrin and ribonucleotide reductase (Shanklin *et al.*, 1997). This work confirmed that the 8-histidine motif is a characteristic of a new family of diiron containing integral membrane protein, of which the alkane hydroxylase is a prime example (Fox *et al.*, 1993; Shanklin, 1994).

Diiron clusters are characterised by the presence of two Fe^{3+} within the active site, co-ordinated by histidine residues, and are thought to facilitate the formation of high energy reaction intermediates. The 8-histidine motif seen in alkane hydroxylase is present in about 70 integral membrane proteins which

catalyse O₂ dependent modification of hydrocarbon substrates, with at least 11 distinct activities (Shanklin *et al.*, 1997). The discovery that diiron clusters exist in both soluble and membrane proteins has greatly expanded the role of diiron centres in O₂ activation biochemistry, and the family as a whole is thought to have an enzymatic versatility approaching that of the P450 enzymes.

The substrate specificity of the alkane hydroxylase is relatively low as indicated by the broad substrate range shown in Figure 1.11, which is by no means comprehensive (van Beilen, 1994). The most preferred substrates are linear aliphatic alkanes. The introduction of one or two methyl substituents into the linear alkane slows, but does not prevent oxyfunctionalisation. However, tertiary carbon substituents are not substrates. Cyclic alkanes up to cyclohexane are substrates as are some alkenes and a range of aromatic compounds, including a heterocyclic compound (van Beilen *et al.*, 1994). Thus, the alkane hydroxylase has the potential to be used as a biocatalyst in the production of either pharmaceuticals or fine chemicals. PCR primers designed from the *alk B* sequence have been used to screen bacterial strains for related proteins, with an eye towards their biocatalytic properties (Smits *et al.*, 1999).

The *alk B* gene has successfully been expressed in a number of *E. coli* hosts, but of particular interest is *E. coli* strain W3110. In this strain the alkane hydroxylase is expressed up to an exceptional 15 % of the total cell protein (Nieboer *et al.*, 1993). This level of expression alters the membrane biogenesis, and the alkane hydroxylase is found in cytoplasmic vesicles. It has been possible to purify the protein as near homogeneous vesicle membranes containing high levels of the alkane hydroxylase (Shanklin *et al.*, 1997).

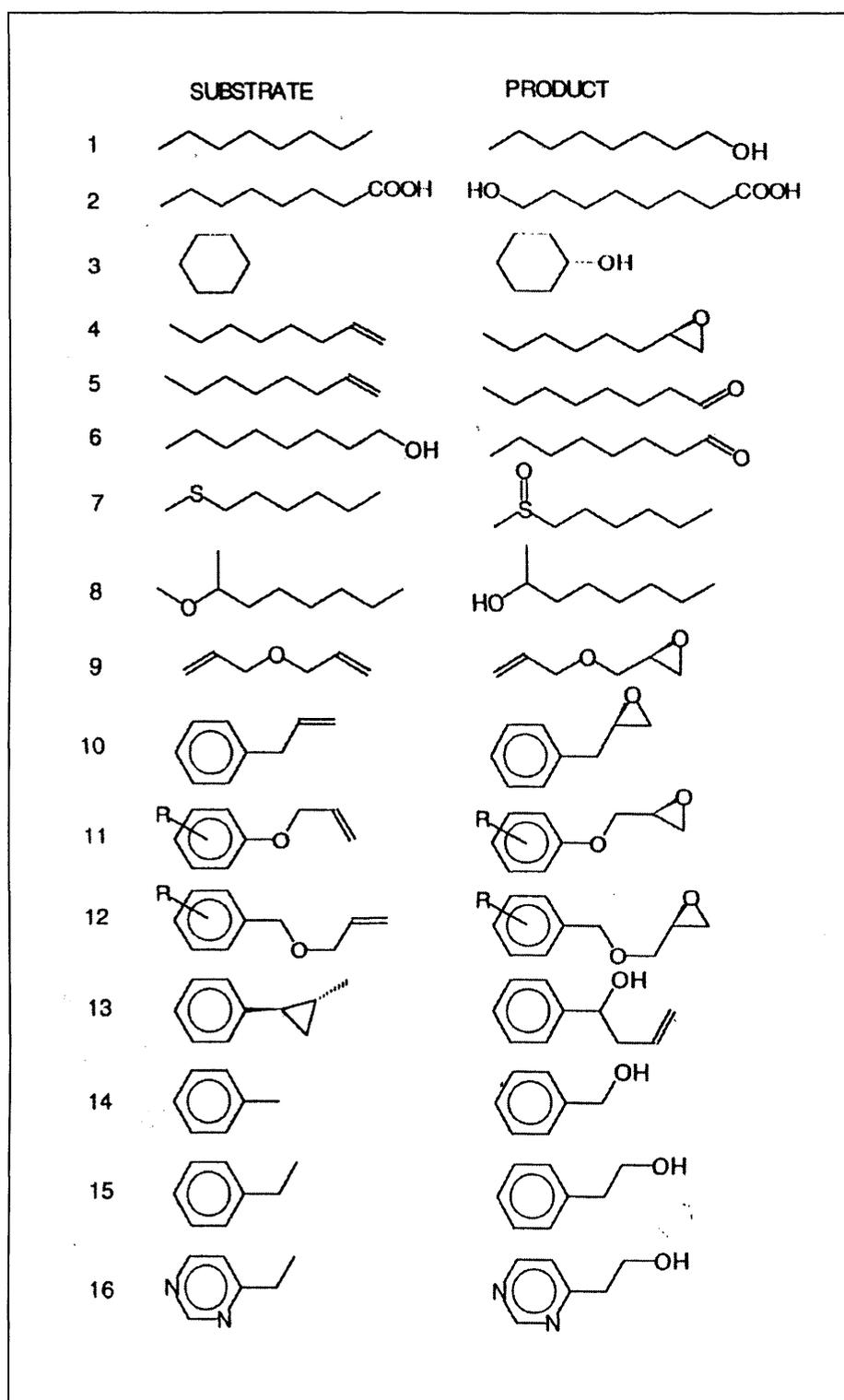


Figure 1.11. Substrate specificity of the alkane hydroxylase. Taken from van Beilen, 1994.

1.6 Aims of this Thesis

The aims of the work described in this thesis are to provide some structural characterisation of the two iron rubredoxin of *Pseudomonas oleovorans*. This rubredoxin is novel in its own right, and a component of a biologically interesting system with respect to the alkane hydroxylase system. Furthermore, the redox partners rubredoxin and rubredoxin reductase have potential as an ideal system for studying biological ET reactions, as do the redox partners rubredoxin and alkane hydroxylase, with respect to ET reactions involving a membrane bound protein. However, while this system possesses some of the characteristics required for detailed studies of biological ET, there is to date no structural information on any of the component proteins. Although, rubredoxins as a family are structurally well characterised there are some very important structural aspects of this rubredoxin that need to be addressed. With this in mind the work described in this thesis concentrates on the structural characterisation of the novel two iron rubredoxin of *Pseudomonas oleovorans*.

Chapter 2

Materials and Methods

2.1 Materials

2.1.1 Chemicals

Ampicillin, cadmium chloride and L-1-chloro-3-[tosylamido]-4-phenyl-2-butanone (TLCK) were supplied by Sigma. Tryptone, yeast extract and agar were supplied by OXOID. Protogel was supplied by National Diagnostics and Temed by Sigma. Sodium dodecyl sulfate (SDS) was from BDH chemicals. Agarose was from Life Technologies Inc. and ethidium bromide was supplied by Bachem. D₂O was supplied from Goss Scientific Instruments Ltd. Cd¹¹³ was supplied by Promochem Ltd.

All other chemicals were supplied by Sigma, Fisher or BDH and were of analytical grade where possible. Water used was glass-distilled and deionised.

2.1.2 Equipment

Thermal cycling was performed on a Techne PHC-3 thermal cycler. UV-visible spectra were recorded with a Hewlett Packard 8452a spectrophotometer or a Jasco V-550 spectrophotometer. CD-spectra were recorded with a Jasco J-715 spectropolarimeter. NMR spectra were recorded with a Bruker AMX-600 or AMX-500 spectrometer. X-ray solution scattering data were collected using the low angle scattering camera on station 2.1 at the Synchrotron Radiation Source (SRS), Daresbury, UK.

2.1.3 Molecular weight markers

DNA marker was from MBI Fermentas (band sizes were 10000, 8000, 6000, 5000, 4000, 3500, 3000, 2500, 2000, 1500, 1000, 750, 500 and 250 base pairs). Protein marker was constructed from stock solutions of trimethylamine dehydrogenase (83 kDa), bovine serum albumin (63 kDa), egg serum albumin (45 kDa), carbonic anhydrase (29 kDa) and lysozyme (14 kDa).

2.1.4 Bacterial strains and media

The rubredoxin expression clone (pKR10; Lee *et al.*, 1997) and *Ps.oleovorans* were obtained from Dr. Ho Joon Lee (University of Leicester). Bacterial strains used were, XL1-Blue (Stratagene; *recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac* [F' *proAB lacI^q ZΔ15* Tn10 (Tet^r)]), BL21 (DE3) (Stratagene; *E.coli* BF' *dcm ompT hsdS (r_B⁻m_B⁻) galλ* (DE3)) and TG1 (Stratagene; *supE thi-1 Δ(lac-pro AB) Δ(mcrB-hsdSM) S (r_K⁻m_K⁻)* [F' *traD36 proAB lacI^q ZΔ15*]).

E.coli strain TG1 was grown in 2xYT media, which contains per litre: 16 g tryptone, 10 g yeast extract and 5 g NaCl, supplemented where appropriate with 100 mg/l ampicillin. Solid media used contained per litre: 16 g tryptone, 5 g yeast extract, 10 g NaCl and 15 g agar, supplemented where appropriate with 100 mg/l ampicillin.

2.1.5 DNA modifying enzymes

Restriction endonucleases *Eco* R1, *Sma* 1, *Hin* dIII, *Nde* 1, *Bgl* II and *Bam* H1, and T4 DNA ligase were supplied by Amersham Life science. *Pfu* DNA polymerase and *Dpn* 1 restriction endonuclease were supplied by Stratagene.

2.1.6 Protein modifying enzymes

Proteinase K and Trypsin were supplied by Boehringer Mannheim. Factor Xa was supplied by Pharmacia.

2.1.7 Chromatographic media and membranes

Pre-swollen diethylaminoethyl cellulose (DE-52) anion exchange media was obtained from Whatman Biosystems Ltd. Phenyl sepharose high performance hydrophobic interaction media and metal chelating column were obtained from Pharmacia. Sephadex G25 columns were supplied by BIO-RAD. YM3 diaflo ultra-filtration membranes and Centricon-3 concentrators were supplied by Amicon.

2.1.8 Primers

All primers were obtained from the Protein and Nucleic Acid Chemistry Laboratory at the University of Leicester. Primers constructed using a Applied Biosystems 394 DNA/RNA synthesizer.

Primer	Sequence 5'-3'
CT1	CAAGTTTAACTGCTGAAGAATTCATGGCGCCGACAAGCT TAGAGAAATTGCC
CT2	GGCAATTTCTCTAAGCTTGTCGGCGCCATGAATTCTTCAG CAGTTAAACTTG
NT1	GCTGAAGCAGTGGTTTAACCGACAAGCTTAGAGAAATTG
NT2	GCAATTTCTCTAAGCTTGTCGGTTAAACCACTGCTTCAGC
NT3	GCTTAGAGAAATTGCCTAGTATTGAGGGGCGCGGCCAAG ATCTATATAAACTCAACC
NT4	GGTTGAGTTTTATATAGATCTTGGCCGCGCCCCTCAATAC TAGGCAATTTCTCTAAGC
NT5	GCTCTACGAGGAAAAGCACCATCACCACCATCACTGACC CGGGGATCCGTCGACC
NT6	GGTCGACGGATCCCCGGGTCAGTGATGGTGGTGATGGTG CTTTTCCTCGTAGAGC
NT7	GGCACCTTATTCCTGAGGATTGGTGCTGCCCCG
NT8	CGGGGCAGCACCAATCCTCAGGAATAAGGTGCC
AlkB 1	AATTGGAGATCTCCATATGCTTGAGA
AlkB 2	GCCGGGCTCTGAGATCTCACATAAC
AlkB 3	GATCCTGCAACATCCCGG
AlkB 4	GCTTTGGCCACGGCGCG
54	TTTTTTTTTTGAATTCATGGCTAGCTATAAATGCCCGGAT
55	TTTTTTTTTTCCCGGGTCACTTTTCCTCGTAGAGCACATA

2.2 Methods

2.2.1 DNA purification

Plasmid DNA was purified using Wizard DNA purification kits from Promega. For screening purposes, DNA obtained using the miniprep kit was used. For preparation of DNA stock solutions (for archives, mutagenesis or transformations) DNA obtained using the maxiprep kit was used.

2.2.2 Agarose gel electrophoresis

The agarose gel was prepared with 0.8 g of agarose dissolved in 110 ml of 1x TAE buffer (50x TAE buffer consists of 242 g tris base, 18.6 g EDTA and 57.1 ml glacial acetic acid made up to 1 litre), and brought to boil in a microwave. The solution was allowed to cool to 50°C before ethidium bromide was added to a concentration of 0.6 µg/ml. The solution was then cast and allowed to set for two hours before being submerged in 1x TAE. The sample (10 µl) in 1x loading buffer (6x loading buffer consists of 0.25 % bromophenol blue, 0.25 % xylene cyanol and 30 % glycerol in 50 mM EDTA) was loaded in to each well; DNA marker (4 µl) was added to a separate well. Gels were run at 150 V for 30 minutes to 2 hours depending on sample and visualised on a UVP transilluminator.

2.2.3 Polyacrylamide gel electrophoresis

For analysis of DNA by polyacrylamide gel electrophoresis a 5 % polyacrylamide gel was used. The following solution was prepared; 4.2 ml protogel, 0.53 ml 3 % ammonium persulfate, 2.5 ml 10x TBE (10x TBE consists of 108 g tris, 55 g borate and 9.3 g EDTA in 1litre, pH 8.2), 17.8 ml water and 7.5 µl Temed. This solution was poured and allowed to set overnight. The gel was placed in 1x TBE and 10 µl of sample in 1x loading buffer was loaded to each well, and 4 µl of DNA marker was added to a separate well. The gel was run at 30 mA for approximately 2 hours, removed from the glass plates, carefully wrapped in Saran wrap and visualized on a UVP transilluminator.

2.2.4 Preparation of competent cells using the rubidium chloride method

250 ml of 2xTY was inoculated using a overnight culture of the desired bacterial strain and incubated at 37 °C with shaking for 14 hours. The cells were harvested by centrifugation at 5000 g for 10 minutes and the supernatant discarded. The cells were re-suspended in 100 ml of pre-cooled transformation buffer I (0.29 g potassium acetate, 1.21 g rubidium chloride, 0.15 g calcium chloride, 0.99 g

manganese chloride, 15 ml glycerol-made up to 100 ml total volume and pH adjusted to 5.8 with dilute acetic acid). Cells were left on ice for 15 minutes prior to centrifugation at 5000 g for 15 minutes. The supernatant was discarded and the cells re-suspended in 10 ml of pre-cooled transformation buffer II (0.02 g MOPS, 0.11 g calcium chloride, 0.01 g rubidium chloride, 1.5 ml glycerol-made up to 10 ml total volume and pH adjusted to 6.5 with 1 M sodium hydroxide). The solution was placed on ice for 15 minutes and 100 μ l and 200 μ l aliquots of the solution were placed into Eppendorf tubes, pre-cooled in dry ice, prior to storage at -70 °C.

2.2.5 Site-directed mutagenesis

Reaction mixtures were prepared with four different concentrations of plasmid DNA; 5 μ l of 10x reaction buffer, 5 μ l of 25 ng/ μ l primer A, 5 μ l of, 25 ng/ μ l primer B, 1 μ l of 10 mM dNTPs, and plasmid DNA providing 5, 10, 20 and 50 ng of template DNA made up in 25 μ l of water.

1 μ l of *pfu* DNA polymerase was added to the reaction, which was then overlaid with 30 μ l of mineral oil. This was then placed in a thermocycler and subjected to the following regime.

Cycles	Temperature °C	Time
1	95	30 seconds
16	95	1 minute
	55	1 minute
	68	10 minutes

After thermocycling, 1 μ l of *Dpn* I restriction enzyme was carefully added below the mineral oil and the solution gently pipetted to aid mixing. The solution was spun on a bench top centrifuge for 1 minute before incubation at 37 °C for 1 hour. After the *Dpn* I digestion, 1-5 μ l of the sample was used to transform competent XL1-Blue cells.

2.2.6 Restriction digest

The following solution was prepared; 10 μ l of DNA, 2.5 μ l of 10x OPA, 1 μ l of restriction enzyme and 11.5 μ l. This was incubated at 37 °C for 3 hours.

2.2.7 DNA ligation

The following solution was prepared; 15 μ l of DNA solution, 7 μ l of water, 2.5 μ l of 10x T4 buffer and 0.5 μ l of T4 DNA ligase. This solution was incubated at 16 °C for 16 hours. 1 μ l of reaction mixture was used to transform competent cells.

2.2.8 Phenol/chloroform extraction

This procedure was used to purify DNA following restriction digests for ligation reactions. To 25 μ l of a DNA solution an equal volume of phenol was added. The solution was mixed vigorously to form an emulsion. The emulsion was centrifuged for 1 minute to give a bi-layered solution, of which the upper aqueous phase was removed to an Eppendorf tube. To the remaining organic phase 25 μ l of TE buffer was added and the above extraction repeated. The two aqueous phases were combined.

To the combined aqueous phases 50 μ l of chloroform was added and this solution was mixed and centrifuged as above. The aqueous phase was removed into 100 μ l of ice-cold ethanol, and the solution placed at -70 °C for 1 hour. The solution was then centrifuged for 10 minutes. The supernatant was discarded and residual supernatant allowed to evaporate. The DNA was then dissolved in 25 μ l of TE buffer.

2.2.9 Determination of protein concentration

Protein concentrations were determined using the BIORAD *DC* protein assay, based on the Lowry assay (Lowry *et al.*, 1951). Bovine serum albumin solutions of known concentrations were used as the standards to construct the calibration curve.

2.2.10 Transformation of competent cells

Competent cells were thawed on ice and a 50 μ l aliquot transferred to a fresh Eppendorf tube and 1-5 μ l of DNA, dependent on its source, was added. The Eppendorf tube was gently flicked and the solution placed on ice for 30 minutes. The solution was heat pulsed at 42 °C for 1 minute and then returned to ice for a further 2 minutes. To the solution 500 μ l of 2xYT pre-heated to 42 °C was added, this was then placed in at 37 °C with shaking for 1 hour. The solution (150 μ l) was then plated onto solid 2xYT, supplemented with ampicillin, and incubated at 37 °C overnight.

2.2.11 SDS polyacrylamide gel electrophoresis

The following solution was prepared; 12 ml protogel, 3.8 ml 1.5 M tris/HCl pH 8.8, 150 μ l 10 % SDS, 50 μ l 10 % ammonium persulphate and 10 μ l temed. This solution was poured between the gel plates until it was 2 cm from the plate edge. A small amount of water saturated butanol was added and the gel allowed to set overnight. To 3 ml of the stacking gel solution (stacking gel solution consists of 33.2 ml protogel, 25 ml 1 M tris/HCl pH 6.9, 2 ml 10 % SDS and 138 ml water) 50 μ l 10 % ammonium persulphate and 5 μ l temed were added. The solution was poured and allowed to set for 1 hour.

For cell culture samples, 3 ml of overnight culture was spun down and resuspended in 50 μ l of water, 100 μ l 10 % SDS and 150 μ l SDS loading buffer (SDS loading buffer consists of 154 mg DTT, 1 ml glycerol, 2 ml 10 % SDS, 170 μ l tris/HCl pH 6.8, 1.63 ml water and 400 μ l bromophenol blue) and boiled for 5 minutes. The solution was then drawn through a 0.5 mm syringe needle 10 times. For other protein samples, an equal volume of SDS loading buffer was added and the sample boiled for 5-10 minutes.

Sample volumes of 10-20 μ l were loaded on to the gel along with 3 μ l of protein molecular weight marker. The gel was run at 30 mA in SDS running buffer (10x SDS running buffer consists of 28.4 g glycine, 60 g tris and 200 ml 10 % SDS) plus 70 μ l 2-mercaptoethanol for 2-4 hours, depending on sample. The stacking gel was removed and the remaining gel placed in staining solution

(staining solution consists of 0.25 % Coomassie brilliant blue R250 in 5:1:5 v/v methanol : acetic acid : water) for 20 minutes. The gel was then placed in de-stain solution (de-stain solution consists of 5:1:5 v/v methanol : acetic acid : water) for 30 minutes.

2.2.12 DNA sequencing

DNA sequencing was performed by the Protein and Nucleic Acid Chemistry Laboratory at the University of Leicester using an Applied Biosystems 377 automated fluorescent sequencer. Pelleted 5 ml overnight cultures of known optical density (OD_{600}) and primers at 0.8 pmole/ μ l were supplied for the sequencing reactions.

2.2.13 N-terminal sequencing

N-terminal sequencing was performed on unstained SDS PAGE by the Protein and Nucleic Acid Chemistry Laboratory at the University of Leicester using an Applied Biosystems 476A protein sequencer.

2.2.14 Electrospray mass spectrometry

Mass measurements were performed by the Protein and Nucleic Acid Chemistry Laboratory at Leicester University, using a MicroMass Platform single quadrupole electrospray mass spectrometer. Cone voltage was 30-32 volts. Data analysis was performed using MassLynx software from MicroMass. Samples were diluted in a solution of 50 % acetonitrile and 0.3 % formic acid at a concentration of 50 pmole/ μ l.

2.2.15 Preparation of Fe-, Cd and apo- forms of the rubredoxin

The rubredoxin was precipitated with 70 % trichloroacetic acid (TCA) and centrifuged at 5000 g for 20 minutes. The precipitate was dissolved in degassed 0.5 M tris base/0.5 M 2-mercaptoethanol at a protein concentration of 5 mg/ml. Argon

was bubbled through the solution for 10 minutes, the solution made air-tight and left at room temperature for 3 hours.

The protein was re-precipitated with TCA and harvested as before. Precipitate was re-dissolved in degassed 0.5 M tris base at a protein concentration of 5 mg/ml, which contained a 5-fold molar excess of ferrous ammonium sulphate for the Fe-form or cadmium chloride for the Cd-form. The solution was quickly flushed with argon, sealed and left at room temperature for 10 minutes. The solution was then put on ice and air admitted. It was left on ice for 3 hours with gentle shaking every 20 minutes.

The solution was concentrated to 2 ml before being applied to a Sephadex G25 column equilibrated with 5 volumes of 50 mM potassium phosphate, pH 7.0.

2.2.16 Steady-state assay of rubredoxin

Steady-state kinetic assays for rubredoxin were conducted using a method modified from (Lee *et al.*, 1998). The reaction mixture (1 ml) comprised 50 mM Tris-HCl buffer, pH 7.8, 83 μ M horse heart cytochrome *c*, 2 μ M rubredoxin reductase and 5-100 nM rubredoxin. The reaction was commenced by addition of NADH (0.3 mM final concentration) at 30 °C. The reduction of cytochrome *c* was followed by observing the absorbance at 550 nm. In calculation of reduction rates a molar coefficient of 21 000 $\text{M}^{-1} \text{cm}^{-1}$ for cytochrome *c* was used.

2.2.17 NMR

The NMR sample was made up in either 90 % H_2O /10 % D_2O or 99 % D_2O . Spectra were recorded at 298 K and 313 K. Proton chemical shifts were referenced to 2,2-dimethylsilapentane-5-sulfonic acid at 0 ppm. The chemical shifts of ^{113}Cd were referenced to 0.1M $^{113}\text{Cd}(\text{ClO}_4)_2$ at 0 ppm. Homonuclear two-dimensional double-quantum-filtered correlated spectroscopy (COSY), total correlated spectroscopy (TOCSY) and nuclear Overhauser effect spectroscopy (NOESY) data were acquired using standard pulse sequences; where appropriate ^{113}Cd -decoupling, using a GARP decoupling sequence, was incorporated during the evolution and acquisition periods. For all experiments, solvent suppression

was achieved by a low-power preirradiation. To obtain ^{113}Cd - ^1H connectivities, a standard heteronuclear single quantum coherence (HSQC) pulse sequence was used with polarisation transfer delays optimised for coupling constants of either 40 or 17 Hz. All heteronuclear experiments were carried out with ^{113}Cd -decoupling during acquisition. All spectra were measured in phase-sensitive modes. COSY data matrices were 4096x512, TOCSY & NOESY data matrices were 2048x512 and ^{113}Cd - ^1H 2D HSQC matrices were 4096x128 real time domain datapoints.

2.2.18 Limited proteolysis

200 μg of rubredoxin and 10 μg of trypsin or proteinase K were incubated at 37 °C and 10 μl samples removed at intervals in to 10 μl of SDS loading buffer and boiled for 5 minutes. Samples were analyzed by SDS-PAGE.

2.2.19 X-ray scattering data collection

X-ray solution scattering data were collected in one session with the low angle scattering camera on station 2.1 at the Synchrotron Radiation Source (SRS), Daresbury, UK using a position-sensitive multiwire proportional counter. At the sample-to-detector distance of 1.25 m and the X-ray wavelength of $\lambda = 1.54 \text{ \AA}$, a momentum transfer interval of $0.005 \text{ \AA}^{-1} \leq s \leq 0.087 \text{ \AA}$ was covered. The modulus of the momentum transfer is defined as $s = 2 \sin \Theta / \lambda$, where 2Θ is the scattering angle. The s-range was calibrated using an orientated specimen of wet rat tail collagen (based on a diffraction spacing of 670 \AA). Samples were contained in a brass cell holding a teflon ring sandwiched by two mica windows that define a sample volume of 120 μl and a thickness of 1.5 mm. The brass cell was maintained at 18 °C during data acquisition. Buffer and sample were measured in alteration, each in a frame of 60 s (amounting to total measuring times between 15 and 60 min depending on sample concentration and changes monitored on-line during experiments).

Chapter 3

Isolation of the Individual Domains

3.1 Introduction

The rubredoxin reductase-rubredoxin electron transfer complex of *Pseudomonas oleovorans* is a physiological complex that donates electrons to a membrane-bound alkane hydroxylase. The rubredoxin is a member of a family of non-heme iron proteins whose active sites comprise an iron atom tetrahedrally coordinated to four cysteine residues (Sieker *et al.*, 1994). Rubredoxins serve as electron carriers and most have a molecular mass of approximately 6 kDa (Blake *et al.*, 1991; Kitamura *et al.*, 1997; Yoon *et al.*, 1999). Although, rubredoxins are presumed to function as electron carriers, the rubredoxin of *Pseudomonas oleovorans* is the only member of this family of proteins for which the electron transfer reactions have been delineated in detail (Lee *et al.*, 1998). Cloning and sequencing of the genes encoding proteins involved in the hydroxylation of alkanes established the role of rubredoxin in transferring electrons from a specific NADH-dependent rubredoxin reductase to the alkane hydroxylase (Eggink *et al.*, 1990; Eggink *et al.*, 1987).

Ps. oleovorans rubredoxin is unique in that it contains two binding sites for iron. Unlike other rubredoxins (molecular mass ~6 kDa), the *Pseudomonas oleovorans* rubredoxin is a 19 kDa protein and the gene encoding the protein, *alk G*, is most likely the product of a gene duplication event (Kok *et al.*, 1989a). The 2Fe form of the rubredoxin is believed to be the physiological form, but it is less stable than the readily isolated 1Fe form (Lee *et al.*, 1997). The 1Fe form is also active in its reaction with rubredoxin reductase (Lee *et al.*, 1998; Lode & Coon, 1971). Cleavage of the 2Fe rubredoxin at residue Met-50 with cyanogen bromide yields a C-terminal peptide (residues 51 to 173) that can be folded around iron (Lode & Coon, 1971) and cobalt (May *et al.*, 1984), and the folded peptide can act as an electron carrier during steady-state hydroxylation reactions (Lode & Coon, 1971; May *et al.*, 1984). The corresponding N-terminal fragment (residues 1 to 50) can be transiently folded around iron, but is highly labile with a half life of only 30 minutes at room temperature (Lode & Coon, 1971). Sequence alignments indicate that the instability of the N-terminal fragment may be a result of the position of cleavage, which would in other rubredoxins disrupt the β -sheet (Eidsness *et al.*, 1997).

Rubredoxins have been studied extensively from the structural viewpoint. Three-dimensional structures with resolutions better than 1.5 Å have been reported (Adman *et al.*, 1991; Bau *et al.*, 1998; Dauter *et al.*, 1996; Frey *et al.*, 1987; Stenkamp *et al.*, 1990) and nuclear magnetic resonance studies have been utilised to investigate metal ion ligation (Henehan *et al.*, 1993; Pountney *et al.*, 1995), contributions made by protein structure to the overall thermostability of rubredoxins (Richie *et al.*, 1996) and solution structure (Bertini *et al.*, 1998; Blake *et al.*, 1992b). However, despite this detailed structural knowledge, little is known about the function of rubredoxins in general. With this deficiency in mind the rubredoxin from *Pseudomonas oleovorans* has been chosen for study, as it has a well-defined biochemical role. A structural determination of the 2Fe rubredoxin, rubredoxin reductase and ultimately the productive electron transfer complex formed by the two proteins will provide valuable information on electron transfer mechanisms in this physiological redox system.

A major question arising from a study of this system is the role of the labile N-terminal domain in recognition of its physiological redox partners and electron transfer. Clearly, detailed functional and structural characterization of the component domains of the 2Fe-rubredoxin is required. With this in mind, this chapter describes the genetic excision and purification of the N and C-terminal domains of *Pseudomonas oleovorans* rubredoxin and studies of the spectroscopic and functional properties of the two domains.

3.2 Molecular Biology

To create expression constructs for the individual domains the middle of the linker region was targeted by site-directed mutagenesis. This region is removed (~ 30 residues) from the structural domain boundaries, as determined by sequence alignments, and therefore should not contribute to domain structure. The expression construct (pKR10; Lee *et al.*, 1997) for di-domain rubredoxin was used as the template for the directed mutagenesis (details of which can be found in Chapter 2). This contains the rubredoxin gene, *alk G*, flanked by unique *Eco* RI and *Sma* I restriction sites at the N and C-terminal regions of the gene respectively, with the gene under the control of the P_{tac} promoter. The primers NT1 and NT2

Chapter 3: Isolation of the Individual Domains

were used to engineer a translation termination codon at the location of the codon for Ala-86 in the wild-type gene, producing an expression construct (designated pKR11) for the N-terminal Fe-binding domain of the di-domain rubredoxin. To create an expression plasmid for the isolated C-terminal Fe-binding domain a *Eco* RI site was first engineered at the location of two codons encoding Ala-83 and Val-84 of the wild-type gene, and the codon for Val-85 was converted to one encoding methionine. These changes were introduced simultaneously using the oligonucleotides CT1 and CT2. Following mutagenesis, a 252 bp DNA fragment encoding the entire N-terminal Fe-binding domain and half the putative interdomain linker was excised by digestion with *Eco* RI, and re-ligation of the resultant plasmid brought Met-85 (introduced in the preceding mutagenesis reaction) to the same location as Met-1 in the wild-type rubredoxin expression plasmid (pKR10) to produce a plasmid construct (designated pKR12) capable of expressing the C-terminal portion of the *alk* G gene.

Lack of expression of the N-terminal domain in cells transformed with pKR11 (see Section 3.3.1) necessitated a new strategy for isolation of the N-terminal domain. This involved constructing a mutant version of the di-domain rubredoxin containing a C-terminal His-tag and a Factor Xa proteolysis site in the putative interdomain linker region. The N-terminal domain can thus be isolated initially as the mutant di-domain (possibly exploiting the C-terminus His-tag) this is then followed by digestion with Factor Xa to yield the individual domains. The individual N and C-terminal domains can then be separated using the C-terminal His-tag. Consequently, a version of the *alk* G gene encoding recombinant rubredoxin containing an additional six histidine residues at the C-terminus was engineered using the primers NT5 and NT6 (new plasmid designated pKR13). A factor Xa proteolysis site (Ile-Glu-Gly-Arg) was then introduced into this gene by mutagenesis of the codons corresponding to residues Ala-96, Asp-97, Val-98 and Lys-99 using the primers NT3 and NT4. The resultant construct was designated pKR14. The position of all the mutations, and the mutagenic primers, in the *alk* G gene are shown in Figure 3.1.

All modified genes contained in the new plasmid constructs were completely re-sequenced on both the coding and non-coding strands (using the primers 54, 55, and the mutagenic primers) to check that spurious changes had not

arisen during the mutagenesis procedures. Large-scale plasmid DNA was prepared from the sequenced clones and transformed into competent *E. coli* cells (strain TG1).

3.3 Purification

3.3.1 Protein Expression

SDS-PAGE analysis of *E. coli* strain TG1 transformed with the N-terminal and C-terminal domain expression plasmids (pKR11 and pKR12, respectively), revealed that only the C-terminal domain was expressed at high levels within the recombinant host (Figure 3.2, panel a). Evidence for expression of the N-terminal domain in the presence or absence of inducer (2 mM IPTG, added at OD₆₀₀ of 0.5) in *E. coli* strain TG1, or in the protease deficient strain of *E. coli* (strain BL21 DE3) was not obtained. The lack of expression of the N-terminal domain likely reflects the poor structural stability of this domain, as inferred from metal binding NMR studies in the full-length di-domain rubredoxin (Lee *et al.*, 1997) and the poor stability of the N-terminal fragment generated by cyanogen bromide cleavage of rubredoxin (Lode & Coon, 1971).

SDS-PAGE analysis of *E. coli* strain TG1 transformed with the mutant di-domain expression constructs (pKR13 and pKR14) indicated that the encoded proteins are expressed at high levels (Figure 3.2, panel b), and can be seen as bands of approximately 20 kDa. The inclusion of 2mM IPTG had no obvious effect on the expression of either mutant di-domain rubredoxin. This finding is consistent with the effect of IPTG on the expression of the recombinant di-domain rubredoxin (Lee *et al.*, 1997).

3.3.2 Protein Purification

Purification of the C-terminal domain rubredoxin was achieved using the protocol described previously for the di-domain rubredoxin (Lee *et al.*, 1997). This involves cell breakage using French press, an ammonium sulphate fractionation

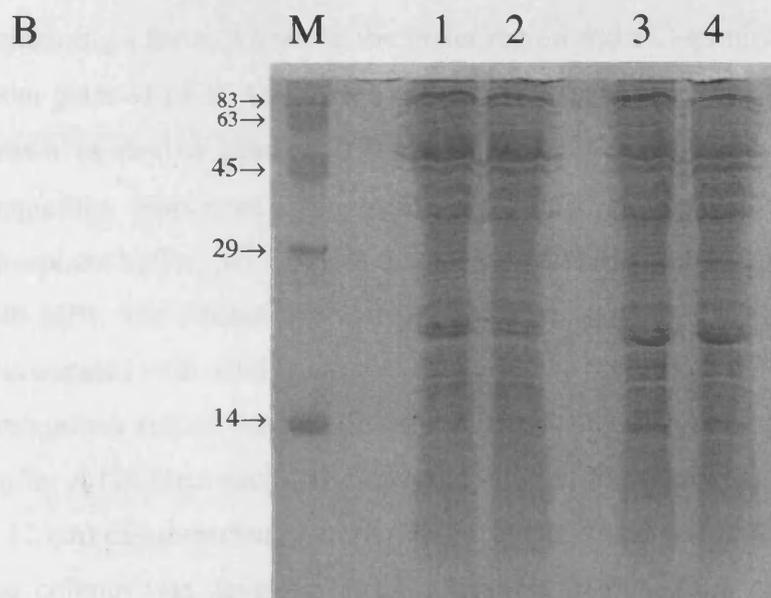
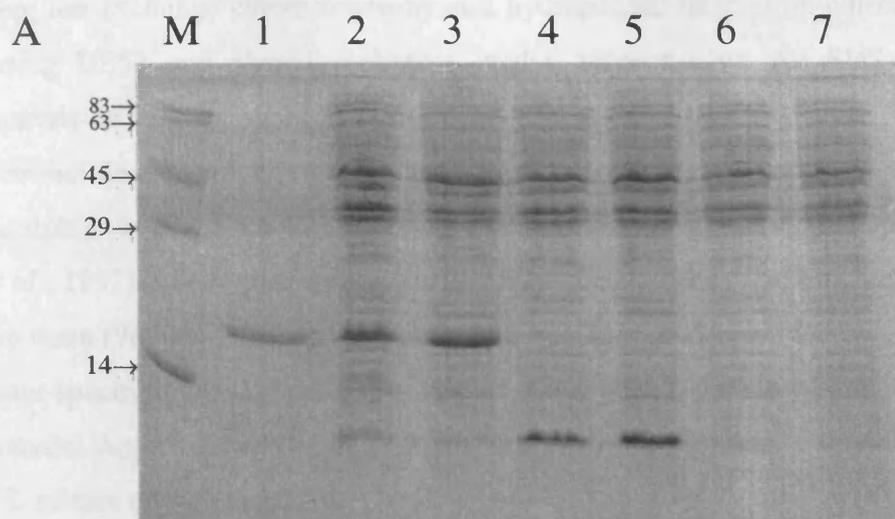


Figure 3.2. SDS PAGE analysis of the expression of the individual domains and the didomain rubredoxin. Panel A, expression of the C-terminal and N-terminal domains in *E. coli* strain TG1. Lane M, molecular weight markers (molecular masses in kDa are indicated to the left). Lane 1, didomain rubredoxin. Lanes 2 & 3, *E. coli* strain TG1 transformed with plasmid pKR10. Lanes 4 & 5, *E. coli* strain TG1 transformed with plasmid pKR12. Lanes 6 & 7, *E. coli* strain TG1 transformed with plasmid pKR11. Lanes 3, 5 & 7 induced with 2 mM isopropyl β -D-thiogalactoside. Panel B, expression of the mutant forms of the didomain rubredoxin. Lane M, molecular weight markers (molecular masses in kDa are indicated to the left). Lanes 1 & 2, *E. coli* strain TG1 transformed with plasmid pKR13. Lanes 3 & 4, *E. coli* strain TG1 transformed with plasmid pKR14. Lanes 2 & 4 induced with 2 mM isopropyl β -D-thiogalactoside.

step, ion exchange chromatography and hydrophobic interaction chromatography (using DE52 and phenyl sepharose media, respectively). An SDS-PAGE gel analysis at various stages of purification process is shown in Figure 3.3. The C-terminal domain migrates as multiple bands due to the formation of intraprotein disulphide bonds, a feature which was also observed for the di-domain protein (Lee *et al.*, 1997). The N-terminal sequence of the protein (Pro, Thr, Ser, Leu, Glu) and the mass (9690 ± 1 Da; expected mass 9691.6 Da), as determined by electrospray mass spectrometry, are consistent with those expected for the isolated C-terminal domain. Approximately 30 mg of purified C-terminal domain was isolated from a 1 L culture of the recombinant host.

The N-terminal domain was isolated from the mutant di-domain rubredoxin containing a factor Xa site in the linker region and a C-terminal His-tag (expressed from plasmid pKR14). *E. coli* strain TG1 transformed with plasmid pKR14 was grown in double strength YT media at 37 °C in the presence of 100 µg/mL ampicillin. Harvested cells were resuspended in buffer A (0.1 M potassium phosphate buffer, pH 7.0) and disrupted in a French press at 4 °C at a pressure of 140 MPa. The extract was clarified by centrifugation at 15000 g for 90 min and fractionated with solid ammonium sulfate. The precipitate with 40 - 60 % saturated ammonium sulfate was dissolved in buffer A and dialysed against two changes of buffer A (10 litres each). The dialysed solution was then applied to a column (5 cm x 12 cm) of ion-exchange resin (DE-52) equilibrated with buffer A. After washing, the column was developed with a gradient of potassium chloride (0 to 0.5 M) contained in buffer A. Mutant di-domain rubredoxin eluted around 0.2 M potassium chloride. Fractions were pooled and applied to a metal chelating column (5 ml HiTrap; Pharmacia) charged with 50 mM nickel sulfate and equilibrated with binding buffer (50 mM potassium phosphate, pH 7.5, 1.0 M sodium chloride and 5 mM imidazole). The column was washed with 100 ml binding buffer and 20 ml wash buffer (50 mM potassium phosphate buffer, pH 7.5, 1.0 M sodium chloride, 60 mM imidazole) and the protein was then eluted with elution buffer (50 mM potassium phosphate buffer, pH 7.0, 1.0 M sodium chloride, 1 M imidazole). Approximately 10 mg of purified mutant di-domain rubredoxin was isolated from a 1L culture of the recombinant host. The N-terminal sequence analysis of the purified protein revealed that the first six residues are Ala, Ser, Tyr, Lys, Cys, and

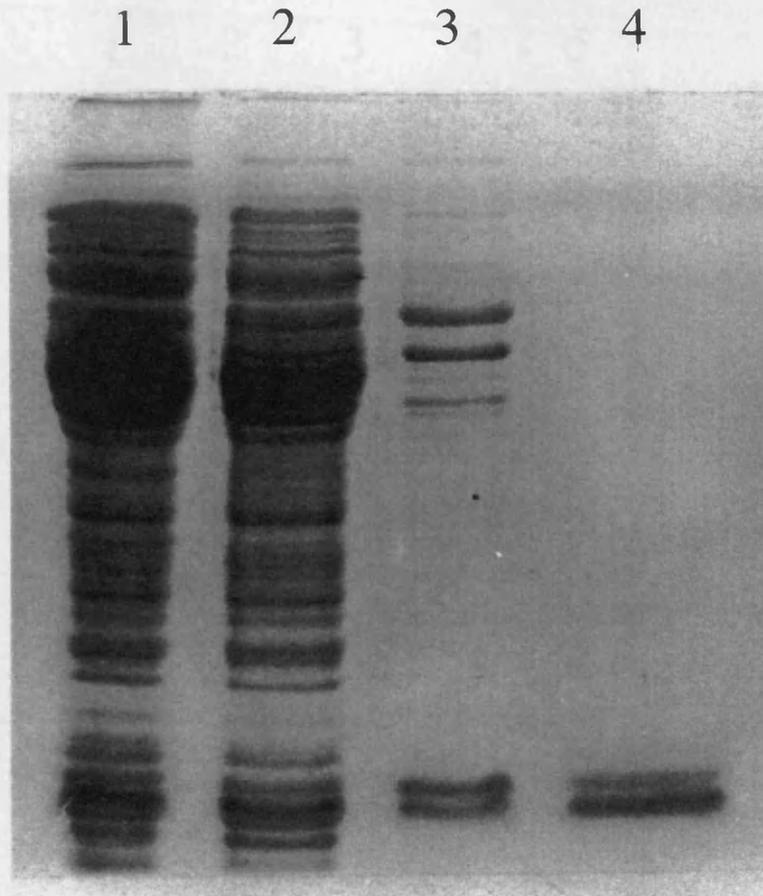


Figure 3.3. Purification of the C-terminal domain rubredoxin. Lane 1, cell free extract. Lane 2, (NH₄)₂SO₄ fraction. Lane 3, pooled fractions after ion-exchange chromatography. Lane 4, pooled fractions after hydrophobic interaction chromatography.

Pro, which is consistent with the predicted sequence of the mutant *alk G* gene. Moreover, analysis by electrospray mass spectrometry indicated the mass of the protein to be 19640 ± 4 Da (expected mass 19632.5 Da). An SDS-PAGE gel analysis at various stages of purification process is shown in Figure 3.4, panel a.

Release of the N-terminal domain from the purified mutant di-domain rubredoxin was achieved by the addition of factor Xa (50 μ g of mutant di-domain per unit of factor Xa) for 24 h at room temperature. Digestion resulted in the appearance of a diffuse band migrating with apparent mass 10 kDa, and the loss of the parental 20 kDa band (Figure 3.4, panel b). Analysis by electrospray mass spectrometry revealed that the diffuse band comprised two components of masses

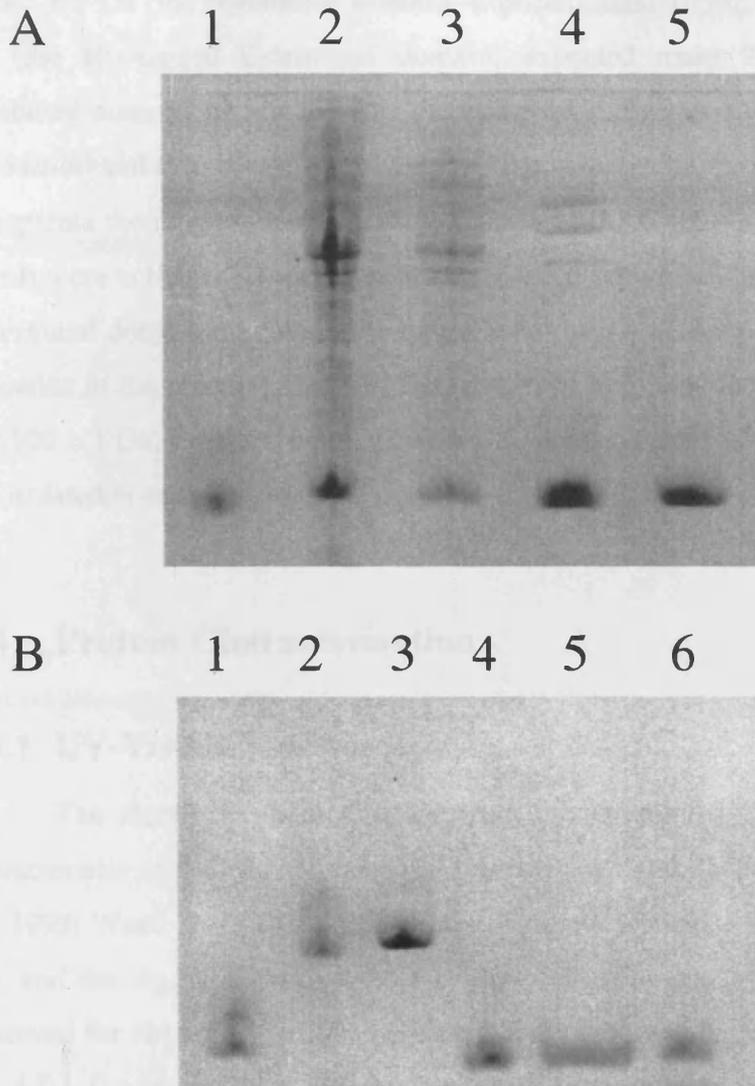


Figure 3.4. Purification of the N-terminal domain rubredoxin. Panel a, purification of the mutant di-domain rubredoxin. Lane 1, purified wild-type di-domain rubredoxin; Lane 2, cell free extract; Lane 3, $(\text{NH}_4)_2\text{SO}_4$ fraction; Lane 4, pooled fractions after ion-exchange chromatography; Lane 5, pooled fractions after metal chelating column. Panel b, isolation of the N-terminal domain by factor Xa proteolysis of the rubredoxin fusion protein. Lane 1, purified C-terminal domain; lane 2, purified wild-type di-domain rubredoxin; lane 3, purified mutant di-domain rubredoxin containing the His-tag and factor Xa cleavage site; lane 4, as lane 3 but after 24 h digestion with factor Xa; lane 5, protein (His-tagged C-terminal domain) retained on metal chelating column following factor Xa digestion of mutant di-domain rubredoxin; lane 6, protein (released N-terminal domain following factor Xa digestion of mutant di-domain) eluted with 'binding buffer' from the metal chelating column.

10497 ± 1 Da (the N-terminal domain, expected mass 10502.6 Da) and 9145 ± 2 Da (the His-tagged C-terminal domain; expected mass 9147.9 Da), with a combined mass of 19,642 ± 3 Da. The proteolytic digest was desalted (to remove imidazole) and then re-applied to the Hi-Trap column (buffer conditions as above) to separate the released N-terminal domain from the C-terminal domain. Fractions (1 ml) were collected as the column was washed with binding buffer. The isolated N-terminal domain migrated as a single band during SDS-PAGE. The N-terminal sequence of the protein (Ala, Ser, Tyr, Lys, Cys, Pro), and the mass of the domain (10,500 ± 1 Da; expected mass 10502.6 Da), are consistent with those expected for the isolated N-terminal domain.

3.4 Protein Characterisation

3.4.1 UV-Visible Spectroscopy

The electronic absorption spectrum of the isolated C-terminal domain is characteristic of the conventional 1Fe rubredoxins (Blake *et al.*, 1991; Henahan *et al.*, 1993; Wastl *et al.*, 2000). Absorption maxima are located at 498, 380 and 280 nm, and the $A_{280}:A_{498}$ ratio is 4.0:1 (Figure 3.5). The spectrum is similar to that observed for chemically cleaved C-terminal domain, which has a $A_{280}:A_{498}$ ratio of 4.3-4.5:1 (Lode & Coon, 1971). The larger ratio reflects the larger size of the chemically cleaved domain, which has a mass of ~ 13 kDa. The spectrum of the genetically excised C-terminal domain was stable over a period of 25 h at 20 °C.

In contrast to the C-terminal domain, the N-terminal domain was colourless 'as purified', and lacked signature in the visible region characteristic of rubredoxins. This is consistent with work on the di-domain rubredoxin where the N-terminal iron is lost during the purification procedure, whereas the C-terminal iron is retained (Lee *et al.*, 1997). However, the apo-N-terminal domain was readily converted to an Fe-bound form by precipitation with trichloroacetic acid using the reconstitution method described previously for the di-domain rubredoxin (Lee *et al.*, 1997). The spectrum of the reconstituted domain has absorption maxima at 448, and 280 nm, and although it has a large shoulder centred around

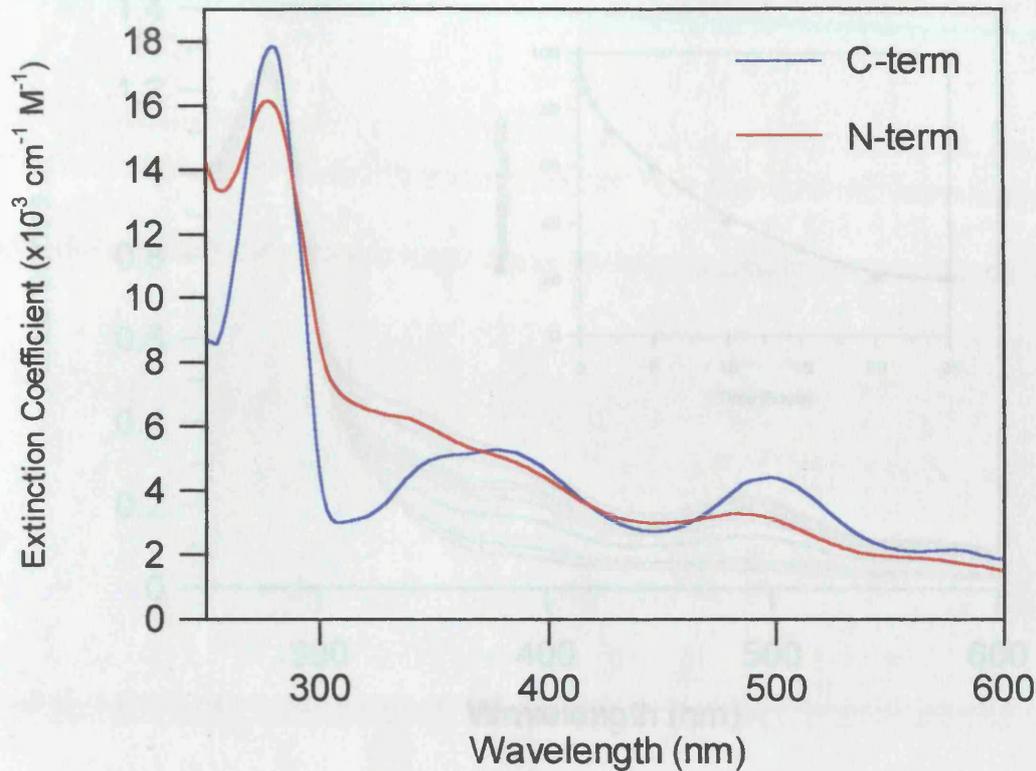


Figure 3.5. UV-visible spectra the N and C-terminal domain rubredoxins. Spectra were recorded with protein in 50 mM potassium phosphate buffer, pH 7.0 at 20 °C.

350-380 nm, it lacks the well-defined maxima at 380 nm seen for other rubredoxins (Figure 3.5). The poor spectral definition in this region likely reflects the structural instability of the N-terminal domain and/or the harsh procedures used to reconstitute the domain with iron. The $A_{280}:A_{488}$ ratio is larger (5.0:1) than the corresponding ratio for the transiently stable, reconstituted N-terminal domain obtained by chemical cleavage with cyanogen bromide (ratio 4.3-4.5:1; Lode & Coon, 1971), again reflecting the additional sequence derived from the linker region present in the genetically constructed N-terminal domain. The smaller N-terminal domain generated by cyanogen bromide cleavage was reported to be very unstable, with a half-life of 30 minutes at room temperature (Lode & Coon, 1971). In contrast, the proteolytically generated N-terminal domain was considerably more stable, having a half-life of about 5 h at 20 °C (as determined by iron loss at 490 nm), as shown in Figure 3.6.

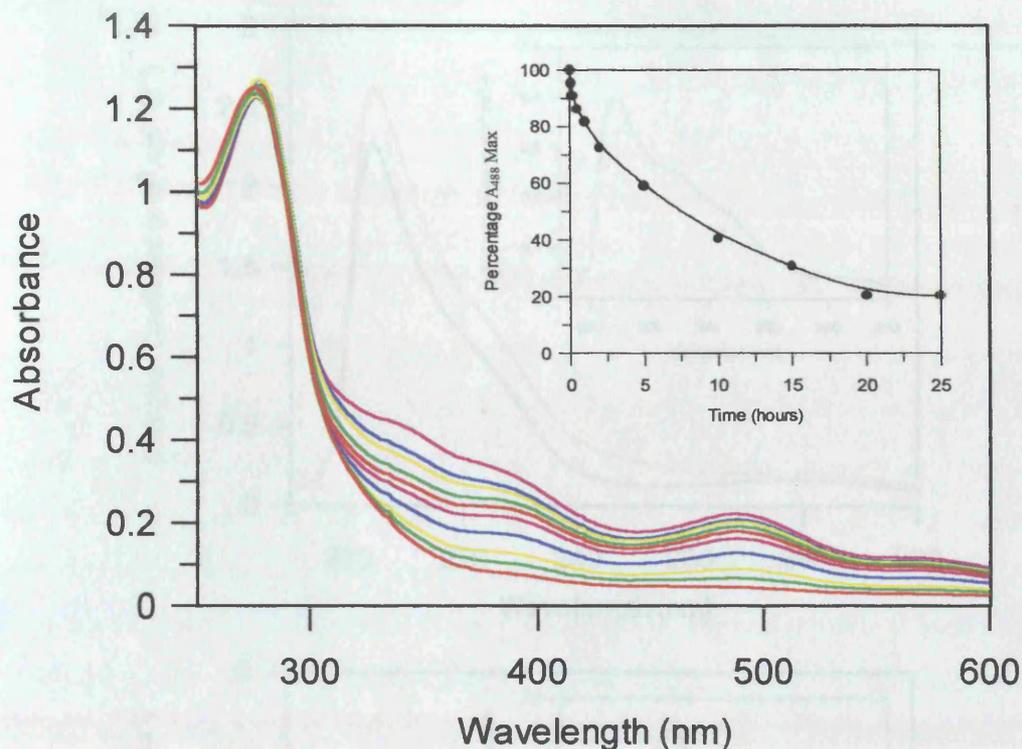


Figure 3.6. Stability of the N-terminal domain rubredoxin. Decay of the spectrum for the N-terminal domain over 25 hours. Inset, A_{488} , as a percentage of A_{488} at time zero, versus time. ●, individual data points; solid line, smooth line fit.

The isolated N- and C-terminal domains could be refolded around cadmium using the method published for the di-domain rubredoxin (Lee *et al.*, 1997). The electronic absorption spectrum of the cadmium-substituted C-terminal domain revealed the presence of a broad absorption between 200 and 300 nm superimposed on the spectrum of the apo-protein, from which the difference spectrum was calculated (Figure 3.7). The difference spectrum could be resolved into four Gaussian components centred at 202, 224, 239 and 276 nm. The ability to deconvolute the difference spectrum into four components mirrors previous work on the parental di-domain rubredoxin from *Pseudomonas oleovorans* (Lee *et al.*, 1997) and also similar studies on the 1Fe rubredoxin from *D. gigas* (Henehan *et al.*, 1993). Jørgensen's electronegativity theory (Jørgensen, 1970) attributes the 239 nm band to a Cd-thiolate charge-transfer transition. The precise assignment of the bands at 202 and 224 nm can not be made at this stage, but they are clearly charge-transfer in character. As discussed for the di-domain protein, the absorption band

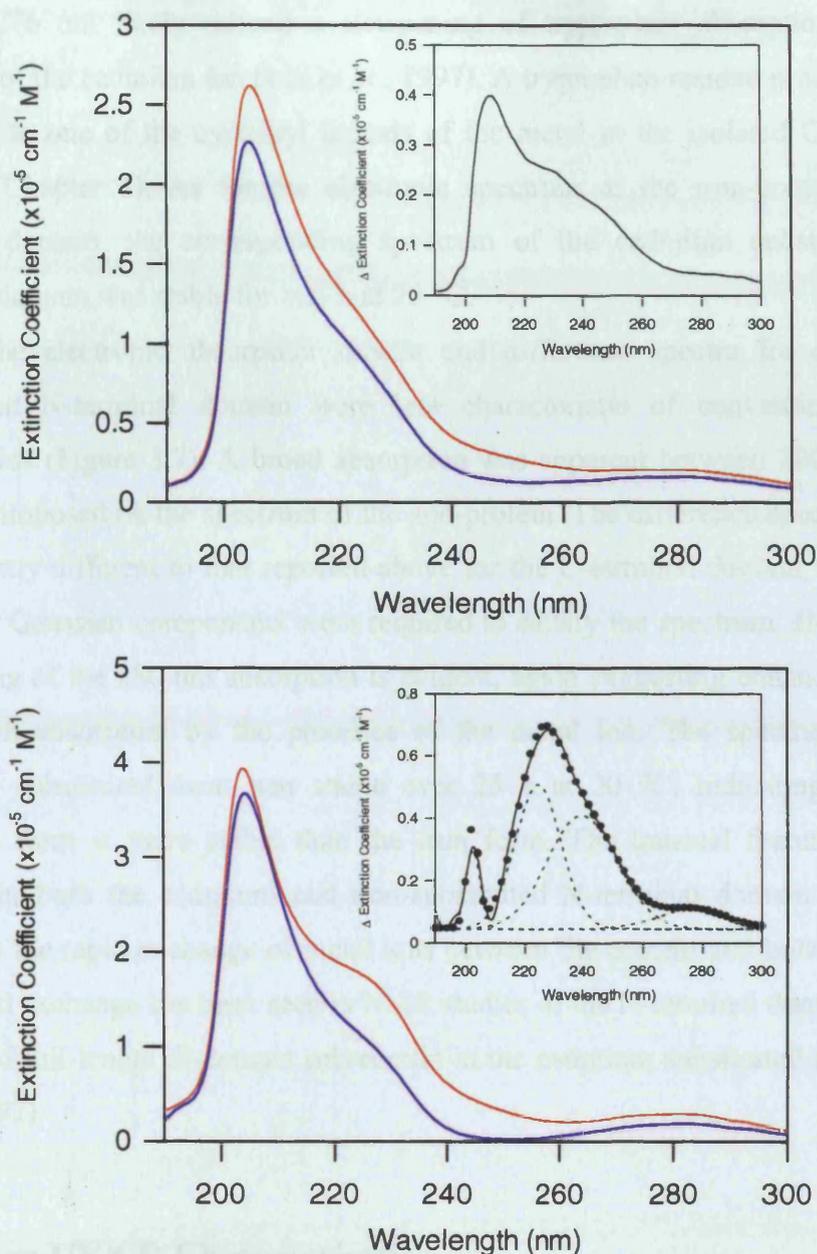


Figure 3.7. Optical spectra and difference spectra of the Cd-substituted forms of the purified N-terminal and C-terminal domains. Upper panel, optical spectra for Cd-substituted (red line) and apo- (blue line) N-terminal domain. Inset, difference absorption spectrum (Cd-substituted minus apo N terminal domain). Lower panel, optical spectra for Cd-substituted (red line) and apo- (blue line) C-terminal domain. Inset, difference absorption spectrum (Cd-substituted minus apo C terminal domain) and Gaussian analysis of the difference spectrum. Broken lines, individual Gaussian components; ●, sum of Gaussian components; solid line, difference spectrum.

seen at 276 nm likely reflects a sharpening of tryptophan absorption by the presence of the cadmium ion (Lee *et al.*, 1997). A tryptophan residue is adjacent in sequence to one of the cysteinyl ligands of the metal in the isolated C-terminal domain (Chapter 1). As for the electronic spectrum of the iron-containing C-terminal domain, the corresponding spectrum of the cadmium substituted C-terminal domain was stable for >25 h at 25 °C.

The electronic absorption spectra and difference spectra for cadmium-substituted N-terminal domain were less characteristic of conventional 1Fe-rubredoxins (Figure 3.7). A broad absorption was apparent between 200 and 300 nm superimposed on the spectrum of the apo-protein. The difference spectrum was significantly different to that reported above for the C-terminal domain, and more than four Gaussian components were required to satisfy the spectrum. However, a sharpening of the 280 nm absorption is evident, again suggesting enhancement of tryptophan absorption by the presence of the metal ion. The spectrum of the cadmium substituted form was stable over 25 h at 20 °C, indicating that the cadmium form is more stable than the iron form. The unusual features of the spectra for both the cadmium and iron-substituted N-terminal domain might be related to the rapid exchange of metal ions between the protein and bulk solution. This rapid exchange has been seen in NMR studies of the N-terminal domain in the context of full-length di-domain rubredoxin in the cadmium substituted form (Lee *et al.*, 1997).

3.4.2 Far-UV CD Characterisation

Circular dichroism spectroscopy was used to investigate the overall structure of the domains and the effects of replacing iron with cadmium. The far-UV spectrum of the iron-containing C-terminal domain has a large negative band at 204 nm, a smaller negative band at 225 nm and a positive transitions at 210-220 nm and 190-200 nm (Figure 3.8). The spectrum is very similar to that reported for the 1Fe rubredoxin of *D. gigas* rubredoxin (Henehan *et al.*, 1993), *P. furiosus* (Cavagnero *et al.*, 1995) and *C. tepidum* (Yoon *et al.*, 1999), suggesting overall structural similarity with these proteins. Also, like the *D. gigas* rubredoxin, the C-terminal domain of *Pseudomonas oleovorans* rubredoxin retains structure on

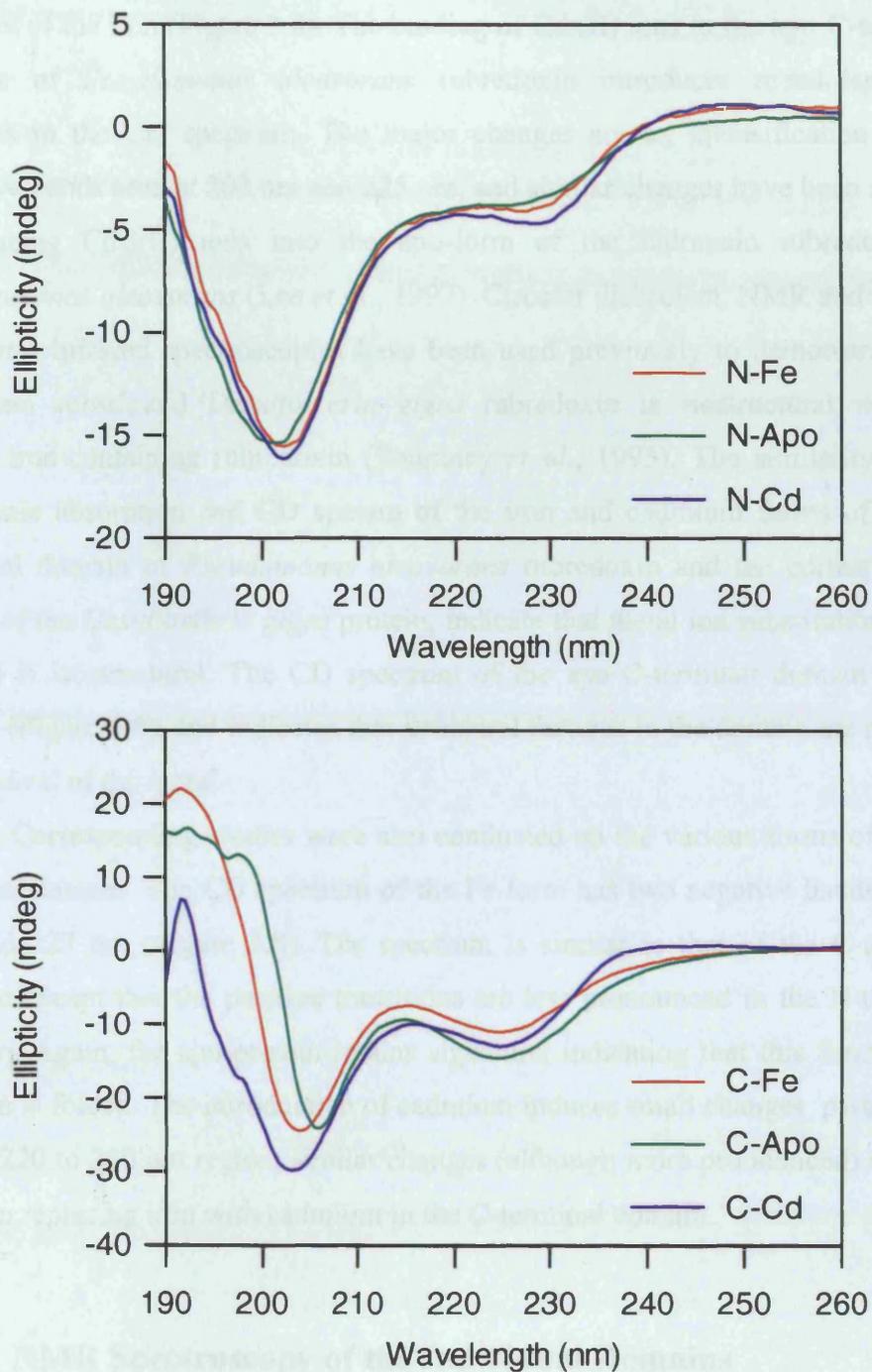


Figure 3.8. Far-UV CD spectra of the Fe- and Cd-substituted N-terminal and C-terminal domains of *Pseudomonas oleovorans* rubredoxin. Spectra were recorded with a 1mm pathlength at a protein concentration of 0.2 mg/ml in 50 mM potassium phosphate buffer, pH 7.0, at 20 °C. Ten scans were averaged for each protein sample.

removal of the iron (Figure 3.8). The binding of Cd (II) ions to the apo-C-terminal domain of *Pseudomonas oleovorans* rubredoxin introduces metal-dependent changes in the CD spectrum. The major changes are an intensification of the negative bands seen at 202 nm and 225 nm, and similar changes have been seen on introducing Cd (II) ions into the apo-form of the didomain rubredoxin of *Pseudomonas oleovorans* (Lee *et al.*, 1997). Circular dichroism, NMR and Fourier transform infrared spectroscopies have been used previously to demonstrate that cadmium substituted *Desulfovibrio gigas* rubredoxin is isostructural with the native iron-containing rubredoxin (Pountney *et al.*, 1995). The similarity of the electronic absorption and CD spectra of the iron and cadmium forms of the C-terminal domain of *Pseudomonas oleovorans* rubredoxin and the corresponding forms of the *Desulfovibrio gigas* protein, indicate that metal ion substitution in the former is isostructural. The CD spectrum of the apo-C-terminal domain is also shown (Figure 3.8), and indicates that structural features in the domain are retained on removal of the metal.

Corresponding studies were also conducted on the various forms of the N-terminal domain. The CD spectrum of the Fe-form has two negative bands at 202 nm and 227 nm (Figure 3.8). The spectrum is similar to that of the C-terminal domain except that the positive transitions are less pronounced in the N-terminal domain. Again, the apo-protein retains signature, indicating that this form of the domain is folded. The introduction of cadmium induces small changes, particularly in the 220 to 240 nm region; similar changes (although more pronounced) are also seen on replacing iron with cadmium in the C-terminal domain.

3.4.3 NMR Spectroscopy of the Individual Domains

1D ^1H NMR spectra were collected for the various forms of the N and C-terminal domains. The spectra acquired for the Cd-substituted C-terminal domain suggests the presence of a stable tertiary structure, including β -sheet secondary structure as evident by the large dispersion of amide resonance (~ 7.0 - 10.0 ppm), and the presence of H^α signals downfield of the water signal (~ 4.7 - 5.5 ppm). This dispersion of resonances and the up-field shift of methyl resonances is similar to that seen for the di-domain protein (unpublished results) and the zinc-substituted

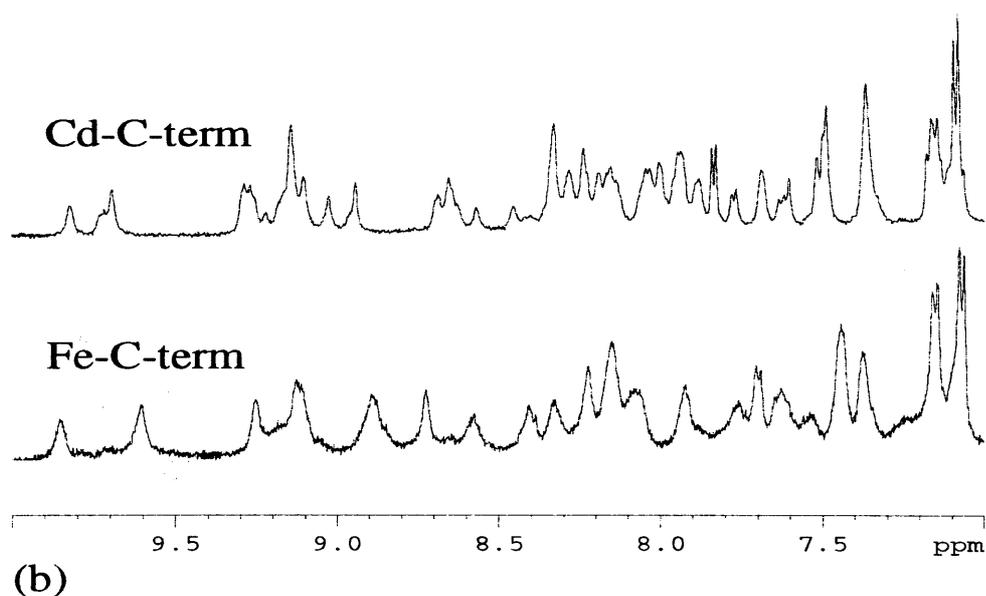
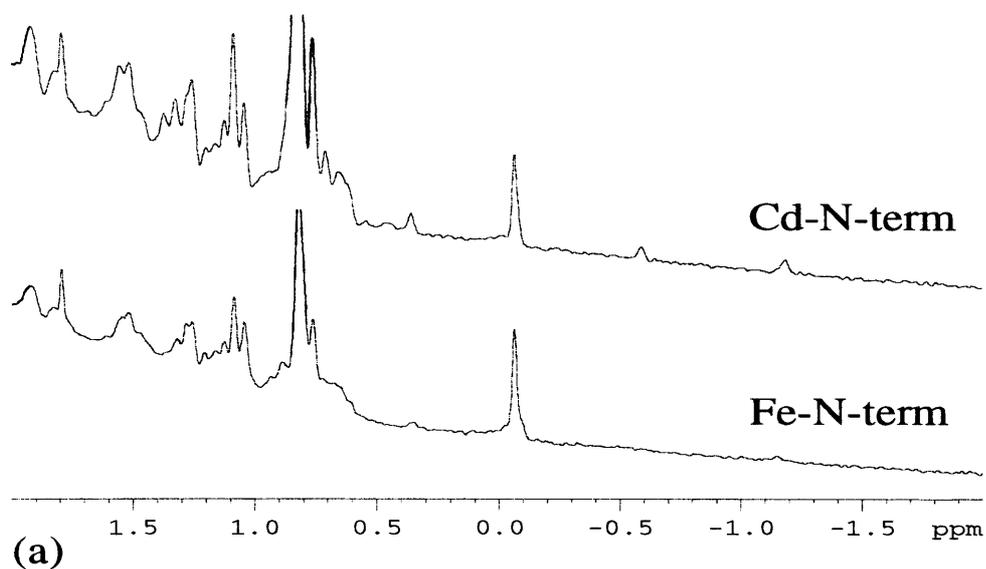


Figure 3.9. 1D ¹H NMR spectra for the Fe- and Cd-substituted N-terminal and C-terminal domains of the *Pseudomonas oleovorans* rubredoxin. Panel a, N-terminal domain rubredoxin. Panel b, C-terminal domain rubredoxin. Spectra were recorded at 25 °C and pH 7.0.

form of *P.furiosus* (Blake *et al.*, 1991). The spectrum for the iron-containing C-terminal domain demonstrates the effect of the Fe (III) on the quality of the spectrum (Figure 3.9). There is significant line broadening and loss of signal in the spectrum of the iron-containing C-terminal domain, due to paramagnetic effects, compared to the spectrum of the Cd-substituted C-terminal. This is similar to effects observed by Blake *et al.* (1991), for Fe (II) containing rubredoxin from *P.furiosus*.

The 1D ^1H NMR spectra of the Cd-substituted N-terminal domain indicates the protein is folded, as shown by the presence of up-field shifted methyl peaks (Figure 3.9). The poor recovery of recombinant N-terminal samples did not yield sufficient protein for a qualitative investigation of secondary structure elements, as was performed for the C-terminal domain. The spectrum of the iron containing N-terminal domain, compared to the Cd-substituted form, displayed line broadening and shifting of peak positions due to paramagnetic effects of the iron.

2D NMR spectra of the C-terminal domain rubredoxin show almost identical overlap with the respective peaks from the di-domain spectra (Figure 3.10). This is evidence that the structure of the C-terminal domain adopts similar conformations when it exists, both as an independent domain, and when present in the di-domain rubredoxin. It also suggests that there is not a major structural interaction between the two domains since such an interaction would perturb the identical shifts, reducing the amount of overlap observed.

The 1D ^{113}Cd NMR spectrum of the C-terminal domain exhibits a single narrow resonance at 732.3 ppm (Figure 3.11). This is in the expected range of 610-750 ppm observed in other tetrahedral tetrathiolate cadmium sites (Summers, 1988). The di-domain rubredoxin contained two ^{113}Cd resonances at 732.2 and 730 ppm with the peak at 730 ppm being broader than that at 732.2 ppm, due to chemical exchange of the cadmium (Lee *et al.*, 1997). This results demonstrates that the ^{113}Cd resonances observed for the di-domain rubredoxin, at 732.2 and 730 ppm, are from ^{113}Cd bound to the C- and N-terminal domain respectively, and confirms the findings of Lee *et al.* (1997) that the N-terminal domain binds metal less strongly than the C-terminal domain. The position of the H^β chemical shifts for the cysteine residues liganded to the Cd^{2+} are identified from the 2D (^1H - ^{113}Cd) HSQC experiment (Figure 3.11).

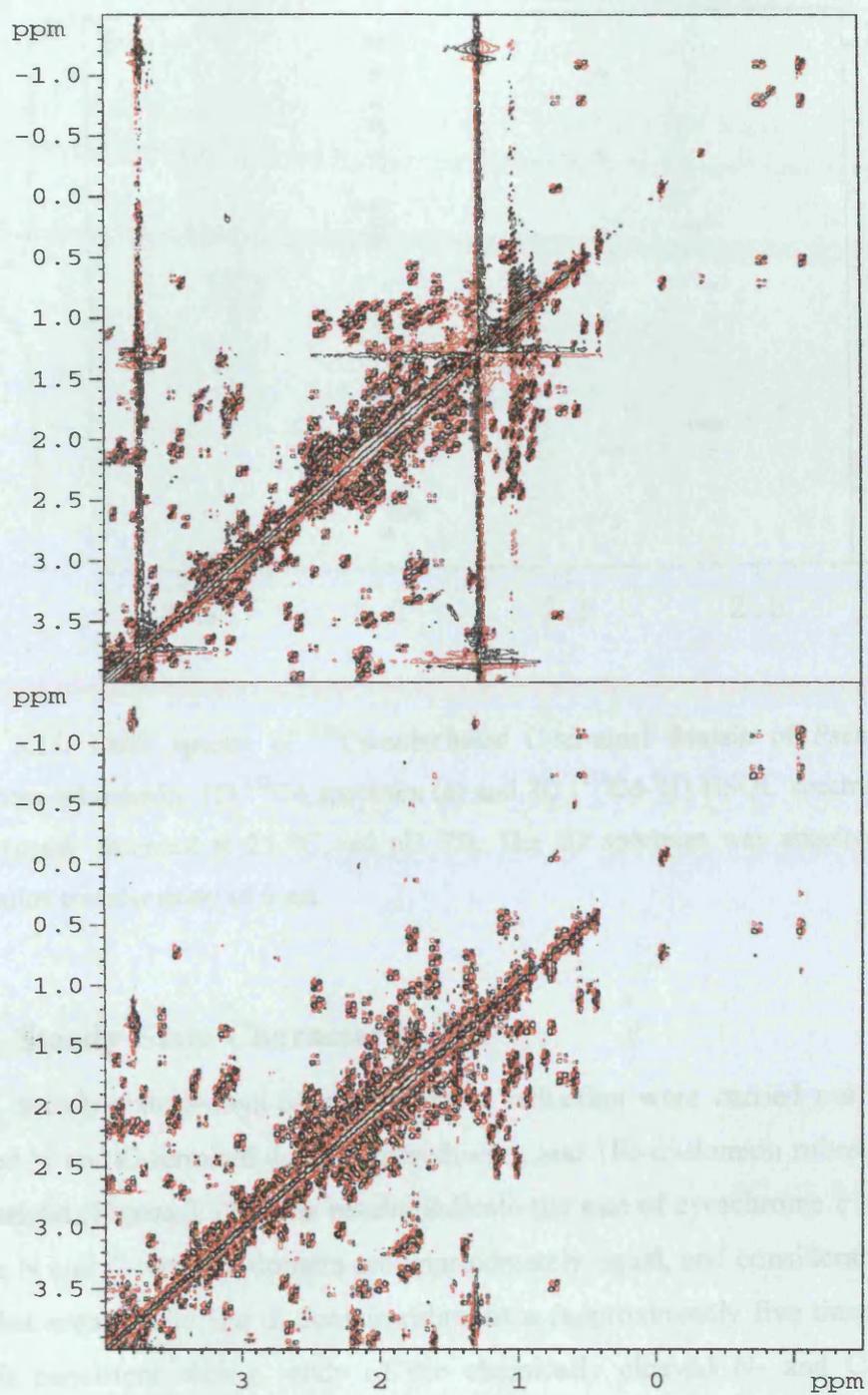


Figure 3.10. 2D COSY spectra for the Cd-substituted C-terminal domain rubredoxin and the Cd-substituted di-domain rubredoxin. Upper panel, di-domain rubredoxin. Lower panel C-terminal domain rubredoxin. Spectra were recorded at 25 °C and pH 7.0.

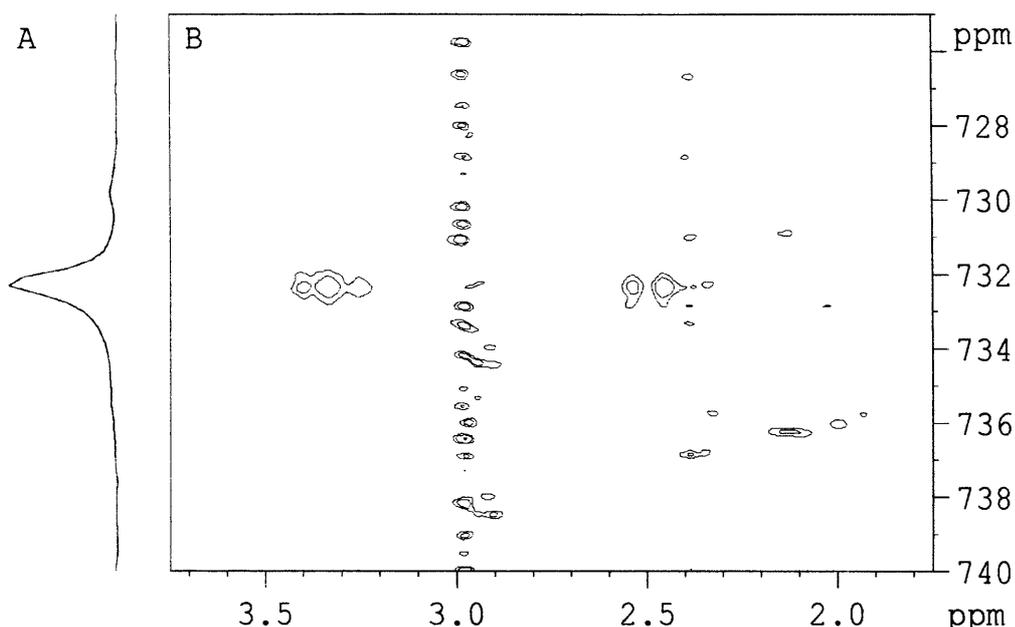


Figure 3.11. NMR spectra of ^{113}Cd -substituted C-terminal domain of *Pseudomonas oleovorans* rubredoxin. 1D ^{113}Cd spectrum (a) and 2D [^{113}Cd - ^1H] HSQC spectrum (b) of *Ps.oleovorans* recorded at 25 °C and pH 7.0. The 2D spectrum was acquired with a polarisation transfer delay of 6 ms.

3.4.5 Steady-State Characterisation

Steady-state assays of cytochrome *c* reduction were carried out with the isolated N and C-terminal domain rubredoxins, and 1Fe-di-domain rubredoxin for comparison (Figure 3.12). The results indicate the rate of cytochrome *c* reduction for the N and C-terminal domain are approximately equal, and considerably lower than that seen for the 1Fe di-domain rubredoxin (approximately five times lower). This is consistent with a study of the chemically cleaved N- and C-terminal domain, where the observed rates for the individual domains were lower than that for the di-domain proteins (approximately three times lower for the C-terminal domain) (Lode & Coon, 1971). In their study the C-terminal domain had greater activity than the N-terminal domain; however the C-terminal domain also contained all of the linker and the N-terminal none.

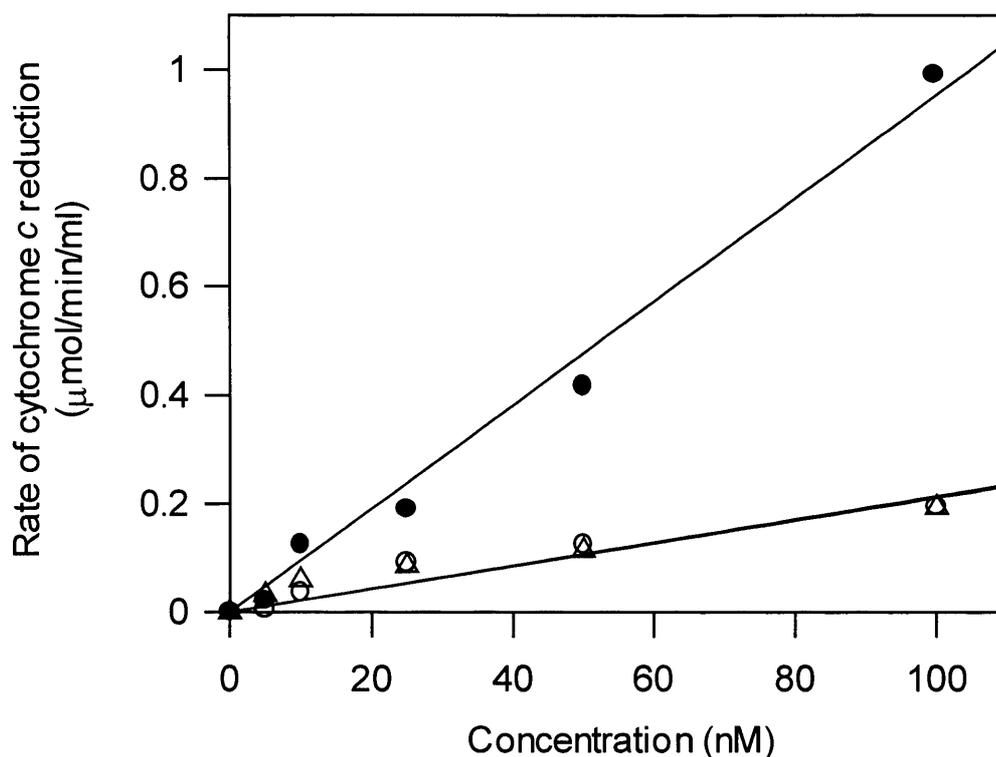


Figure 3.12. Reduction of cytochrome *c* by rubredoxin reductase mediated by the N-terminal and C-terminal domains of rubredoxin and the 1Fe di-domain rubredoxin. Filled circles, 1Fe di-domain rubredoxin; open circles, N-terminal domain; open triangles, C-terminal domain.

3.4.6 Limited Proteolysis

Limited proteolysis was performed on the C-terminal domain to ascertain if the N-terminal region (putative linker sequence) is predominantly unstructured, as predicted from the sequence alignments (Chapter 1). Digestion with trypsin released several bands that were stable for up to 10 minutes. The red coloration of the solution, which is attributed to the ligation of iron by the C-terminal domain, was retained within the first 20 minutes of the digestion with trypsin, but thereafter lost. Electrophoretic analysis by SDS PAGE revealed that beyond 20 minutes digestion, the major high molecular weight band (band X) corresponding to the truncated and structured C-terminal domain is degraded (Figure 3.13). Analysis of the N-terminal sequences of the major digestion products released during the early stages of proteolysis revealed that cleavage occurs within the conjectured linker

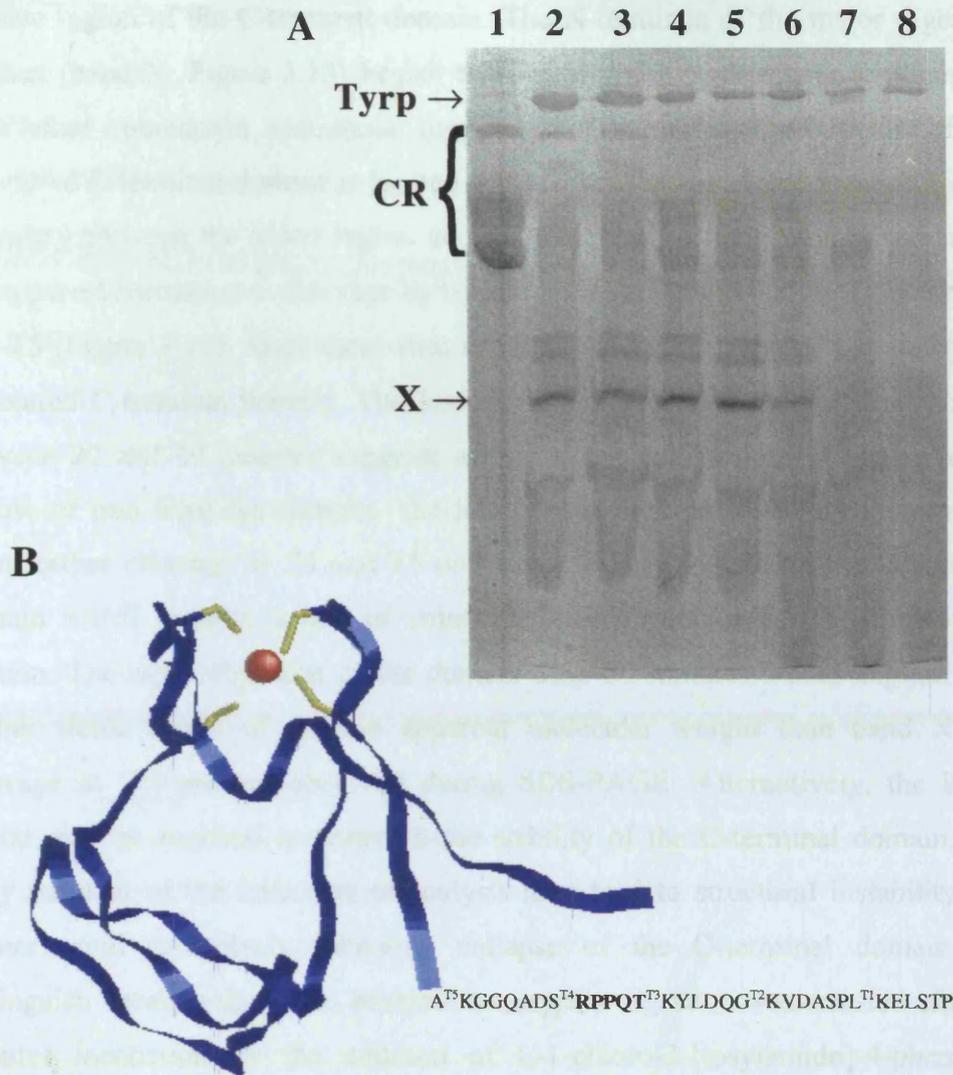


Figure 3.13. Limited proteolysis of the purified C-terminal domain. Panel A, SDS PAGE of proteolysis products. Lane 1, purified C-terminal domain; lanes 2 to 8, samples analysed at various time points throughout the limited proteolysis (0.5, 1, 2, 5, 10, 20 and 40 minutes respectively). The multiple bands, indicated by CR, are the C-terminal domain rubredoxin. The band labelled "tryp" has been identified as trypsin. The band X is the major digestion product referred to in the text. Panel B, schematic of the C-terminal domain rubredoxin. Structured domain shown as blue ribbon, linker region shown as text. T1 to T5 indicate location of potential trypsin cleavage sites. T3 is the cleavage site giving rise to band 'X' in panel A, with amino acids identified by N-terminal sequencing shown in bold. Structure of *Pyrococcus furiosus* rubredoxin used to represent the structured domain (PDB code 1CAD).

peptide region of the C-terminal domain. The N-terminus of the major digestion product (band X; Figure 3.13) begins at residue Thr-106. A sequence alignment with other rubredoxin sequences suggests that the N-terminal region of the structured C-terminal domain is located at Ala-118. This predicted location for the boundary between the linker region and the structured domain might account for the apparent resistance to cleavage by trypsin between 0 and 10 minutes at sites T4 and T5 (Figure 3.13), since these sites are likely afforded some protection by the structured C-terminal domain. The loss of red coloration of the reaction mixture between 20 and 40 minutes suggests a loss of structural stability leading to the release of iron from the domain. The loss of iron beyond 20 minutes may result from further cleavage at T4 and T5 or an exposed trypsin site in the C-terminal domain which leads to a loss of structural integrity and rapid digestion of the domain. The rapid digestion of the domain after 20 minutes would explain why further stable bands of smaller apparent molecular weight than band X (*i.e.* cleavage at T3) are not observed during SDS-PAGE. Alternatively, the linker region may be required to maintain the stability of the C-terminal domain, and early removal of the linker by proteolysis may lead to structural instability and further rapid proteolysis following collapse of the C-terminal domain. To distinguish between these two possibilities, trypsin digestion was arrested after 10 minutes incubation by the addition of L-1-chloro-3-[tosylamido]-4-phenyl-2-butanone (TLCK). Following addition of the inhibitor, the red coloration of the proteolysed C-terminal domain (cleaved at site T3) was stable for several hours, indicating that the absence of the linker region does not impart instability on the structured C-terminal domain.

3.5 Discussion

The work described in this Chapter demonstrates the feasibility of isolating functional individual domain regions of the di-domain rubredoxin of *Ps. oleovorans*. This has been achieved for the C-terminal domain by the direct genetic excision of the N-terminal domain region, and expression of the intact C-terminal domain. The N-terminal region could not be expressed directly, likely reflecting the poor structural stability of this domain, as illustrated by metal binding NMR

studies in the full-length di-domain rubredoxin, and the poor stability of the N-terminal fragment generated by cyanogen bromide cleavage of the di-domain rubredoxin (Lee *et al.*, 1997; Lode & Coon, 1971). The N-terminal domain was thus obtained via expression, purification and proteolytic digestion of a mutant form of the di-domain rubredoxin.

The isolated domains are active in transferring electrons from rubredoxin reductase to horse cytochrome *c* and solution studies indicate that each domain is folded and assembled with the cofactor iron as expected for the simpler 1Fe rubredoxins of other species. These studies reveal that the N-terminal domain is less stable than the C-terminal domain. However, the stability of the N-terminal domain described here is increased compared with that isolated by cyanogen bromide cleavage of the di-domain rubredoxin (Lode & Coon, 1971). The increased stability can be attributed to the additional residues in the putative linker included in this work (residues 2-99 as opposed to 2-51 in the work of Lode & Coon, (1971)). Precise identification of the linker region required for increased stability of the N-terminal domain requires further study since the limited proteolysis studies of the C-terminal domain indicated that a substantial portion of the linker region does not contribute to domain stability.

The quantities of the C-terminal domain obtained are suitable for kinetic studies using stopped-flow methods. Through the use of a larger metal chelating column it should be possible to obtain comparable quantities of the N-terminal domain. Therefore, this work should pave the way for a detailed kinetic analysis of the functional properties of the individual domains, which will provide valuable insight into the electron transfer properties of this system. Additionally, the ability to substitute Cd for Fe (thereby avoiding the problems associated with paramagnetism) in the individual domains will allow detailed study of solution structure of the domains by NMR spectroscopy. Determination of the solution structure of the C-terminal domain is the focus of Chapter 4.

Chapter 4

NMR Analysis of the C-Terminal Domain

4.1 Introduction

The rubredoxin of *Pseudomonas oleovorans* was first isolated over 30 years ago and since has been thoroughly characterised (Lee *et al.*, 1997; May *et al.*, 1984; Peterson *et al.*, 1967). Despite this there is little structural information on this rubredoxin, while the structures of several rubredoxins have been solved by both X-ray crystallography and NMR techniques (Bertini *et al.*, 1998; Blake *et al.*, 1992b; Dauter *et al.*, 1992; Watenpaugh *et al.*, 1979).

Structural work on the *Pseudomonas oleovorans* rubredoxin has been hindered by the low yields achieved following purification from the host strain (Benson *et al.*, 1971). Work by Lee *et al.* (1997), allowed large quantities of the recombinant rubredoxin to be obtained. This protein gave high quality 2D NMR spectra but with considerable overlap, which is predictable for a 20 kDa protein. Attempts at labelling the protein, with both minimal media and pre-labelled media were unsuccessful (Scrutton, unpublished results). With this in mind Chapter 3 described the isolation of the individual domains of the *Pseudomonas oleovorans* rubredoxin. At 10 kDa these proteins are more applicable to structural determinations using 2D NMR.

This chapter describes the determination of the structure of the C-terminal domain *Pseudomonas oleovorans* rubredoxin using NMR techniques. This approach involves the acquisition of 2D NMR spectra, assignment of spin systems, identification of secondary structure elements and structure calculations. The structure calculations are performed using ambiguous restraints for iterative assignments (ARIA) methodology (Nilges *et al.*, 1997; Pascual *et al.*, 1997). ARIA is a suite of programmes and performs an automated assignment of selected peak lists based on chemical shifts, and is able to merge several peak lists (from spectra collected at different experimental conditions). The structure calculations are performed in X-PLOR, using a simulated annealing protocol adapted to interpret ambiguous restraints (Brünger, 1992).

4.2 Assignment Strategy

The assignment strategy follows the standard methods outlined by Wüthrich (Wüthrich, 1986). Correlated spectroscopy (COSY) allows detection of the couplings between hydrogens that are two or three covalent bonds apart, and this is used to identify the H^N - H^α couplings. Every residue exhibits a H^N - H^α coupling in what is known as the fingerprint region with the exception of proline, which lacks a H^N . Glycine, has two H^α , shows two couplings in this region. The H^N - H^α couplings identified in the COSY experiment are used as a starting point for the analysis of the Total correlated spectroscopy (TOCSY) spectrum. TOCSY allows detection of couplings within the same spin system; in protein studies this equates to couplings within the same amino acid residue, the spin systems are classified on the basis of the coupling pattern. The residues can be grouped as those with (i) unique spin systems, such as Ile and Val; (ii) AMX spin systems, which correspond to residues with H^α - $H^{\beta 2}$, $H^{\beta 3}$, such as Asp or Phe and (iii) those with longer spin systems, such as Arg and Lys.

The spin systems are then linked together using nuclear Overhauser effect spectroscopy (NOESY). NOESY reveals dipolar coupling between nuclei less than 5.5 Å apart, and these need not be connected directly by covalent bonds. Hence, peaks occurring in COSY and TOCSY spectra may also appear in the NOESY spectrum, but it is the additional peaks that are of interest as they can only occur by through-space coupling. From the NOESY spectrum it is possible to identify interactions between neighbouring amino acids, allowing identification of specific amino acids through sequence comparisons, a process known as sequence specific assignments. The remaining couplings identified in the NOESY spectrum are those that occur between nuclei, which although not close in sequence, have been brought together (5.5 Å apart or less) by the folding of the protein. From the NOESY spectrum a series of specific pair of hydrogen atoms with associated distance restraints can be built up. It is with this information that the protein structure is calculated.

4.3 Spin system assignments

All the NMR analysis described in this chapter was performed on the Cd-substituted form of the C-terminal domain. This avoids the problem of paramagnetic effects experienced with the iron-containing form of the protein (the Fe^{3+} ion causes line broadening, see Chapter 3). Studies of other rubredoxins have demonstrated that metal substitution does not affect protein structure (Blake *et al.*, 1992a; Blake *et al.*, 1994; Blake *et al.*, 1992b; Day *et al.*, 1992; Richie *et al.*, 1996). Spin system assignments were made using standard methods; COSY and TOCSY spectra were used to identify the characteristic patterns and chemical shifts of the various amino acids (Wüthrich, 1986).

As established in Section 3.3, the C-terminal domain contains 87 amino acids. These amino acids can be classified as follows; 39 residues with unique spin systems (eight Gly, six Ala, four Ile, six Thr, six Leu, seven Pro and two Val); 28 residues with AMX spin systems (corresponding to the $\text{H}^\alpha\text{-H}^{\beta 2}, \text{H}^{\beta 3}$ of ten Asp, five Cys, five Tyr, three Ser, two Phe, two Trp and one His); 20 residues with “longer” spin systems (eight Glu, seven Lys, three Gln and two Arg).

The COSY spectra of the “fingerprint” region contains 44 well-resolved $\text{H}^{\text{N}}\text{-H}^\alpha$ crosspeaks and 20 more in the heavily overlapped area between 8.5 and 7.5 ppm (Figure 4.1). Therefore, as many as 20 $\text{H}^{\text{N}}\text{-H}^\alpha$ crosspeaks are unresolved. This provides assignment difficulties for these unresolved $\text{H}^{\text{N}}\text{-H}^\alpha$ crosspeaks. However comparisons of the COSY and NOESY spectra demonstrates a lack of NOE connectivities to the H^{N} protons of residues in the overlapped region (Figure 4.1). This suggests that these residues are part of a random coil (the chemical shifts also support this) where a lack of structure does not allow sufficient time in any one conformation for NOE connectivities to form. It is possible to speculate that these residues are from the linker region included in the C-terminal domain, but this can be confirmed from the sequential sequence assignments.

Correlation of the $\text{H}^{\text{N}}\text{-H}^\alpha$ crosspeaks with the TOCSY spectra allowed identification of the following unique spin systems: five Gly, three Ala, three Ile, three Thr, two Leu and one Val. The AMX systems belonging to aromatic residues were assigned using the NOESY crosspeaks between the $\text{H}^{\beta 2}, \text{H}^{\beta 3}$ protons and the

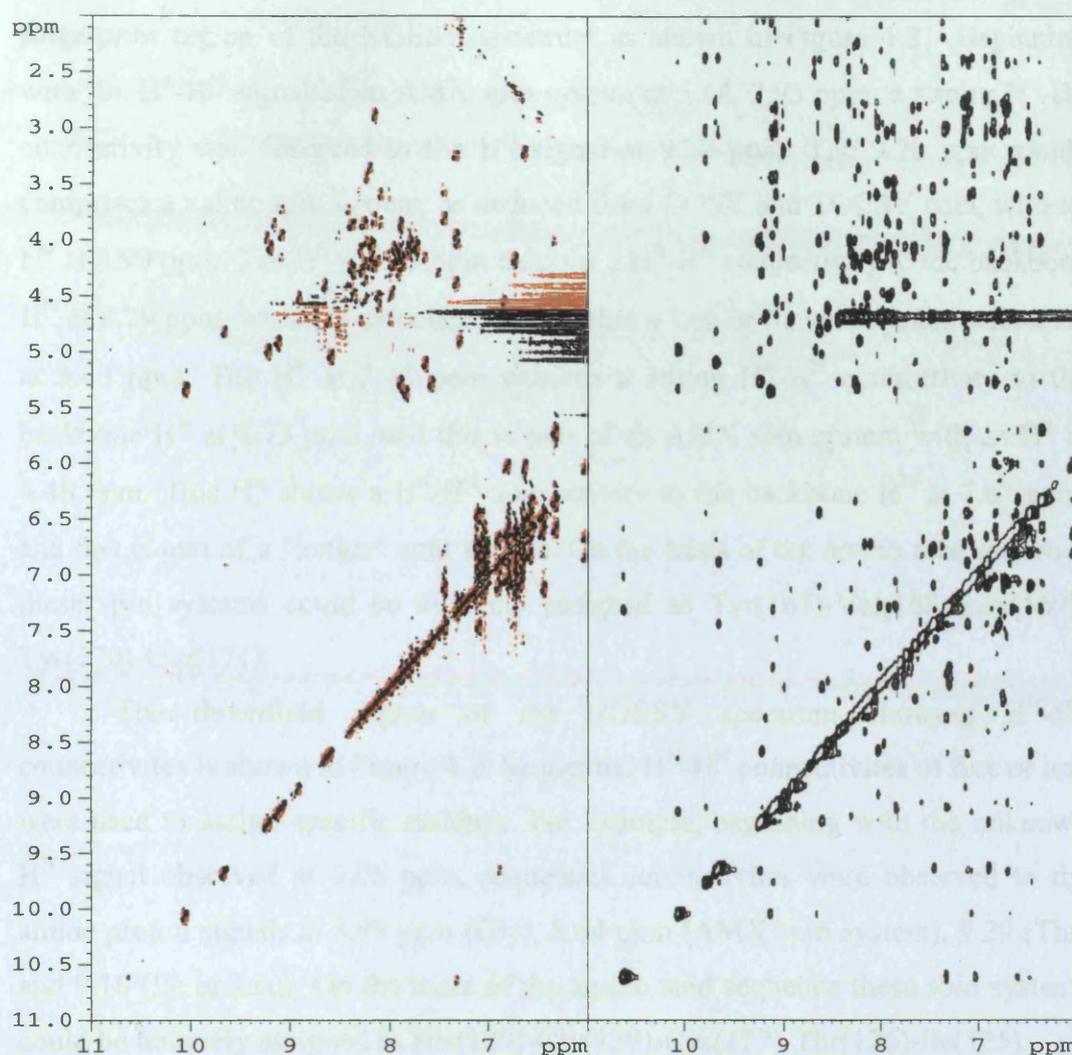


Figure 4.1. Sections of COSY and NOESY spectra. Spectra collected at 25 °C and pH 6.0.

aromatic side chain protons. The remaining spin systems were classified either as AMX or longer spin systems.

4.4 Sequential sequence assignments

Sequential assignments were made using 2D NOESY data collected in 90% H₂O/10% D₂O solution with a mixing time of 100 ms. Sequential sequence assignments were made using H^α-H^N and/or H^N-H^N connectivities.

Several stretches of sequential H^α - H^N connectivities were observed in the fingerprint region of the NOESY spectrum as shown in Figure 4.2. Beginning with the H^α - H^N signal of an AMX spin system at 5.44, 7.93 ppm, a strong H^α - H^N connectivity was observed to the H^N signal at 9.20 ppm. The 9.20 ppm amide comprises a valine spin system, as deduced from COSY and TOCSY data, with an H^α at 4.99 ppm. The H^α at 4.99 ppm exhibits a H^α - H^N connectivity to the backbone H^N at 8.29 ppm, which is associated with either a Leu or Ile spin system with a H^α at 3.43 ppm. The H^α at 3.43 ppm exhibits a strong H^α - H^N connectivity to the backbone H^N at 8.73 ppm, and this is part of an AMX spin system with an H^α at 4.48 ppm. This H^α shows a H^α - H^N connectivity to the backbone H^N at 7.82 ppm, and this is part of a "longer" spin system. On the basis of the amino acid sequence these spin systems could be uniquely assigned as Tyr(167)-Val(168)-Leu(169)-Tyr(170)-Glu(171).

The downfield region of the NOESY spectrum showing H^N - H^N connectivities is shown in Figure 4.2. Sequential H^N - H^N connectivities of five or less were used to assign specific residues. For example, beginning with the unknown H^N signal observed at 9.06 ppm, sequential connectivities were observed to the amide proton signals at 8.09 ppm (Gly), 8.64 ppm (AMX spin system), 9.29 (Thr) and 9.16 (Ile or Leu). On the basis of the amino acid sequence these spin systems could be uniquely assigned as His(129)-Gly(128)-Cys(127)-Thr(126)-Ile(125).

All the residues assigned fall within the region predicted to form the structured domain by sequence alignments (Chapter 1). The resonances of the 32 residues that form the putative linker region were not assigned. This is due firstly to these residues giving rise to heavily overlapping resonances, especially in the fingerprint region. Secondly there is a significant lack of NOE connectivities, most notably H^N - H^N , involving the backbone protons of residues in the overlapped region (Figure 4.1). Although this has provided assignment difficulties for these residues, it is also strong evidence that they have little or no structure. The H^α - H^N chemical shifts do not show the dispersion expected if they were part of a structured region, which gives rise to the extensive overlap. Additionally, the lack of H^N - H^H connectivities indicate that these residues do not spend sufficient time in any one conformation for an NOE to develop.

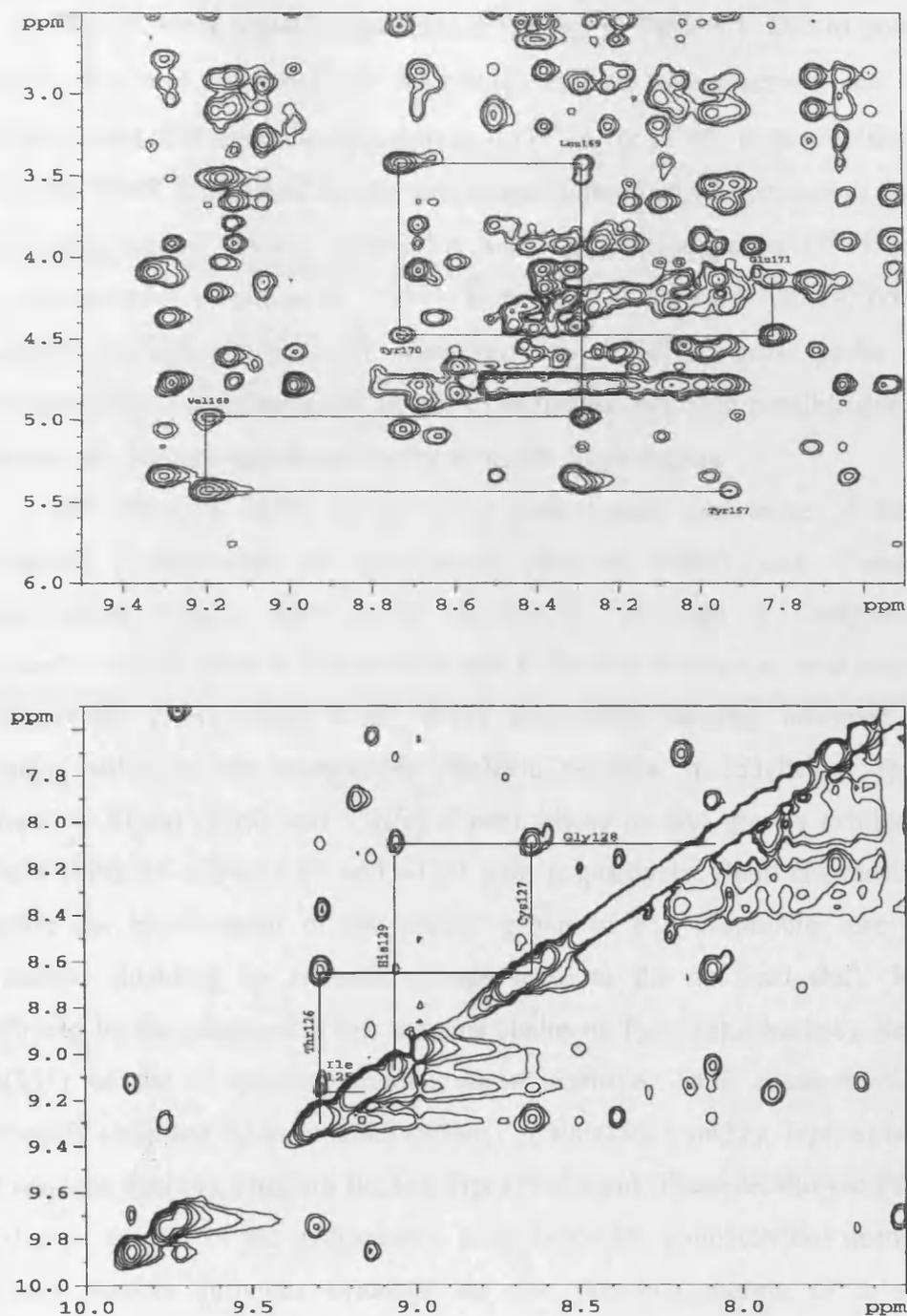


Figure 4.2. NOESY spectra for the C-terminal domain connectivities. Upper panel, sequential H^α - H^N connectivities for residues Tyr(167)-Glu(171). Residue type and number are shown above or beneath the H^α - H^N crosspeaks. Lower panel, sequential H^N - H^N connectivities for residues Ile(125)-His(129). Residue type and number are shown on lines corresponding to the H^N chemical shifts. Spectra collected at 25 °C and pH 6.0 with a mixing time of 100 ms.

Chapter 4: NMR Analysis of the C-Terminal Domain

The ^1H NMR signal assignments are shown in Table 4.1. Due to excessive overlap only one resonance for a proline residues was observed, the H^α of Pro(144). Asp(153) has no assignments as no $\text{H}^\alpha\text{-H}^{\text{N}}$ or $\text{H}^{\text{N}}\text{-H}^{\text{N}}$ connectivities were identified. NMR analysis of the Zn-substituted rubredoxin of *Pyrococcus furiosus* (Pfrd) demonstrated that the comparable Asp(34) lacked sequential $\text{H}^\alpha\text{-H}^{\text{N}}$ or $\text{H}^{\text{N}}\text{-H}^{\text{N}}$ connectivities (Blake *et al.*, 1991). However, for the Pfrd Asp(34) could be identified through a process of elimination as all $\text{H}^{\text{N}}\text{-H}^\alpha$ cross peaks in the fingerprint region were assigned. In this work this has not been possible due to the presence of ~30 unassigned resonances from the linker region.

The chemical shifts of the C-terminal domain and those of the zinc substituted rubredoxins of *Pyrococcus furiosus* (Pfrd) and *Clostridium pasteurianum* (Cprd) show some similarities, although a comprehensive comparison of the shifts is not possible due to the different amino acid sequences (Blake *et al.*, 1991; Richie *et al.*, 1996). Particularly striking, however, is the chemical shifts of the comparable aliphatic residues at 151(Ile) (C-terminal domain), 32(Leu) (Pfrd) and 33(Ile) (Cprd) whose methyl groups exhibit large upfield shifts of -1.23, -1.61 and -1.29 ppm respectively. Such chemical shifts indicate the involvement of the methyl group in a hydrophobic core where substantial shielding by aromatic groups leads to the up field shift. This is confirmed by the observation that the side chains of Tyr(131), Phe(148), Ile(151), Trp(155) of the C-terminal domain show extensive NOE connectivities; as previously observed between residues Tyr(12), Phe(29), Leu(32), Trp(36) of Pfrd; and residues Tyr(13), Phe(30), Ile(33), Trp(37) of Cprd. These residues in Pfrd and Cprd exist as part of the hydrophobic core, hence the connectivities seen in C-terminal domain provides evidence for the presence therein of a similar hydrophobic core.

Although more complete ^1H chemical shift assignments were obtained for Cprd and Pfrd, in the C-terminal domain there is extensive overlap from resonances belonging to the linker region, limiting the ^1H assignments to those shown in Table 4.1 (Blake *et al.*, 1991; Richie *et al.*, 1996). The overlap in the ^1H region (3.5-1.0 ppm) made it extremely difficult to assign ^1H resonances, and beyond, for a number of residues.

Table 4.1. ^1H HMR chemical shifts of Cd-substituted Crub at 25 °C and pH 6.0^a.

^b Residue	H ^N	^c H ^α	^d H ^β	Others
118	Ala			
119	Tyr	8.38	4.80	3.21, 2.86 H ^{δ1,2} 7.20; H ^{ε1,2} 6.80
120	Leu	8.98	4.59	1.77, 1.19 H ^γ 0.77
121	Lys	7.98	5.35	1.85, 1.59
122	Trp	9.30	5.06	2.93, 2.79 H ^{ε1} 9.72; H ^{ε2} 7.52; H ^{η2} 7.17; H ^{ε3} 6.98; H ^{δ1} 6.96; H ^{ε3} 6.74
123	Ile	10.14	5.39	1.21 (H ^{γ2}) ₃ 0.85; H ^{γ12} 0.73; H ^{γ13} 0.61; (H ^{δ1}) ₃ 0.42
124	Cys	8.29	3.52	2.95, 2.33
125	Ile	9.16	4.03	2.08 H ^{γ12} 1.59; H ^{γ13} 1.46; (H ^{γ2}) ₃ 1.05; (H ^{δ1}) ₃ 0.85
126	Thr	9.29	4.14	4.38 (H ^{γ2}) ₃ 1.16
127	Cys	8.64	5.10	3.39, 2.47
128	Gly	8.09	4.21, 3.67	
129	His	9.06	4.16	
130	Ile	7.73	5.17	1.45 H ^{γ12} 1.61; (H ^{γ2}) ₃ 0.64; H ^{γ13} 0.49; (H ^{δ1}) ₃ 0.42
131	Tyr	9.76	4.89	3.02, 2.86 H ^η 8.69; H ^{δ1,2} 7.17; H ^{ε1,2} 6.52
132	Asp	8.68	4.79	2.87, 2.14
133	Glu	8.50	3.81	2.49, 1.96
134	Ala	8.22	2.97	1.25
135	Leu	6.81	4.25	1.70, 1.52
136	Gly	8.17	3.92, 3.67	
137	Asp	9.13	4.77	3.05, 2.90
138	Glu	9.28	3.93	2.10
139	Ala	8.39	4.18	1.59
140	Glu	7.38	4.07	1.95, 1.48
141	Gly	7.53	3.90, 3.63	
142	Phe	7.64	5.35	3.11, 2.90 H ^{δ1,2} 6.90; H ^{ε1,2} 6.69; H ^ε 6.57
143	Thr	8.31	4.08	4.24 (H ^{γ2}) ₃ 1.30
144	Pro		3.56	
145	Gly	7.96	4.18, 3.13	
146	Thr	7.38	4.10	4.09 H ^{γ1} 6.43; (H ^{γ2}) ₃ 0.82
147	Arg	9.33	4.05	2.12, 1.72
148	Phe	9.86	3.37	2.46, 1.96 H ^ε 6.64 H ^{ε1,2} 6.39; H ^{δ1,2} 5.76
149	Glu	9.14	3.80	2.14, 1.92
150	Asp	7.40	4.46	2.67, 2.52
151	Ile	6.71	3.30	0.57 H ^{γ13} -0.98; H ^{γ12} 0.37; (H ^{γ2}) ₃ -0.19; (H ^{δ1}) ₃ -1.23
152	Pro			
153	Asp			
154	Asp	8.19	4.57	2.88, 2.56
155	Trp	7.72	4.23	3.10, 2.90 H ^{ε1} 10.61; H ^{δ1} 7.39; H ^{ε2} 7.11; H ^{η2} 6.87; H ^{ε3} 6.71; H ^{ε3} 6.09
156	Cys	6.92	4.61	2.31, 2.21
157	Cys	9.14	3.51	3.11, 2.91
158	Pro			
159	Asp	8.08		
160	Cys	7.35	4.97	3.29, 2.53
161	Gly	7.91	4.21, 3.56	
162	Ala	9.18	4.23	1.58
163	Thr	8.16	4.68	4.84 (H ^{γ2}) ₃ 1.32
164	Lys	8.59	4.04	1.97, 1.86
165	Glu	8.05	4.31	2.40 2.16 1.99

^b Residue	H ^N	^c H ^α	^d H ^β	Others
166 Asp	8.07	4.79		
167 Tyr	7.93	5.44	3.68, 2.51	H ^{δ1,2} 7.13; H ^{ε1,2} 6.76
168 Val	9.20	4.99	2.32	(H ^{γ1}) ₃ 1.00; (H ^{γ2}) ₃ 0.85
169 Leu	8.29	3.43	1.22, 1.10	H ^γ 0.98; (H ^{δ1}) ₃ 0.41; (H ^{δ2}) ₃ 0.28
170 Tyr	8.73	4.48	2.60, 2.48	H ^{δ1,2} 6.87; H ^{ε1,2} 6.73
171 Glu	7.82	4.14		1.96 1.78 1.65
172 Glu	7.95	3.94	1.91, 1.74	
173 Lys	7.87			1.75 1.20

^a relative to DSS at 0 ppm

^b Only residues 118-173 shown for clarity.

^c H^{α2} & H^{α3} for glycine residues.

^d (H^β)₃ for alanine; H^β for threonine, isoleucine and valine; H^{β2} & H^{β3} for all other residues.

4.5 Secondary structure identification

The secondary structure elements of the C-terminal domain are summarised in Figure 4.3. Medium to strong H^N-H^N sequential connectivities, and in some cases H^N-H^{N(i+2)}, observed for residues Ile(125)-Gly(128), Glu(133)-Gly(136), Glu(138)-Phe(142), Phe(148)-Ile(151), Asp(159)-Ala(162), Lys(164)-Tyr(167) indicate the presence of tight turns.

The strong H^α-H^N connectivities observed for residues Leu(120)-Ile(125), His(129)-Glu(133) and Tyr(167)-Glu(171) indicate that these residues exist in a regular conformation. Characteristic long range H^N-H^N, H^α-H^N and H^α-H^α connectivities, together with hydrogen bonds indicated by slowly exchanging N^H protons, demonstrate that these residues comprise a three stranded anti-parallel β-sheet as shown in Figure 4.4.

The secondary structure elements of the C-terminal domain are similar to those observed in the NMR studies of Pfrd and Cprd. Residues Ala(1)-Lys(6); Tyr(10)-Glu(14); Phe(48)-Glu(52) of Pfrd; residues Lys(3)-Cys(6); Tyr(11)-Asn(14); Phe(49)-Val(52) of Cprd; and residues Leu(120)-Ile(125); His(129)-Glu(133); Tyr(167)-Glu(171) of the C-terminal domain comprise a three-stranded anti-parallel β-sheet. Sequence alignment of these proteins demonstrates that the β-sheets occur in the same position within the proteins (Figure 4.5). Also, comparable β-sheet connectivities to those shown in Figure 4.4 are observed in Pfrd and Cprd.

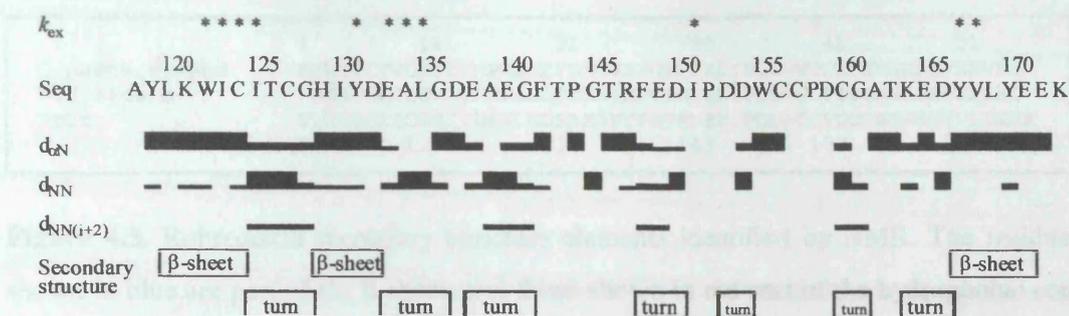


Figure 4.3. Summary of sequential H^{α} - H^N , H^N - H^N and H^N - $H^{N(i+2)}$ connectivities observed in the NOESY spectra of the C-terminal domain. The thickness of the bar reflects the relative intensity of the crosspeak observed. Secondary structure elements are shown at the bottom. Above the amino acid sequence * indicates residues with slow hydrogen exchange rates at the backbone amide after 24 hours incubation at room temperature in D_2O .

Structural changes from NOE connectivities alone it would be impossible to suggest for the C-terminal domain will have a structure to other

4.6 Structural Analysis

4.6.1 NOESY Spectra

The NOESY spectra of the C-terminal domain are shown in Figure 4.3. The

NOESY spectra show crosspeaks between protons that are close to each other in

the protein structure. The crosspeaks are shown in Figure 4.3. The

crosspeaks are shown in Figure 4.3. The crosspeaks are shown in Figure 4.3.

Figure 4.4. Schematic representation of the three-stranded anti-parallel β -sheet observed in the C-terminal domain. Double-headed arrows indicate observed NOE connectivities, sequential H^{α} - H^N connectivities shown in red, sequential H^N - H^N connectivities shown in blue and non sequential connectivities shown in green. Slowly exchanging H^N protons involved in hydrogen bonds are shown in bold, with hydrogen bonds shown as grey bands.

	1	11	21	31	41	51
C. pasteurianum	MKKYTCTVCGYIYNPEDGDPDNGVNP	GTDFKDI	PDDWVCPLCGV	GKDKQ	FEEV	EE-
P. furiosus	-AKWVCKICGYIYDE	DAGDPDNGIS	PGTKF	EELPDDWVC	PICGAPKSE	FEKLED-
PoCt	YLKWICITCGHIYDE	ALGDEAEGFT	PGTRF	EEDI	PDDWCCPDCG	ATKEDYVLYEEK
	119	129	139	149	159	169

Figure 4.5. Rubredoxin secondary structure elements identified by NMR. The residues shown in blue are part of the β -sheet, and those shown in red part of the hydrophobic core (Blake *et al.*, 1991; Richie *et al.*, 1996). PoCt, is the C-terminal domain sequence.

The turns seen in the C-terminal domain are also akin to those seen in Pfrd and Cprd. There are tight turns in all proteins between β -sheets one and two, after β -sheet two and prior to β -sheet three. An additional structural element is the hydrophobic core, which is present in all proteins. Therefore, on the basis of the structural findings from NOE connectivities alone it would be reasonable to suggest that the C-terminal domain will have a similar structure to other rubredoxins (Pfrd and Cprd being structurally typical of rubredoxins).

4.6 Structural restraints

4.6.1 NOE connectivities

Three peak lists were used to generate NOE structural restraints. Two of these peak lists were selected using automated peak picking programmes from NOESY spectra collected at 298 K and 310 K for a 10 mM sample at pH 6.0. Also a peak list was selected manually from the 298 K data. Manually selecting a peak list allowed the input of overlapping peaks, which can be selected as a single peak by the automated method. The peak lists and chemical shift table were converted into ARIA/X-PLOR format (Nilges *et al.*, 1997).

4.6.2 Hydrogen bonds

Using data from the amide exchange experiment (Figure 4.3) and through analysis of the NMR structures, 7 hydrogen bonds were identified. Six of these bonds are shown in Figure 4.4, with an additional bond between the backbone CO

of Lys(164) and amide proton of Tyr(167). A summary of the hydrogen bonds is shown in Table 4.2. The hydrogen-acceptor distance was restrained by $1.95 \pm 0.25\text{\AA}$, and donor-acceptor distance by $2.95 \pm 0.25\text{\AA}$ (Nilges *et al.*, 1997). The use of two distance restraints for each hydrogen bond ensures that the donor-hydrogen-acceptor angle is always larger than 110° . The four remaining slow exchanging hydrogens have not been assigned to hydrogen bonds. These residues exist in regions without well-defined secondary structure, which makes it difficult to identify the acceptor group. It may also be the case that these residues are buried. In the case of Ile(151) the side chain groups exists in a hydrophobic core, this hydrophobicity could extend to the amide group.

Table 4.2. Amino acids involved in hydrogen bonds.

Donor	Acceptor
Trp(122)	Tyr(131)
Ile(123)	Val(168)
Cys(124)	His(129)
Ile(125)	Asp(166)
Tyr(131)	Trp(122)
Tyr(167)	Lys(164)
Val(168)	Ile(123)

4.6.3 Metal binding sites

Restraints were used to enforce the Cd-S co-ordinate bond in the metal binding site. The Cd-S distance was restrained by $2.5 \pm 0.05\text{\AA}$, and the Cd-C ^{β} by $3.5 \pm 0.25\text{\AA}$ (C ^{β} refers to the beta carbon of the cysteine residue) (Baleja *et al.*, 1997). These restraints were used for cysteines 124, 127, 157 and 160. The use of two distance restraints for each cysteine ensures that the S-Cd-S angle remains within allowed limits.

4.7 Structure generation

The lack of identified resonances for residues 87-118 results in no structural restraints being placed on these residues. Therefore, to reduce processing time, with respect to the simulated annealing protocol, only residues 118-173 were

included for the structural calculations. Structures were generated by implementing the ARIA suite of programmes (Nilges *et al.*, 1997; Pascual *et al.*, 1997). Structures in iteration zero were generated using a modelled C-terminal domain structure as input (generated using MODELLER-4, with the high resolution (>1.5 Å) crystal structures of *Clostridium pasteurianum* (Dauter *et al.*, 1996), *Desulfovibrio gigas* (Frey *et al.*, 1987), *Desulfovibrio vulgaris* (Adman *et al.*, 1991), and *Pyrococcus furiosus* (Bau *et al.*, 1998), rubredoxins as input). Subsequent iterations began by ordering the structures from the previous iteration. The best seven structures (selected on the basis of total energy) were used to select restraints and the best ten structures used as a starting structure. An additional ten structures were calculated using the model as the starting structure, thus 20 structures were generated in each iteration. All the starting structures are randomised along the backbone prior to the simulated annealing protocol. For the final iteration 100 structures were generated and the 20 structures of lowest energy were selected as the NMR ensemble. The NOE restraints used to generate the final structures were checked manually for errors, and those restraints excluded were also scrutinised to ensure no obvious restraints had been excluded i.e. those selected manually. The analysis of excluded NOE restraints indicated that they fall into three categories real NOE connectivities, presumably originating from unassigned hydrogen's, noise peaks and inaccurately picked peaks (mainly in areas of high peak density).

During the iterative process the structures show increasing levels of convergence, as illustrated by the decreasing RMSD values, relative to the mean co-ordinates, shown in Figure 4.6 and Table 4.3. The exception is iteration 6 where the RMSD values are larger than that of iteration 5 (Table 4.3). This may represent the structure/structural calculations overcoming large local energy minima. The convergence of structures coincides with an increase in the number of unambiguous restraints (from 212 to 737) and a decrease in the number of ambiguous restraints (from 1732 to 141). This is typical of ARIA-based structure determination and is a function of the expulsion of violated restraints during the selection process, and a tightening up of the ambiguous restraint selection (Nilges *et al.*, 1997). This is borne out by the decrease in the NOE and bond angle RMSD values, which fall as violated restraints are deselected (Figure 4.6).

Table 4.3. Structure statistics.

Itn ^a	Restraints					RMSD	
	Unamb ^b	Ambig ^c	H-bonds ^d	Co ^e	Total ^f	Backbone ^g	Heavy atom ^h
0	212	1723	14	8	1957	4.02	4.58
1	210	1635	14	8	1867	3.82	4.22
2	324	1404	14	8	1750	3.06	3.65
3	477	805	14	8	1304	2.10	2.59
4	562	559	14	8	1143	1.17	1.63
5	630	386	14	8	1038	0.88	1.32
6	683	271	14	8	976	0.93	1.42
7	699	227	14	8	948	0.77	1.31
8	739	145	14	8	906	0.67	1.18
F	737	141	14	8	900	0.65	1.17

^a Iteration number.

^b Number of unambiguous restraints.

^c Number of ambiguous restraints.

^d Number of hydrogen bond restraints.

^e Number of co-ordination bond restraints.

^f Total number of restraint used to generate the structures.

^{g,h} Backbone and heavy atom RMSD relative to the mean co-ordinates for the seven structures of lowest energy, calculated with XPLOR (Brünger, 1992).

The structure selected as the NMR ensemble (the 20 structures of lowest energy in iteration F; shown in Figure 4.7) were generated using 878 NOE based restraints (737 unambiguous and 141 ambiguous), 14 restraints for hydrogen bonds and the 8 restraints used to simulate the metal co-ordination bonds. This is a total of 900 restraints, which approximates to 16 restraints per residue. This is a reasonable number of restraints per residue for a NMR structure, and is comparable to the number of restraints per residue used to calculate the NMR structure of the rubredoxin from *Pyrococcus furiosus* (Blake *et al.*, 1992b). Additionally, the RMSD values relative to the mean co-ordinates of 0.71 and 1.29, for backbone and heavy atom respectively, are reasonable for an NMR structure (Blake *et al.*, 1992b; Nilges *et al.*, 1997; Pascual *et al.*, 1997). The NMR ensemble for the C-terminal domain is shown in Figure 4.7. A detailed analysis of the C^α RMSD and the distribution of restraints indicates a good correlation between the two, as shown in Figure 4.8. The C^α RMSD is high in areas with a low number of restraints, and is highest among the core residues for Pro(152) which lack any restraints and is flanked by Asp(153) which also lack restraints. The C^α RMSD is high for the residues at the start and end of the molecule, which is typical for NMR structures.

The areas of lowest C^α RMSD are those involved in β -sheets formation, and again this would be expected.

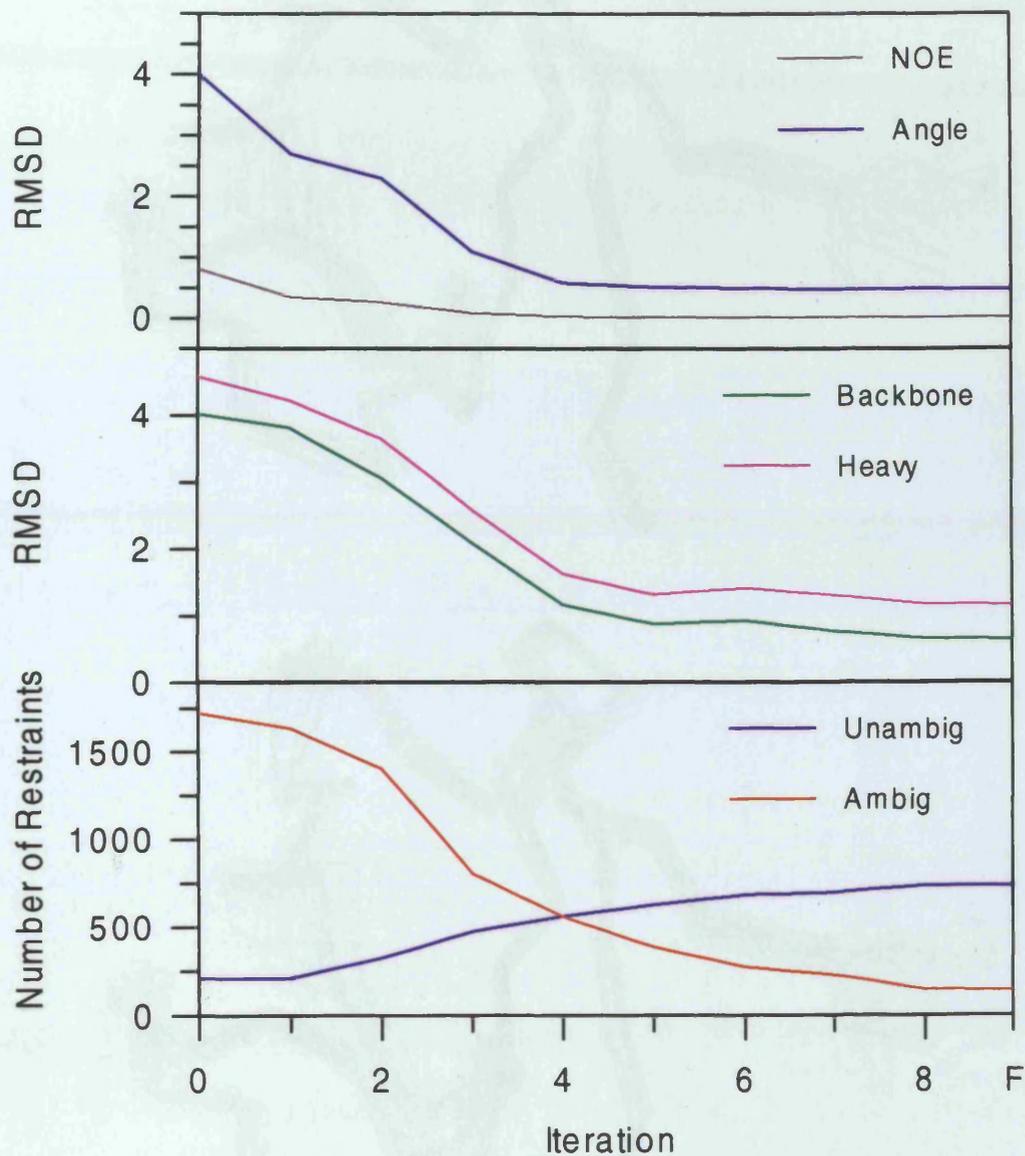


Figure 4.6. Statistics from structure generation. The RMSD values for NOE and bond angles were calculated with XPLOR, as were the backbone and heavy atom RMSD values relative to the mean co-ordinates, which were calculated for the seven lowest energy structures of each iteration (Brünger, 1992).



Figure 4.7. The ensemble of 20 NMR structures for the C-terminal domain rubredoxin. Taken from the 20 lowest energy structures from the final iteration.

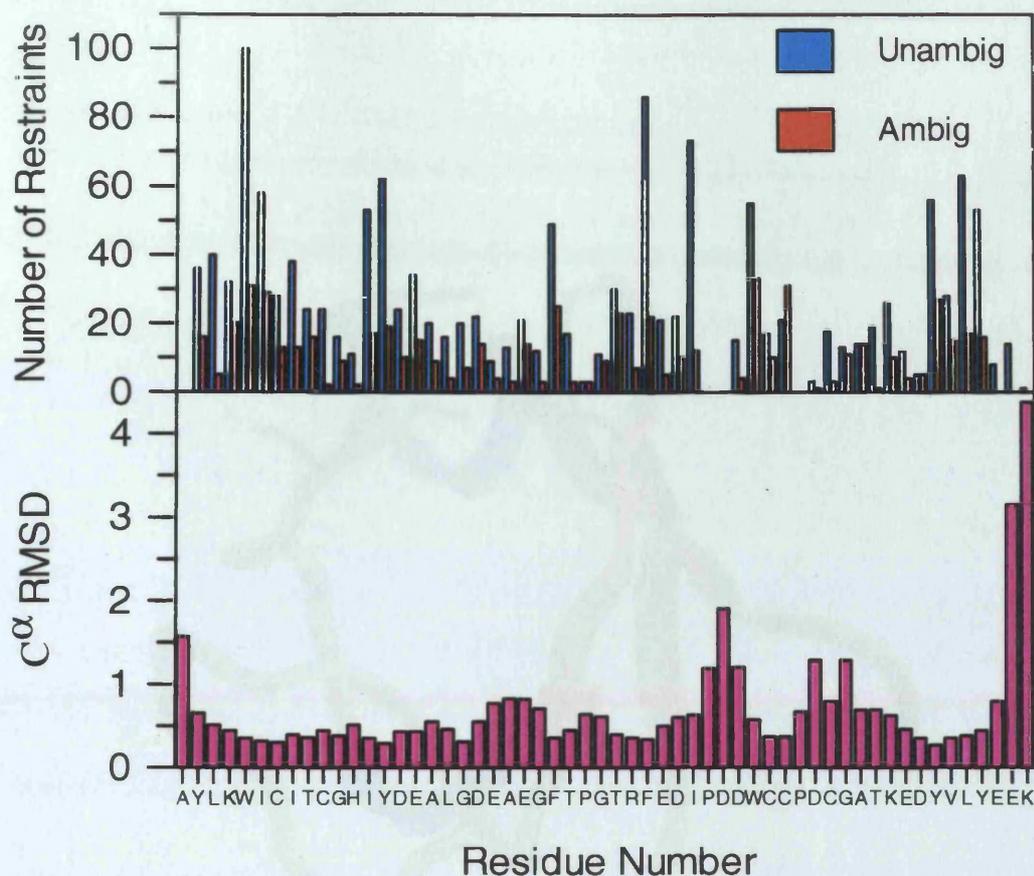


Figure 4.8. Plots of structural parameters. Upper panel, the distribution of unambiguous and ambiguous restraints. Lower panel, C^α RMSD relative to the mean co-ordinates calculated with XPLOR (Brünger, 1992).

4.8 Structure description and analysis

The structure determined for the C-terminal domain is shown in Figure 4.9. It contains a three-stranded anti-parallel β -sheet and a number of tight turns, some showing helical character (revealed by analysis of the NMR ensemble by PROCHECK-NMR; Laskowski *et al.*, 1996). The β -sheets were identified during the signal assignments. However, there were no characteristic NOE patterns indicating the presence of helical secondary structure. This is possibly due to the helical regions being relatively short, at two and three residues respectively. The structure is typical of rubredoxins, a fact borne out by the comparative backbone

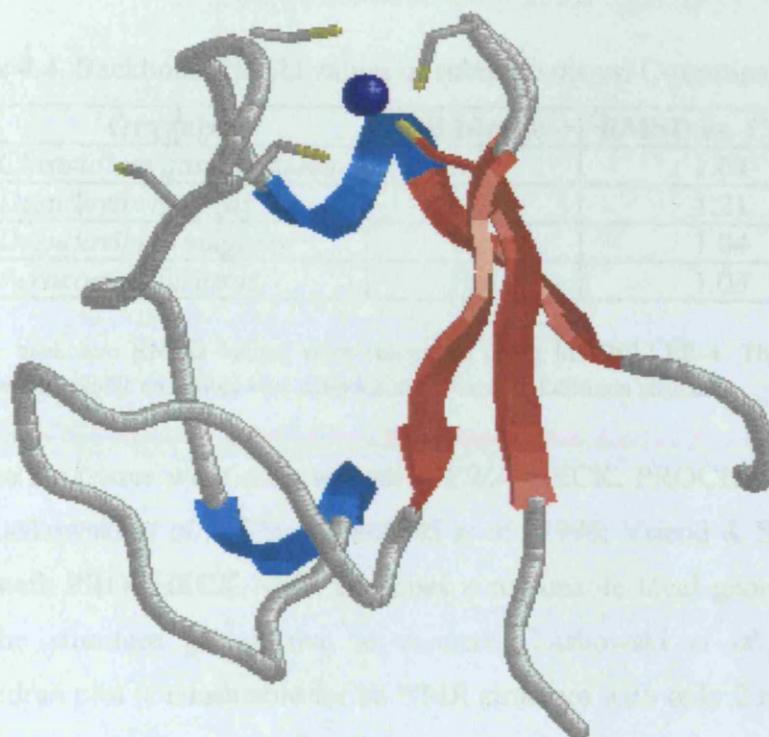


Figure 4.9. The structure of the C-terminal domain. The cadmium is shown in blue. β -sheets shown in red. Helical regions shown in cyan. Cysteines shown as wireframes with the sulphur atoms shown in yellow. Structure shown is the lowest energy structure from the NMR ensemble.

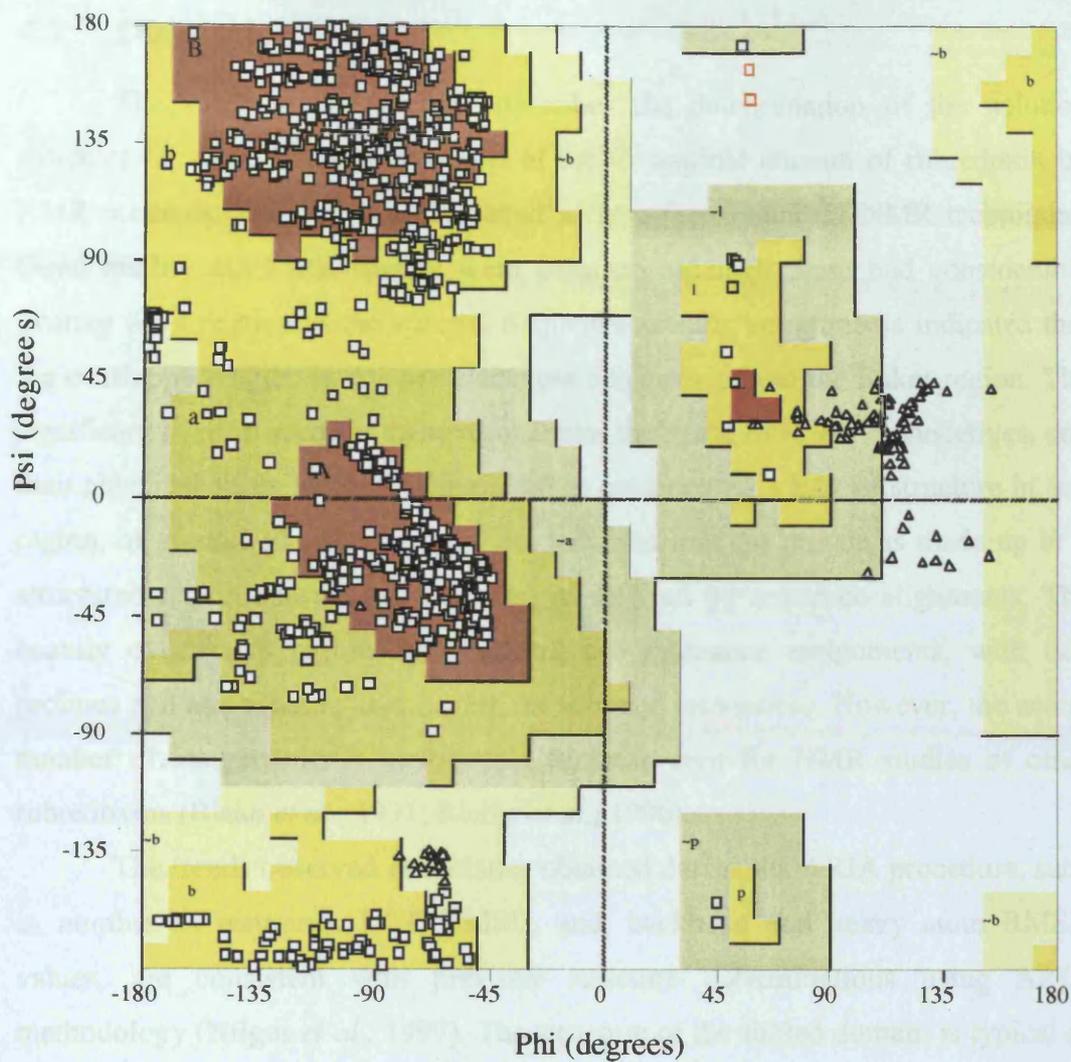
RMSD values between the C-terminal domain and other rubredoxins; the structures used for the RMSD calculations have been solved at high resolution (>1.5 Å) by X-ray crystallography (Dauter *et al.*, 1996) (Adman *et al.*, 1991; Bau *et al.*, 1998; Frey *et al.*, 1987). The comparative RMSD values are all around 1.0 Å, indicating a high level of similarity between the structures (Table 4.4).

Table 4.4. Backbone RMSD values of rubredoxins vs. C-terminal domain.

Organism	PDB Idcode	RMSD vs. Cterm ^a
<i>Clostridium pasteurianum</i>	5rxn	1.04
<i>Desulfovibrio gigas</i>	1rdg	1.21
<i>Desulfovibrio vulgaris</i>	7rxn	1.04
<i>Pyrococcus furiosus</i>	1brf	1.03

^aAll residue backbone RMSD values were calculated using MODELLER-4. The lowest energy structure from the NMR ensemble was used for the C-terminal domain structure.

The structures were analysed using PROCHECK, PROCHECK-NMR and WhatIf (Laskowski *et al.*, 1993; Laskowski *et al.*, 1996; Vriend & Sander, 1993). Analysis with PROCHECK-NMR indicates a reasonable local geometry with no part of the structure giving rise to concern (Laskowski *et al.*, 1996). The Ramachandran plot is reasonable for an NMR structure with only 2 residues in the disallowed region (Figure 4.10). This can be converted into an equivalent resolution of 2.7 Å, which would be expected as NMR structures are generally considered to be low resolution. The average overall G factor for the NMR ensemble is -0.425, and this is a good figure with ideal structures giving values below -0.5 (the G factor was calculated by running PROCHECK for each individual structure of the ensemble and averaging the G factors obtained; Laskowski *et al.*, 1993). Analysis of the packing, which indicates whether the residues feel at home within the protein structure, was performed with WhatIf (Vriend & Sander, 1993). This analysis indicated no individual or stretches of residues with bad packing, and that the overall packing environment is acceptable.



Plot statistics

Residues in most favoured regions [A,B,L]	677	73.6%
Residues in additional allowed regions [a,b,l,p]	209	22.7%
Residues in generously allowed regions [-a,~b,~l,~p]	32	3.5%
Residues in disallowed regions	2	0.2%
<hr/>		
Number of non-glycine and non-proline residues	920	100.0%
Number of end-residues (excl Gly and Pro)	40	
Number of glycine residues (shown as triangles)	100	
Number of proline residues	60	
<hr/>		
Total number of residues	1120	

Figure 4.10. Ramachandran plot of the NMR ensemble for the C-terminal domain. Ramachandran plot produced by PROCHECK-NMR (Laskowski *et al.*, 1996).

4.9 Discussion

The work in this Chapter describes the determination of the solution structure for the Cd-substituted form of the C-terminal domain of rubredoxin by NMR methods. This has been achieved using conventional ^1H NMR techniques. Good quality 2D NMR spectra were obtained although these had considerable overlap for a portion of the spectra. Sequence specific assignments indicated that the overlapped region is due to resonances originating from the linker region. The significant overlap seen for these resonances, their lack of $\text{H}^{\text{N}}\text{-H}^{\text{N}}$ connectives, and their chemical shifts has been interpreted as evidence for a lack of structure in this region, of around 30 amino acids. This indicates that the protein is made up of a structured region and a flexible linker, as inferred by sequence alignments. The heavily overlapped region has hindered the resonance assignments, with two prolines and one aspartic acid having no assigned resonances. However, the actual number of assignments is comparable to those seen for NMR studies of other rubredoxins (Blake *et al.*, 1991; Richie *et al.*, 1996).

The trends observed in statistics obtained during the ARIA procedure, such as number of restraints, NOE RMSD, and, backbone and heavy atom RMSD values, are consistent with previous structure determinations using ARIA methodology (Nilges *et al.*, 1997). The structure of the folded domain is typical of rubredoxins as indicated by the comparative RMSD values. It contains an anti-parallel three stranded β -sheet and a number of tight turns. The number of restraints used to generate the structures, and the RMSD values obtained for the final structures compare well with the NMR structure determination of the rubredoxin of *Pyrococcus furiosus* (Blake *et al.*, 1992b). Structure analysis with PROCHECK-NMR and WhatIf indicate a reasonable local geometry and good packing. Furthermore the G value obtained of -0.425 is excellent for an NMR structure.

The determination of the solution structure for the C-terminal domain represents the first major structural investigation of the rubredoxin of *Pseudomonas oleovorans*. This work strengthens the hypothesis of the linker being a structureless region connecting two rubredoxin type domains, and will allow the construction of mutants based on a reliable structure. The determination of the

Chapter 4: NMR Analysis of the C-Terminal Domain

structure of the C-terminal domain will now facilitate the analysis of the structure of the di-domain rubredoxin using small angle X-ray scattering methods. This work is the focus of Chapter 5.

Chapter 5

SAXS Analysis

5.1 Introduction

This chapter describes small angle X-ray scattering (SAXS) analyses of the 1Fe and 2Fe-forms of the di-domain rubredoxin, and its C-terminal domain. SAXS is a low resolution technique (~ 15 Å) that can provide structural information for proteins in solution, and is especially useful for investigating the arrangement of domains in multidomain proteins. The technique is used to generate an envelope structure for rubredoxin that allows the identification of molecular features. The generation of an envelope structure is achieved on the basis of experimental scattering data alone using the method described by Grossmann & Hasnain (1997).

The SAXS studies represented here provide some degree of structural validation for the NMR solution structure determined for the C-terminal domain rubredoxin (Chapter 4), and comprise the first major structural investigations of the complete di-domain rubredoxin. Of particular importance is whether the two core domains of the di-domain protein interact with each other, and if so, to what degree. The SAXS experiments also provide some information about the linker region, for which to date there is very little structural information.

5.2 SAXS

5.2.1 SAXS theory and applications

Small angle X-ray scattering (SAXS) is a well-established experimental technique that has been available for more than 50 years. Theoretical descriptions were developed by Debye, Fournet and Guinier (Guinier & Fournet, 1955). SAXS benefited strongly from the use of Synchrotron Radiation. Experiments using this intense and collimated radiation began in the 1970s and allowed a variety of studies generally not possible with conventional X-ray equipment. The main advantage of SAXS is that it provides structural information on proteins in solution (solution X-ray scattering). This negates the requirement of crystalline samples needed for X-ray crystallography, and the resonance assignments needed for NMR structure determinations. The disadvantage, compared to a crystal structure or NMR structure, is that it is a low resolution technique (~ 15 Å).

While this study focuses on the determination of the molecular shape deduced from the SAXS profile alone, SAXS can also be used to investigate protein oligomerisation, folding and complex formation (between proteins or proteins and nucleic acids) (Gallagher *et al.*, 1999; Grossmann *et al.*, 1993; Kojima *et al.*, 2000; Krueger *et al.*, 1999). SAXS is particularly useful for investigating conformational flexibility as demonstrated in the study of methane monooxygenase (MMO) and electron-transferring flavoprotein (ETF) (Gallagher *et al.*, 1999; Jones *et al.*, 2000). In both of these studies SAXS provided valuable insight into the complex formation process by demonstrating that dramatic conformational changes occur in these proteins. In particular the MMO study demonstrates the power of this technique as the MMO enzyme has a mass of over 300 kDa. Through manipulations of the crystal structure it has been possible to identify the specific component of the enzyme that controls its reactivity (Gallagher *et al.*, 1999).

Solution X-ray scattering is based on the difference in scattering mass density between solute and solvent i.e. protein and buffer. The electrons of the solute scatter a portion of the incident X-ray beam and the intensity of the scattered X-rays is measured as a function of scattering angle or scattering vector \vec{s} . The magnitude of the scattering vector $|\vec{s}|$ is defined as:

$$|\vec{s}| = s = 2 \sin \theta / \lambda$$

Equation 5.1

where 2θ is the scattering angle, and λ the X-ray wavelength. This results in a scattering profile, which is expressed as a plot of intensity ($I(s)$) versus scattering angle expressed in terms of the reciprocal space quantity (s). Information can be obtained from the scattering data about the size and shape of a protein. At very small angles, in the so-called Guinier region, the curvature of the scattering profile provides information on the overall dimensions of the molecule. From this region, the radius of gyration (R_g) of the molecule can be determined, which is equivalent to the root mean square distance of the atoms from the electronic centre of gravity. At extended or wider angles the profile provides information on the substructure of the molecule (e.g. subunit arrangement of a multidomain protein). The pair distribution function ($p(r)$) obtained from the whole scattering curve by

Fourier transform methods represents the probability of finding a particular distance between two scattering centres (atoms) within the molecule. From this the maximum molecular dimension (D_{max}) can be evaluated. The determination of these parameters is based on the assumption of dilute, monodisperse and non-interacting molecules in solution

The interpretation of X-ray scattering data is greatly advanced from the direct deduction of structural details based upon the assumptions of specific models, such as a structure determined by X-ray crystallography. However, a model-independent approach, based on the multipole expansion method using spherical harmonics, has been developed by Grossman & Hasnain (1997). This allows the determination of the molecular envelope structure of proteins in solution, from high quality scattering data alone. The theory behind this method has been established for more than two decades, but only recently has it been made practically possible due to advances in computer technology (Stuhrmann, 1970). This method has been demonstrated to produce biologically meaningful molecular envelope structures in the absence of any high resolution structure, and it allows the direct comparison between protein structures in solution and solid state (Grossmann & Hasnain, 1997; Grossmann *et al.*, 1999; Jones *et al.*, 2000).

5.2.2 Molecular shape computation

The experimental scattering data is processed using the indirect Fourier transform method, as implemented by the programme GNOM (Semenyuk & Svergun, 1991). This calculates the distance distribution function ($p(r)$), which can be back transformed to give a smooth scattering profile (I_S) extrapolated to zero scattering angle, the forward scattering intensity (I_0), and the radius of gyration (R_g). From I_S , the molecular volume (V) can be obtained via the procedures described in Feigin & Svergun, (1987) which allows one to scale I_S to absolute intensity. The input for the shape restoration programme is the scaled smooth scattering profile.

Under the assumption that the scattering profile results from a globular molecule of homogeneous (i.e. uniform) density within its boundary, its shape can

be defined by the angular envelope function $F(\theta, \varphi)$, such that $F(\theta, \varphi)$ is unity inside the molecular boundary and vanishes elsewhere. This function can be expanded into a series of spherical harmonics $Y_{lm}(\theta, \varphi)$ according to Stuhrmann (1970)(Stuhrmann, 1970):

$$F(\theta, \varphi) = \sum_{l=0}^L \sum_{m=-l}^l f_{lm} Y_{lm}(\theta, \varphi)$$

Equation 5.2

with f_{lm} being complex multipole coefficients, L representing the highest multipole order (i.e. the resolution of the shape representation), and l & m being integers. The scattering curve can be described in terms of multipole coefficients by a computational procedure developed by Svergun & Stuhrmann (1991) (Svergun & Stuhrmann, 1991). Since L determines the resolution and thus the number of free variables for the envelope description $[(L+1)^2 - 6]$ i.e. arbitrary rotations and translations of the molecule do not alter the scattering curve and therefore lead to a reduction of six variables], an upper limit is imposed on L in order to match the information content (number of independent variables) of a scattering profile. Typically, scattering data allows the determination of 15-20 multipole coefficients (Svergun *et al.*, 1996). Consequently, unique shape calculations can be performed up to $L=4$ (i.e. 19 parameters). In the case of molecular symmetry, certain multipole coefficients will disappear, allowing for shape restorations with more details (up to $L=7$ is feasible depending on the degree of symmetry).

The molecular envelope structures produced can be compared with any available structures, by superposition, on a graphics computer (e.g. running Insight II (Biosym/MSI) on a Silicon Graphics workstation).

5.3 SAXS analysis of the C-terminal domain rubredoxin

SAXS analysis was performed on the C-terminal domain rubredoxin, which was expressed and purified as described in Chapter 3. Scattering profiles were collected at protein concentrations of 1 mg/ml and 12.8 mg/ml. The collection of scattering profiles at low and high concentrations allows the operator to distinguish between interparticle effects (due to the close proximity of molecules at high concentration) and aggregation, both of which are characterised by increases in intensity at low scattering angles. While any aggregation will be evident in both high and low concentration samples, any interparticle effects observed are reduced at low concentration. The profile obtained for the high concentration, and the intraparticle distance distribution function, are shown in Figure 5.1. The scattering profile shows that the protein is mono-dispersed, with no protein aggregation or interparticular effects evident (as revealed by comparison with the low concentration data, which was qualitatively similar). Also evident in the scattering profile is a sharp drop in intensity at very low scattering angles. This seen in all scattering profiles and is due to the 'beamstop', which is a physical barrier protecting the detector from the intense incident X-ray beam (those X-rays not scattered by the sample).

Using data from the high concentration sample the molecular shape calculations were carried out up to fourth order harmonics ($L=4$), assuming a non-symmetrical particle. The molecular envelope is shown in Figure 5.2. The theoretical scattering calculated from the restored shape is shown in Figure 5.1. There is an excellent correlation between experimental and theoretical data, indicating that the calculated shape satisfies the experimental finding (the residual value between the smoothed experimental data and the harmonics fit 1.0 %).

The molecular envelope for the C-terminal domain reveals two apparent domains (see Figure 5.2). The larger of the two domains can effectively accommodate the C-terminal core domain (the structure as determined in Chapter 4), displaying features that are in good agreement with the overall structure obtained from NMR. This indicates that it is the structured part of the C-terminal domain that gives rise to the larger domain of the envelope structure.

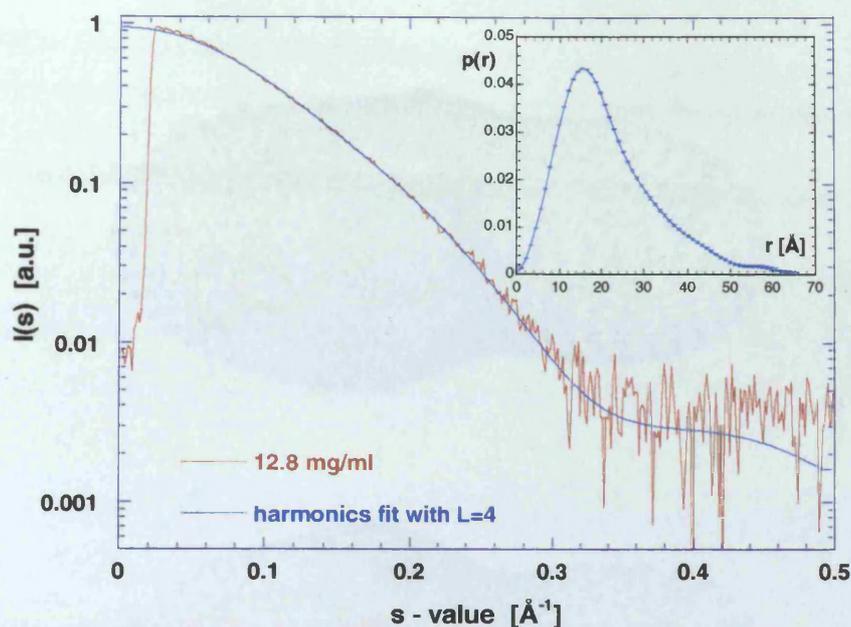


Figure 5.1. SAXS profiles and $p(r)$ function for C-terminal domain rubredoxin. The calculated structural parameters are $R_g = 16.9 \text{ \AA}$, $D_{max} = 64 \text{ \AA}$ and $V = 17500 \text{ \AA}^3$.

The origin of the smaller domain is the flexible linker region. The linker is, in essence, an unstructured ~ 30 amino acid chain linked to a structured domain (as demonstrated by NMR; Chapter 4). Although the linker is considerably longer than the structured domain, its highly flexible nature causes a significant reduction in scattering contrast compared to the homogenous core domain (it "mixes" with the solvent and consequently the full-length linker remains essentially invisible). The flexibility of the linker is reduced in the vicinity of the peptide segment that it is anchored to the core domain. Therefore, the smaller domain represents an area of decreased flexibility of the linker region close to the structured domain, and indicates much of the linker is flexible to a degree that precludes its detection by SAXS analysis.

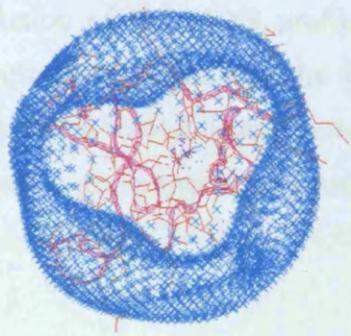
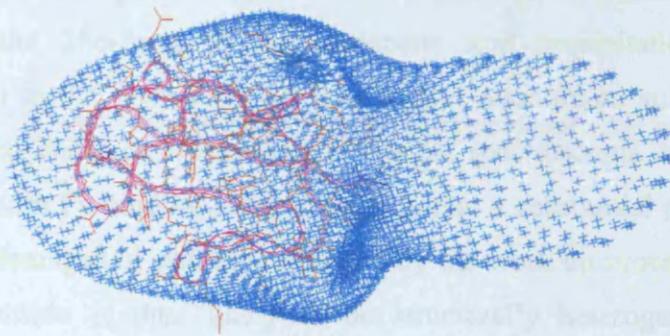
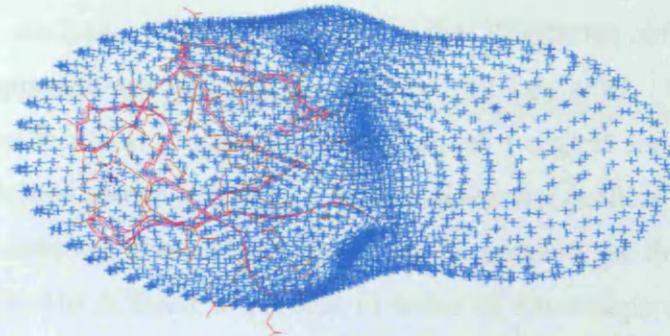


Figure 5.2. Three orientations of the molecular envelope for the C-terminal domain rubredoxin calculated from the SAXS data. The shape of the C-terminal domain rubredoxin calculated at a harmonics resolution of $L=4$. The superimposed backbone ribbons are taken from the NMR structure determined in chapter 3. The NMR structure was superimposed using Insight (Biosym/MSI).

5.4 SAXS analysis of the di-domain rubredoxin

5.4.1 2Fe di-domain rubredoxin

SAXS analysis was performed on the 2Fe-forms of the di-domain rubredoxin, expressed and purified as described by Lee *et al.* (1997). Scattering profiles were collected at protein concentrations of 1 mg/ml and 23 mg/ml, with the profiles obtained shown in Figure 5.3. Both scattering profiles for high and low concentration samples, show strong increases in intensity at the low scattering angles which proved difficult to process in terms of monodisperse particles. This suggests that the sample is heterogeneous (i.e. a polydisperse mixture of particles).

The di-domain protein is purified as the 1Fe-form, which is subsequently converted to the 2Fe-form by trichloroacetic acid precipitation followed by resolubilization in the presence of excess ferrous ammonium sulphate. This is a relatively harsh procedure. Although UV-visible spectroscopic analysis indicates occupancy of both Fe-binding sites, it is likely that a substantial proportion of the protein is left damaged or partially unfolded by the work up procedure (Lee *et al.*, 1997). The sample is thus likely to be structurally heterogeneous, and this suggested by the SAXS profile at low scattering angles. Given the difficulties associated with interpretation of the SAXS profiles for the 2Fe-form of the di-domain rubredoxin, comparable studies with the 1Fe-form were performed, thus negating the harsh precipitation/resolubilisation step in preparation of the sample. This therefore gives rise to inhomogeneities and can also precipitate protein aggregation.

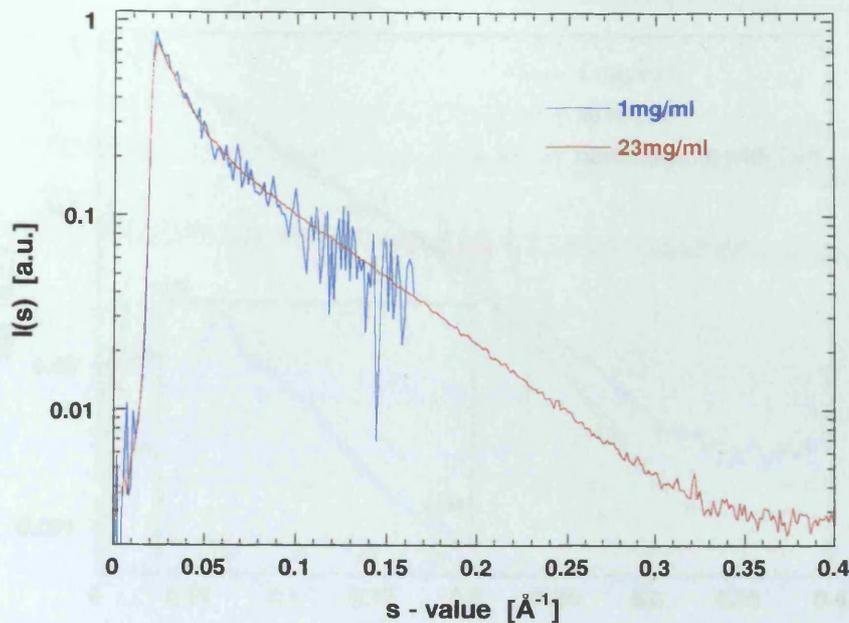


Figure 5.3. SAXS profiles recorded for 2Fe di-domain rubredoxin at low and high concentration.

5.4.2 1Fe di-domain rubredoxin

SAXS analysis was performed on the 1Fe-forms of the di-domain rubredoxin, expressed and purified as described by Lee *et al.* (1997). Scattering profiles were collected at protein concentrations of 1 mg/ml and 30 mg/ml. The composite profile obtained with high and low concentration samples, and the intraparticle distance distribution function, is shown in Figure 5.4. The scattering profile shows that the protein is monodisperse, with no protein aggregation or interparticular effects evident.

A composite of the data collected at high and low was used for the molecular shape calculations. The molecular shape calculations were carried out up to fourth order harmonics ($L=4$), assuming a molecule with no axes of symmetry. The molecular boundary is shown in Figure 5.5. The theoretical scattering calculated from the restored envelope is shown in Figure 5.4. Good agreement between experimental and theoretical data is provided up to approximately $s = 0.2 \text{ \AA}^{-1}$, indicating that at higher scattering angles significant contributions from

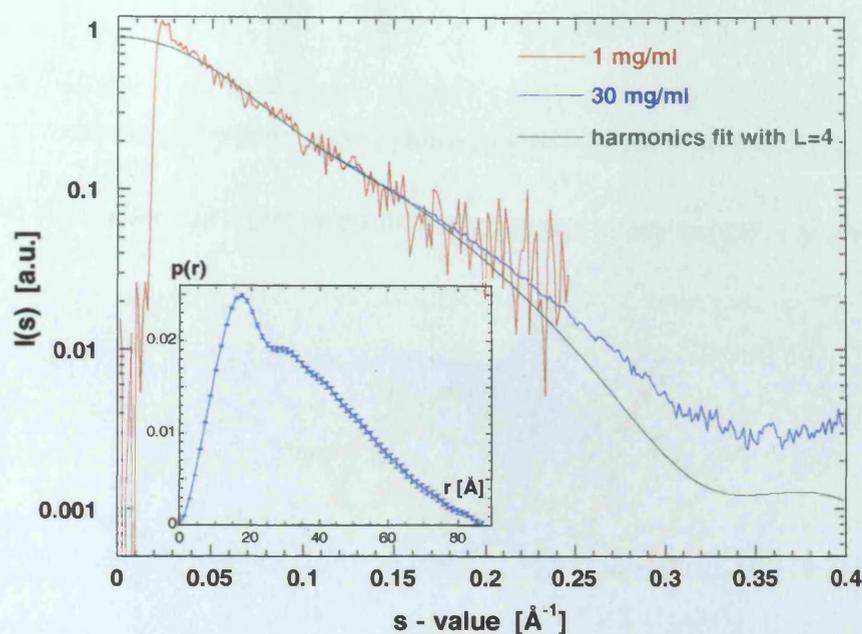


Figure 5.4. SAXS profile and $p(r)$ function for 1Fe di-domain rubredoxin. The calculated structural parameters are $R_g = 26.3 \text{ \AA}$, $D_{max} = 87 \text{ \AA}$ and $V = 34000 \text{ \AA}^3$.

inhomogeneities of this highly flexible, low molecular weight protein are relevant (the residual value between the smoothed experimental data and the harmonics fit is 2.4 %). The theoretical profile at high angles therefore represents the "shape scattering" only (approximating a particle with homogeneous density). Interestingly, these inhomogeneities appear to be less pronounced considering the shape scattering of the C-terminal domain rubredoxin alone (see Fig. 5.1), which highlights the different effects of the linker segments on the scattering contrast. Whereas in the C-terminal domain the linker is more or less invisible (due to only one anchoring point it can merge almost entirely with solvent), in the full-length rubredoxin it is restricted at both ends by the structured core domains causing a more significant disturbance to the homogeneous shape scattering.

The molecular envelope structure for the 1Fe di-domain rubredoxin consists of two large domains (see Figure 5.5) arranged in a "L-shaped" manner. Although each of these envelope domains could accommodate both core domains (from C-terminal and N-terminal; the latter was constructed by a direct amino acid replacement of the C-terminal domain structure, performed in Insight), this would

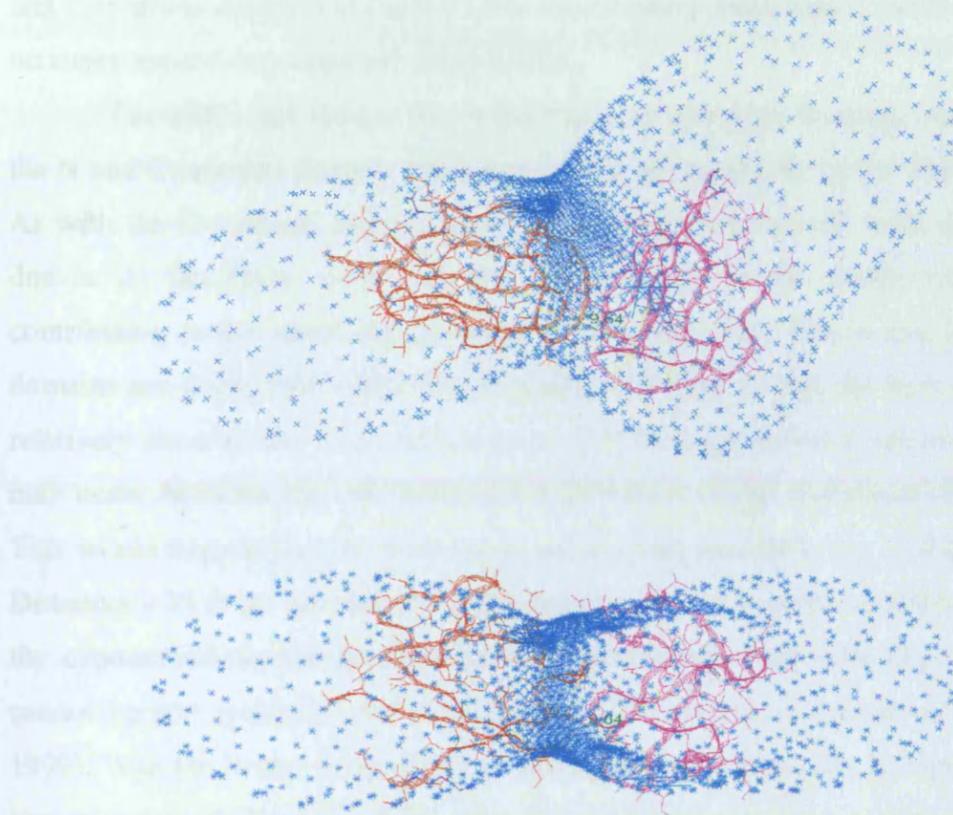


Figure 5.5. Two orientations of the molecular envelope structure for the 1Fe di-domain rubredoxin calculated from the SAXS data. The shape of the 1Fe di-domain rubredoxin calculated at a harmonics resolution of $L=4$. The superimposed backbone ribbons are taken from the NMR structure and a modelled structure (C- and N-terminal respectively). The backbone ribbons were superimposed using Insight (Biosym/MSI).

only be achievable with extensive close contacts between the two domains. The NMR results suggested such an interaction does not occur (see Chapter 4). Therefore, the N and C-terminal domains are fitted into separate envelope domains as shown in Figure 5.5. It should be noted that the position and orientation of the N and C-terminal domains in Figure 5.5 is one of many possibilities, and is not based on experimental data obtained using SAXS.

The additional volume for each of the two envelope domains, compared to the N and C-terminal domain structures, can be accounted for by the linker region. As with the C-terminal domain, parts of the linker has "mixed" with the solvent due to its flexibility, with only the regions close to the structured domains contributing to the molecular boundary (see Section 5.3). The N and C-terminal domains are positioned within the molecular envelope so that the iron atoms are relatively close as the work of Lee *et al.* (1998) suggested that electron transfer may occur between the two irons of the 2Fe-form of the di-domain rubredoxin. This would require that the iron atoms are in close proximity i.e. $\sim 10\text{-}20\text{ \AA}$ apart. Distances $>20\text{ \AA}$ are inconsistent with electron transfer in protein molecules, since the exponential decline in the electron wavefunction over such large distances cannot support good electronic coupling between the two iron atoms (Page *et al.*, 1999). With the N and C-terminal domains positioned as they are in Figure 5.5 the iron atoms would be $\sim 10\text{ \AA}$ apart. Also, the additional volume is around the bottom of the N and C-terminal domains, where the linker originates in both cases. The distance between the carboxyl and amino termini of the N and C-terminal domain respectively, from which the linker originates and terminates, is $\sim 35\text{ \AA}$, this can easily be spanned by the linker region which is over 60 residues in length.

5.5 Discussion

The SAXS analysis of the C-terminal domain rubredoxin indicates that the protein consists of a structured domain with a flexible region. This is in good agreement with the NMR structure for the C-terminal domain (see Chapter 4). Also, demonstrated here is that the linker region only contributes to the envelope model where its flexibility is limited due to its connection with the structured 'core' domain.

The SAXS data for the 2Fe-form of the di-domain rubredoxin indicated that the protein is not monodisperse, possibly due to the harsh conditions of the protein preparation. Although it may be possible in future work to remove any damaged/unfolded protein, the work of Lee *et al.* (1997), suggested that the 2Fe- and 1Fe-form of the di-domain rubredoxin are essentially isostructural. Therefore, a study of the 1Fe-form should provide a good approximation of the solution structure of the 2Fe-form, which is thought to be the physiological form of the protein.

The SAXS analysis of the 1Fe-form of the di-domain protein is consistent with two folded domains interacting with each other. It also demonstrates that there is a region of increased flexibility in this protein, which does not contribute to the envelope model. This is in harmony with the work of Lee *et al.* (1997), which indicates that the di-domain rubredoxin consists of two folded domains and a flexible region (this was determined by a simple comparison of the chemical shifts of the di-domain protein with those of the typical single domain rubredoxins). In their study they also demonstrated that the di-domain protein had intermediate linewidths compared to those seen for 'typical' globular proteins of 19 kDa and the single domain rubredoxins of 6 kDa. This would indicate that while the protein is separated into domains considerably smaller than 19 kDa, it cannot adequately be described as a pair of independent, non-interacting rubredoxin type domains joined by a long amino acid linker. Therefore, the picture of two folded domains, joined by a highly flexible linker, interacting with each other is in agreement with both the SAXS and NMR analysis of the di-domain protein.

The arrangement of the N and C-terminal domains, in the di-domain protein shown in Figure 5.5, is consistent with kinetic analysis of the independent domains (see Chapter 3), and the work of Lee *et al.* (1998). Although the independent domains are both able to transfer electrons from rubredoxin reductase to horse cytochrome *c*, the respective rates are substantially lower than that for the di-domain protein, suggesting an interaction between the two domains. The exact nature of the contribution of this interaction in the electron transfer reaction requires further research. A possible explanation is that both N and C-terminal domains are involved in the interaction between the di-domain rubredoxin and

rubredoxin reductase. Therefore, this interaction is reduced with the independent domains. This would also account for the fact that the independent domains are functional with respect to electron transfer, although at a reduced rate. Additionally, it is thought that electron transfer can take place between the two irons of the 2Fe-form of the di-domain rubredoxin (Lee *et al.*, 1998). The arrangement of the domains, shown in Figure 5.5, leaves the iron atoms approximately 10 Å apart, which is a reasonable distance for supporting relatively fast electron transfer.

Chapter 6

Discussion

6.1 Discussion

The work described in this thesis has provided a structural characterisation of the di-domain rubredoxin of *Pseudomonas oleovorans*. This protein and its redox partner rubredoxin reductase have previously been identified as an ideal system for investigating interprotein ET (Lee *et al.*, 1996). Additionally, the rubredoxin is a component of the alkane hydroxylase system, which has great potential with respect to synthetic chemistry (Smits *et al.*, 1999). A detailed structural analysis of this rubredoxin is essential in characterising the ET reactions and understanding the structural factors influencing the molecular assembly of the alkane hydroxylase system.

As discussed in Chapter 3, the individual N and C-terminal domains of rubredoxin have been isolated. The C-terminal domain was expressed from a mini-gene encoding the C-terminal domain, created from the di-domain rubredoxin expression vector. The vector directed the high-level production of the C-terminal domain, which was purified with a simple procedure developed for the di-domain protein (Lee *et al.*, 1997). The C-terminal domain was isolated as the iron-form, which is consistent with work carried out on the di-domain protein. In steady-state assays, the recombinant C-terminal domain is redox-active and able to transfer electrons from reduced rubredoxin reductase to cytochrome *c*. A limited proteolysis of the C-terminal domain suggested that a major portion of the linker region does not contribute to domain structure, supporting the idea of the di-domain protein comprising two folded domains separated by a flexible linker region. The absorption spectrum and dichroic features of the CD spectrum for the iron and cadmium-substituted C-terminal domain are similar to those reported for the iron and cadmium-substituted *Desulfovibrio gigas* rubredoxin (Henehan *et al.*, 1993). Difference absorption spectroscopy of the cadmium-substituted C-terminal domain revealed the presence of four Gaussian-resolved maxima at 202, 225, 240 and 276 nm; the 240 nm band is attributable from Jørgensen's electronegativity theory to a CysS-Cd(II) charge-transfer excitation. Comparison of 2D NMR spectra collected for the di-domain and the C-terminal domain proteins indicated that the structure of the C-terminal domain is similar when it exists as both an independent domain, and when present in the di-domain rubredoxin. The NMR

data also suggest that there is not a major structural interaction between the two domains, in the di-domain protein.

Attempts to directly express the N-terminal domain from a mini-gene were unsuccessful. However, the N-terminal domain was isolated through cleavage of an engineered 2Fe-rubredoxin in which a factor Xa proteolysis site had been introduced into the putative interdomain linker (a C-terminal His-tag aided the separation of the individual domains). The N-terminal domain was purified as the apoprotein, which again is consistent with work carried out on the di-domain protein. The apoprotein was readily converted to the iron and cadmium forms following precipitation with TCA and resolubilisation in the presence of ferrous ammonium sulphate and cadmium chloride, respectively. The characterisation by absorption spectroscopy is less typical of other rubredoxins; the UV-visible spectrum has a large shoulder and the difference spectrum of the Cd-substituted form cannot be resolved by four Gaussian components. These differences likely reflect the rapid exchange of metal ion between the protein and the bulk solvent, which supports studies of the di-domain protein (Lee *et al.*, 1997). Steady-state assays demonstrate the protein is redox-active. The NMR spectroscopy performed on the N-terminal domain protein was limited compared to that performed on the C-terminal owing to the poor recovery of recombinant N-terminal samples, which were insufficient for a more detailed structural investigation. However, NMR spectroscopy experiments revealed the protein is folded. The N-terminal domain, however, is less stable than the isolated C-terminal domain, a finding that is consistent with the known properties of the full-length 2Fe and Cd-substituted rubredoxin. However, inclusion of more of the N-terminal domain sequence has significantly increased the stability of the protein, compared to the previous attempt to isolate the individual domains via chemical cleavage with cyanogen bromide (Lode & Coon, 1971).

The solution structure of the C-terminal domain was determined using 2D NMR methods, as discussed in Chapter 4. The chemical shifts, where applicable, were consistent with those seen for other rubredoxins, as were the secondary structure elements (Blake *et al.*, 1991; Richie *et al.*, 1996). The resonances of the residues that form the linker region have not been assigned. They are heavily overlapped and lack of H^N - H^N sequential NOE connectivities. This indicates that

the linker region is devoid of structure, which is consistent with the limited proteolysis performed in Chapter 3. The structure was generated using the ARIA method, and statistics from this procedure mirrored previous ARIA structure determinations (Nilges *et al.*, 1997; Pascual *et al.*, 1997). The structure determined for the folded region of the C-terminal domain has a three stranded anti-parallel β -sheet, a number of tight turns, and is typical of the rubredoxin structures previously solved (Adman *et al.*, 1991; Bau *et al.*, 1998; Dauter *et al.*, 1996; Frey *et al.*, 1987).

SAXS analysis was performed on both the C-terminal domain and the di-domain proteins, as described in Chapter 5. This low resolution technique can be used to provide structural information on proteins in solution (Grossmann & Hasnain, 1997). The SAXS analysis of the C-terminal domain rubredoxin indicates that the protein consists of a structured domain with a flexible region. This is in good agreement with the NMR structure for the C-terminal domain (see Chapter 4). Also, SAXS demonstrated that the linker region only contributes to the envelope structure where its flexibility is limited due to its connection with the folded domain.

The SAXS data for the 2Fe-form of the di-domain rubredoxin indicated that the protein is not suitable for SAXS studies, possibly due to the harsh conditions of the protein preparation. However, previous studies have suggested that the 2Fe- and 1Fe-form of the di-domain rubredoxin are essentially isostructural (Lee *et al.*, 1997). Therefore, a study of the 1Fe-form, which should provide a good approximation of the solution structure of the physiological 2Fe-form, was performed. The SAXS analysis of the 1Fe-form of the di-domain protein indicated that it comprises two folded domains that interact with each other, and the intervening flexible region does not contribute to the domain structure. The arrangement of the C- and N-terminal domains of the di-domain protein described in Chapter 5 are consistent with the kinetic analysis of the independent domains (see Chapter 3) and the work of Lee *et al.* (1998). Although the independent domains are both able to transfer electrons from rubredoxin reductase to horse cytochrome *c*, the respective rates are substantially lower than that for the di-domain protein, suggesting an interaction between the two domains.

6.2 Future work

The structure of the N-terminal domain needs to be determined. This will indicate whether the structure is rubredoxin-like, the contribution of the linker to the structured domain, and possibly the origin of the structural instability of this domain, compared to the C-terminal domain. With this achieved it will be feasible to produce a structure for the "whole" rubredoxin by fitting the two domain structures into the SAXS structure for the di-domain protein.

With this accomplished it will then be possible to investigate complex formation between rubredoxin and rubredoxin reductase. Clearly, at some point the structure of rubredoxin reductase will have to be solved for a full characterisation of the ET reaction between the two proteins (this will be achieved through X-ray crystallography for this 55 kDa protein and crystallisation trials are underway). However, in the absence of such a structure a range of experiments can still be performed. Using the amino acid resonance assignments, solvent exchange experiments, similar to those conducted for the CCP:CC system (Section 1.3.2; Jeng *et al.*, 1994; Yi *et al.*, 1994), can be used to identify residues involved in the protein-protein interface. Cross-linking experiments, as used in the studies of all the systems described in Section 1.3 (Kumar & Davidson, 1990; Bisson & Capaldi, 1981; Morand *et al.*, 1989), are also useful in identifying the sites of interactions in protein-protein complexes. The availability of a reliable structure along with information from cross-linking and solvent exchange studies will allow the design of mutants, on the basis of experimental data. Kinetic analysis of such mutants can then be used to further characterise complex formation and the ET reaction.

Whether it will be possible to probe the interaction between the rubredoxin and the alkane hydroxylase with similar experiments remains to be seen. The alkane hydroxylase is a membrane bound protein and therefore is technically difficult to work with. However, it may be possible to use the mutants identified in the reaction with rubredoxin reductase to investigate if the rubredoxin uses the same binding site to interact with both proteins. However, the alkane hydroxylase is not suitable for conventional stopped-flow analysis due to mixing problems associated with membrane bound proteins. Thus, analysis of the ET reactions

involving alkane hydroxylase will need to use techniques such as laser flash photolysis.

Detailed kinetic analysis of the two individual domains may also provide some interesting results. Characterisation of the ET reaction between rubredoxin reductase and rubredoxin indicated that the N-terminal domain may not play an active role in electron transfer, and that the ET event is preceded by a rate-limiting adiabatic event (Lee *et al.*, 1998). The work, in this thesis however, has demonstrated that the N-terminal domain is redox-active and able to transfer electrons from reduced rubredoxin reductase to cytochrome *c*. Therefore, it will be important to investigate whether this adiabatic event is present in the ET reactions of the individual domains. Furthermore it will be important to see if rubredoxin reductase or alkane hydroxylase favour interactions with either domain. These studies may provide insight into the unusual nature of this di-domain rubredoxin.

References

Adman, E. T., Sieker, L. C. & Jensen, L. H. (1991) Structure of rubredoxin from *Desulfovibrio vulgaris* at 1.5 Å resolution. *Journal of Molecular Biology* **217**, 337-352.

Aquino, A. J. A., Beroza, P., Beratan, D. N. & Onuchic, J. N. (1995) Docking and electron transfer between cytochrome c_2 and the photosynthetic reaction center. *Chem. Phys.* **197**, 277-288.

Armstrong, F. A., Driscoll, P. C., Hill, H. A. O. & Redfield, C. (1986) ^1H NMR studies of $\text{Cr}(\text{NH}_3)_6^{3+}$ binding to spinach plastocyanin. *Journal of Inorganic Biochemistry* **28**, 171-180.

Bachmayer, H., Benson, A.M., Yasunobu, K.T., Garrard, W.T. & Whiteley, H.R. (1968a) Non-heme iron proteins. IV. Structural studies of *Micrococcus aerogenes* rubredoxin. *Biochemistry* **7**, 986-996.

Bachmayer, H., Yasunobu, K. T., Peel, J. L. & Mayhew, S. G. (1968b) Non-heme iron proteins. V. The amino acid sequence of rubredoxin from *Peptostreptococcus elsdenii*. *The journal of Biological Chemistry* **243**, 1022-1030.

Baleja, J. D., Thanabal, V. & Wagner, G. (1997) Refined Solution Structure of the DNA-Binding of GAL4 and the use of $3\text{J}(^{113}\text{Cd}, ^1\text{H})$ in Structure Determination. *Journal of Biomolecular NMR* **10**, 397-401.

Banci, L., Bertini, I., Eltis, L. D., Felli, I. C., Kastrau, D. H. W., Luchinat, C., Piccioli, M., Pierattelli, R. & Smith, M. (1994) The three-dimensional structure in solution of the paramagnetic high-potential iron-sulfur protein I from *Ectothiorhodospira halophila* through nuclear magnetic resonance. *European Journal of Biochemistry* **225**, 715-725.

Baptist, J. N., Ghilson, R. K. & Coon, M. J. (1963) Hydrocarbon oxidation by a bacterial enzyme system. I. Products of octane oxidation. *Biochimica et biophysica acta* **69**, 40-47.

Bau, R., Rees, D. C., Kurtz, D. M., Scott, R. A., Huang, H., Adams, M. W. W. & Eidsness, M. K. (1998) Crystal structure of rubredoxin from *Pyrococcus furiosus* at 0.95 Å resolution and the structures of N-terminal methionine and formylmethionine variants of Pf Rd. Contributions of N-terminal interactions to thermostability. *Journal of Biological Inorganic Chemistry* **3**, 484-493.

Benson, A., Tomoda, K., Chang, J., Matseuda, G., Lode, E. T., Coon, M. J. & Yasunobu, K. T. (1971) Evolutionary and phylogenetic relationships of rubredoxin-containing microbes. *Biochemical and Biophysical Research Communications* **42**, 640-646.

Benson, S., Oppici, M., Shapiro, J. & Fennewald, M. (1979) Regulation of membrane peptides by the *Pseudomonas* plasmid *alk* regulon. *Journal of Bacteriology* **140**, 754-762.

Beoku-Betts, D., Chapman, S. K., Knox, C. V. & Sykes, A. G. (1985) Kinetic studies on 1:1 electron-transfer reactions involving blue copper proteins. 11. Effects of pH, competitive inhibition, and chromium(III) modification on the reaction of plastocyanin with cytochrome *f*. *Inorganic Chemistry* **24**, 1677-1681.

Beratan, D. N., Betts, J. N. & Onuchic, J. N. (1991) Protein electron transfer rates set by bridging secondary and tertiary structure. *Science* **252**, 1285-1288.

Beratan, D. N., Onuchic, J. N., Winkler, J. R. & Gray, H. B. (1992) Electron-tunneling pathways in proteins. *Science* **258**, 1740-1741.

Bertini, I., Kurtz, D. M., Eidsness, M. K., Liu, G. H., Luchinat, C., Rosato, A. & Scott, R. A. (1998) Solution structure of reduced *Clostridium pasteurianum* rubredoxin. *Journal of Biological Inorganic Chemistry* **3**, 401-410.

Betts, J. N., Beratan, D. N. & Onuchic, J. N. (1992) Mapping electron tunneling pathways: an algorithm that finds the minimum length/maximum coupling pathway between electron donors and acceptors in proteins. *Journal of the American Chemical Society* **114**, 4043-4046.

Bisson, R. & Capaldi, R. A. (1981) Binding of arylazidocytochrome *c* to yeast cytochrome *c* peroxidase. *Journal of Biological Chemistry* **256**, 4362-4367.

Blake, P. R., Day, M. W., Hsu, B. T., Joshuaator, L., Park, J. B., Hare, D. R., Adams, M. W. W., Rees, D. C. & Summers, M. F. (1992a) Comparison of the X-ray structure of native rubredoxin from *Pyrococcus furiosus* with the NMR structure of the zinc-substituted protein. *Protein Science* **1**, 1522-1525.

Blake, P. R., Lee, B., Summers, M. F., Park, J. B., Zhi, H. Z. & Adams, M. W. W. (1994) Heteronuclear magnetic resonance studies of Zn, ¹¹³Cd, and ¹⁹⁹Hg substituted *P. furiosus* rubredoxin: implications for biological electron-transfer. *New Journal of Chemistry* **18**, 387-395.

Blake, P. R., Park, J. B., Bryant, F. O., Aono, S., Magnuson, J. K., Eccleston, E., Howard, J. B., Summers, M. F. & Adams, M. W. W. (1991) Determinants of protein hyperthermostability: purification and amino acid sequence of rubredoxin from the hyperthermophilic archaeobacterium *Pyrococcus furiosus* and secondary structure of the zinc adduct by NMR. *Biochemistry* **30**, 10885-10895.

Blake, P. R., Park, J. B., Zhou, Z. H., Hare, D. R., Adams, M. W. W. & Summers, M. F. (1992b) Solution-state structure by NMR of zinc-substituted rubredoxin from the marine hyperthermophilic archaeobacterium *Pyrococcus furiosus*. *Protein Science* **1**, 1508-1521.

Brooks, H. B. & Davidson, V. L. (1994a) Free energy dependence of the electron transfer reaction between methylamine dehydrogenase and amicyanin. *Journal of the American Chemical Society* **116**, 11201-11202.

Brooks, H. B. & Davidson, V. L. (1994b) Kinetic and thermodynamic analysis of a physiological intermolecular electron-transfer reaction between methylamine dehydrogenase and amicyanin. *Biochemistry* **33**, 5696-5701.

Brünger, A. (1992) X-PLOR 3.1. A System for x-ray crystallography and NMR, Yale University Press, New Haven.

Cavagnero, S., Zhou, Z. H., Adams, M. W. W. & Chan, S. I. (1995) Response of rubredoxin from *Pyrococcus furiosus* to environmental changes: implications for the origin of hyperthermostability. *Biochemistry* **34**, 9865-9873.

Chakrabarty, A. M., Chou, G. & Gunsalus, I. C. (1973) Genetic regulation of octane dissimilation plasmid in *Pseudomonas*. *Journal of Bacteriology* **118**, 815-820.

Chance, B. & Nishimura, M. (1960) On the mechanism of chlorophyll-cytochrome interaction: the temperature sensitivity of light induced cytochrome oxidation in *Chromatium*. *Proceedings of the National Academy of Sciences of the United States of America* **46**, 10-24.

Chen, J. C. & Mortenson, L. E. (1992) Two open reading frames (Orfs) identified near the hydrogenase structural genes in *Azotobacter vinelandii*, the 1st Orf may encode for a polypeptide similar to rubredoxins. *Biochimica Et Biophysica Acta* **1131**, 122-124.

Chen, L. Y., Durley, R., Poliks, B. J., Hamada, K., Chen, Z. W., Mathews, F. S., Davidson, V. L., Satow, Y., Huizinga, E., Vellieux, F. M. D. & Hol, W. G. J. (1992) Crystal structure of an electron-transfer complex between methylamine dehydrogenase and amicyanin. *Biochemistry* **31**, 4959-4964.

Chen, L. Y., Durley, R. C. E., Mathews, F. S. & Davidson, V. L. (1994) Structure of an electron transfer complex: methylamine dehydrogenase, amicyanin, and cytochrome *c*_{551i}. *Science* **264**, 86-90.

Chen, L. Y., Mathews, F. S., Davidson, V. L., Huizinga, E. G., Vellieux, F. M. D., Duine, J. A. & Hol, W. G. J. (1991) Crystallographic investigations of the tryptophan-derived cofactor in the quinoprotein methylamine dehydrogenase. *Febs Letters* **287**, 163-166.

Condon, E. U. (1926) A theory of intensity distribution in band systems. *Phys. Rev.* **28**, 1182-1201.

Curry, W. B., Grabe, M. D., Kurnikov, I. V., Skourtis, S. S., Beratan, D. N., Regan, J. J., Aquino, A. J. A., Beroza, P. & Onuchic, J. N. (1995) Pathways, pathway tubes, pathway docking, and propagators in electron transfer proteins. *Journal of Bioenergetics and Biomembranes* **27**, 285-293.

Dauter, Z., Sieker, L. C. & Wilson, K. S. (1992) Refinement of rubredoxin from *Desulfovibrio vulgaris* at 1.0 Å with and without restraints. *Acta Crystallographica Section B-Structural Science* **48**, 42-59.

Dauter, Z., Wilson, K. S., Sieker, L. C., Moulis, J. M. & Meyer, J. (1996) Zinc- and iron-rubredoxins from *Clostridium pasteurianum* at atomic resolution: a high-precision model of a ZnS₄ coordination unit in a protein. *Proceedings of the National Academy of Sciences of the United States of America* **93**, 8836-8840.

Davidson, V. L. & Jones, L. H. (1991) Intermolecular electron-transfer from quinoproteins and its relevance to biosensor technology. *Analytica Chimica Acta* **249**, 235-240.

Davidson, V. L. & Jones, L. H. (1995) Complex formation with methylamine dehydrogenase affects the pathway of electron-transfer from amicyanin to cytochrome *c*_{551i}. *Journal of Biological Chemistry* **270**, 23941-23943.

Davidson, V. L. & Jones, L. H. (1996) Electron transfer from copper to heme within the methylamine dehydrogenase-amicyanin-cytochrome *c*-551i complex. *Biochemistry* **35**, 8120-8125.

Davidson, V. L., Jones, L. H. & Zhu, Z. Y. (1998) Site-directed mutagenesis of Phe 97 to Glu in amicyanin alters the electronic coupling for interprotein electron transfer from quinol methylamine dehydrogenase. *Biochemistry* **37**, 7371-7377.

Day, M. W., Hsu, B. T., Joshuator, L., Park, J. B., Zhou, Z. H., Adams, M. W. W. & Rees, D. C. (1992) X-Ray crystal-structures of the oxidized and reduced forms of the rubredoxin from the marine hyperthermophilic archaebacterium *Pyrococcus furiosus*. *Protein Science* **1**, 1494-1507.

Dowe, R. J., Vitello, L. B. & Erman, J. E. (1984) Sedimentation equilibrium studies on the interaction between cytochrome *c* and cytochrome *c* peroxidase. *Archives of Biochemistry and Biophysics* **232**, 566-573.

Drepper, F., Hippler, M., Nitschke, W. & Haehnel, W. (1996) Binding dynamics and electron transfer between plastocyanin and photosystem I. *Biochemistry* **35**, 1282-1295.

Eggink, G., Engel, H., Meijer, W. G., Otten, J., Kingma, J. & Witholt, B. (1988) Alkane utilization in *Pseudomonas oleovorans*: structure and function of the regulatory locus *alkR*. *Journal of Biological Chemistry* **263**, 13400-13405.

Eggink, G., Engel, H., Vriend, G., Terpstra, P. & Witholt, B. (1990) Rubredoxin reductase of *Pseudomonas oleovorans*: structural relationship to other flavoprotein oxidoreductases based on one NAD and 2 FAD fingerprints. *Journal of Molecular Biology* **212**, 135-142.

Eggink, G., Lageveen, R. G., Altenburg, B. & Witholt, B. (1987) Controlled and functional expression of the *Pseudomonas oleovorans* alkane utilizing system in *Pseudomonas putida* and *Escherichia coli*. *Journal of Biological Chemistry* **262**, 17712-17718.

Eidsness, M. K., Richie, K. A., Burden, A. E., Kurtz, D. M. & Scott, R. A. (1997) Dissecting contributions to the thermostability of *Pyrococcus furiosus* rubredoxin: β -sheet chimeras. *Biochemistry* **36**, 10406-10413.

Erman, J. E. & Vitello, L. B. (1980) The binding of cytochrome *c* peroxidase and ferricytochrome *c*. *Journal of Biological Chemistry* **225**, 6224-6227.

Evenson, J. W. & Karplus, M. (1993) Effective coupling in biological electron transfer: exponential of complex dependence. *Science* **262**, 1247-1249.

Feigin, L. A. & Svergun, D. I. (1987) Structure analysis by small-angle X-ray and Neutron scattering. Plenum Publishing Corp, New York.

Fennewald, M., Benson, S., M., O. & Shapiro, J. (1979) Insertion element analysis and mapping of the *Pseudomonas* plasmid *alk* regulon. *Journal of Bacteriology* **139**, 940-952.

Finzel, B. C., Poulos, T. L. & Kraut, J. (1984) Crystal structure of yeast cytochrome *c* peroxidase refined at 1.7-Å resolution. *Journal of Biological Chemistry* **259**, 3027-3036.

Fox, B. G., Shanklin, J., Somerville, C. & Munck, E. (1993) Stearoyl-acyl carrier protein Δ^9 desaturase from *Ricinus communis* is a diiron-oxo protein. *Proceedings of the National Academy of Sciences of the United States of America* **90**, 2486-2490.

Franck, J. (1925) Elementary processes of photochemical reactions. *Trans. Faraday Soc.* **21**, 536-542.

Frey, M., Sieker, L., Payan, F., Haser, R., Bruschi, M., Pepe, G. & Legall, J. (1987) Rubredoxin from *Desulfovibrio gigas*: a molecular-model of the oxidized form at 1.4 Å resolution. *Journal of Molecular Biology* **197**, 525-541.

Gaillard, J., ZhuangJackson, H. Y. & Moulis, J. M. (1996) Nuclear-magnetic-resonance determination of the electron self-exchange rate constant of *Clostridium pasteurianum* rubredoxin. *European Journal of Biochemistry* **238**, 346-349.

Gallagher, S. C., Callaghan, A. J., Zhao, J., Dalton, H. & Trewhella, J. (1999) Global conformational changes control the reactivity of methane monooxygenase. *Biochemistry* **38**, 6752-6760.

Gamow, G. (1928) Zur Quanten Theories des Atom Kernes. *Z. Physics* **51**, 204-212.

Geissdorfer, W., Frosch, S. C., Haspel, G., Ehrt, S. & Hillen, W. (1995) Two genes encoding proteins with similarities to rubredoxin and rubredoxin reductase are required For conversion of dodecane to lauric acid in *Acinetobacter calcoaceticus* ADP1. *Microbiology* **141**, 1425-1432.

Geissdorfer, W., Kok, R. G., Ratajczak, A., Hellingwerf, K. J. & Hillen, W. (1999) The genes *rubA* and *rubB* for alkane degradation in *Acinetobacter* sp. strain ADP1 are in an operon with *estB*, encoding an esterase, and *oxyR*. *Journal of Bacteriology* **181**, 4292-4298.

Geren, L., Hahm, S., Durham, B. & Millett, F. (1991) Photoinduced electron transfer between cytochrome *c* peroxidase and yeast cytochrome *c* labeled at Cys 102 with (4-bromomethyl-4'-methylbipyridine)[bis(bipyridine)]ruthenium²⁺. *Biochemistry* **30**, 9450-9457.

Gholson, R. K., Baptist, J. N. & Coon, M. J. (1963) Hydrocarbon oxidation by a bacterial enzyme system. II. Co-factor requirements for octanol formation from octane. *Biochemistry* **2**, 1155-1159.

Gomes, C. M., Silva, G., Oliveira, S., LeGall, J., Liu, M. Y., Xavier, A. V., RodriguesPousada, C. & Teixeira, M. (1997) Studies on the redox centers of the terminal oxidase from *Desulfovibrio gigas* and evidence for its interaction with rubredoxin. *Journal of Biological Chemistry* **272**, 22502-22508.

Gray, J. C. (1992) Cytochrome *f*: structure, function and biosynthesis. *Photosynthesis Research* **34**, 359-374.

Gray, K. A., Davidson, V. L. & Knaff, D. B. (1988) Complex formation between methylamine dehydrogenase and amicyanin from *Paracoccus denitrificans*. *Journal of Biological Chemistry* **263**, 13987-13990.

Gray, K. A., Knaff, D. B., Husain, M. & Davidson, V. L. (1986) Measurement of the oxidation-reduction potentials of amicyanin and *c*-type cytochromes From *Paracoccus denitrificans*. *Febs Letters* **207**, 239-242.

Grossmann, J. G., Abraham, Z. H. L., Adman, E. T., Neu, M., Eady, R. R., Smith, B. E. & Hasnain, S. S. (1993) X-ray scattering using synchrotron radiation shows nitrite reductase from *Achromobacter xylosoxidans* to be a trimer in solution. *Biochemistry* **32**, 7360-7366.

Grossmann, J. G. & Hasnain, S. S. (1997) X-ray scattering studies of metalloproteins in solution: a quantitative approach for studying molecular conformations. *Journal of Applied Crystallography* **30**, 770-775.

Grossmann, J. G., Hasnain, S. S., Yousafzai, F. K., Smith, B. E., Eady, R. R., Schindelin, H., Kisker, C., Howard, J. B., Tsuruta, H., Muller, J. & Rees, D. C. (1999) Comparing crystallographic and solution structures of nitrogenase complexes. *Acta Crystallographica Section D* **55**, 727-728.

Guinier, A. & Fournet, G. (1955) *Small-Angle Scattering of X-Rays*. Chapman & Hall, Ltd., London.

Gupta, R. K. & Yonetani, T. (1973) Nuclear magnetic resonance study of the interaction of cytochrome *c* with cytochrome *c* peroxidase. *Biochim Biophys Acta* **292**, 502-508.

Gurney, R. W. (1931) The quantum mechanics of electrolysis. *Proc. R. Soc. Lond.* **134**, 137-154.

Guss, J. M., Bartunik, H. D. & Freeman, H. C. (1992) Accuracy and precision in protein structure analysis: restrained least-squares refinement of the structure of poplar plastocyanin at 1.33 Å resolution. *Acta crystallographica. Section B, Structural Science* **48**, 790-811.

Haehnel, W., Jansen, T., Gause, K., Kloesgen, R. B., Stahl, B., Michl, D., Huvermann, B., Karas, M. & Herrmann, R. G. (1994) Electron transfer from plastocyanin to photosystem I. *EMBO journal* **13**, 1028-1038.

Hahm, S., Durham, B. & Millett, F. (1992) Photoinduced electron transfer between cytochrome *c* peroxidase and horse cytochrome *c* labeled at specific lysines with (dicarboxypyridine)(bisbipyridine)ruthenium(II) *Biochemistry* **31**, 3472-3477.

Hazzard, J. T., Moench, S. J., Erman, J. E., Satterlee, J. D. & Tollin, G. (1988) Kinetics of intracomplex electron transfer and of reduction of the components of covalent and noncovalent complexes of cytochrome *c* and cytochrome *c* peroxidase by free flavin semiquinones. *Biochemistry* **27**, 2002-2008.

He, S., Modi, S., Bendall, D. S. & Gray, J. C. (1991) The surface-exposed tyrosine residue Tyr83 of pea plastocyanin is involved in both binding and electron transfer reactions with cytochrome *f*. *EMBO journal* **10**, 4011-4016.

Henehan, C. J., Pountney, D. L., Zerbe, O. & Vasak, M. (1993) Identification of cysteine ligands in metalloproteins using optical and NMR-spectroscopy: cadmium-substituted rubredoxin as a model $[\text{Cd}(\text{CysS})_4]^{2-}$ center. *Protein Science* **2**, 1756-1764.

Husain, M. & Davidson, V. L. (1985) An inducible periplasmic blue copper protein from *Paracoccus denitrificans*. *Journal of Biological Chemistry* **260**, 4626-4629.

Husain, M. & Davidson, V. L. (1986) Characterization of two inducible periplasmic *c*-type cytochromes from *paracoccus denitrificans*. *Journal of Biological Chemistry* **261**, 8577-8580.

Husain, M. & Davidson, V. L. (1987) Purification and properties of methylamine dehydrogenase from *Paracoccus denitrificans*. *Journal of Bacteriology* **169**, 1712-1717.

Jeng, M. F., Englander, S. W., Pardue, K., Rogalskyj, J. S. & McLendon, G. (1994) Structural dynamics in an electron-transfer complex. *Nature Structural Biology* **1**, 234-238.

Jenney, F. E., Verhagen, M., Cui, X. Y. & Adams, M. W. W. (1999) Anaerobic microbes: oxygen detoxification without superoxide dismutase. *Science* **286**, 306-309.

Jørgensen, C. K. (1970) Electron transfer spectra. *Prog.Inorg.Chem.* **12**, 101-157.

Jones, M., Basran, J., Sutcliffe, M. J., Grossmann, J. G. & Scrutton, N. S. (2000) X-ray scattering studies of *Methylophilus methylotrophus* (sp.W₃A₁) electron-transferring flavoprotein. *The Journal of Biological Chemistry* **275**, 21349-21354.

Kang, C. H., Ferguson-Miller, S. & Margoliash, E. (1977) Steady state kinetics and binding of eukaryotic cytochrome *c* with yeast cytochrome *c* peroxidase. *Journal of Biological Chemistry* **252**, 919-926.

Kim, K. L., Kang, D. S., Vitello, L. B. & Erman, J. E. (1990) Cytochrome *c* peroxidase catalyzed oxidation of ferrocycytochrome *c* by hydrogen peroxide: ionic strength dependence of the steady-state rate parameters. *Biochemistry* **29**, 9150-9159.

Kitamura, M., Koshino, Y., Kamikawa, Y., Kohno, K., Kojima, S., Miura, K., Sagara, T., Akutsu, H., Kumagai, I. and Nakaya, T. (1997) Cloning and expression of the rubredoxin gene from *Desulfovibrio vulgaris* (Miyazaki F): comparison of the primary structure of desulfoferrodoxin. *Biochimica et Biophysica Acta-Gene Structure and Expression* **1351**, 239-247.

Kojima, M., Tanokura, M., Maeda, M., Kazumoto, K., Amemiya, Y., Kihara, H. & Takahashi, K. (2000) pH-dependent unfolding of aspergillopepsin II studies by small-angle X-ray scattering. *Biochemistry* **39**, 1364-1372.

Kok, M., Oldenhuis, R., Vanderlinden, M. P. G., Meulenberg, C. H. C., Kingma, J. & Witholt, B. (1989a) The *Pseudomonas oleovorans alkBAC* operon encodes two structurally related rubredoxins and an aldehyde dehydrogenase. *Journal of Biological Chemistry* **264**, 5442-5451.

Kok, M., Oldenhuis, R., Vanderlinden, M. P. G., Raatjes, P., Kingma, J., Vanlelyveld, P. H. & Witholt, B. (1989b) The *Pseudomonas Oleovorans* alkane hydroxylase gene. Sequence and expression. *Journal of Biological Chemistry* **264**, 5435-5441.

Krueger, K. J., McCarry, B. S., Hang, A. H.-J., Shriver, J. W., Trehella, J. & Edmondson, S. P. (1999) The solution structure of the Sac7d/DNA complex: a small angle X-ray scattering study. *Biochemistry* **38**, 10247-10255.

Kumar, M. A. & Davidson, V. L. (1990) Chemical cross-linking study of complex formation between methylamine dehydrogenase and amicyanin from *Paracoccus denitrificans*. *Biochemistry* **29**, 5299-5304.

Kümmerle, R., Zhuang-Jackson, H., Gaillard, J. & Moulis, J. M. (1997) Site-directed mutagenesis of rubredoxin reveals the molecular basis of its electron transfer properties. *Biochemistry* **36**, 15983-15991.

Kusunose, M., Kusunose, E. & Coon, M. J. (1964) Enzymatic ω -oxidation of fatty acids. I. Products of octanoate, decanoate and laurate oxidation. *Journal of Biological Chemistry* **239**, 1374-1380.

Lageveen, R. G. (1986) Oxidation of aliphatic compounds by *Pseudomonas oleovorans*. Biotechnological applications of the alkane hydroxylase system. Ph.D thesis. University of Groningen, the Netherlands.

Langen, R., Chang, I.-J., Germanas, J. P., Richards, J. H., Winkler, J. P. & Gray, H. B. (1995) Electron tunneling in proteins: coupling through a β strand. *Science* **268**, 1733-1735.

Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993) PROCHECK: a program to check the stereochemical quality of protein structures. *Journal of Applied Crystallography* **26**, 283-291.

Laskowski, R. A., Rullmann, J. A. C., MacArthur, M. W., Kaptein, R. & Thornton, J. M. (1996) AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. *Journal of Biomolecular NMR* **8**, 477-486.

Lee, H. J., Basran, J. & Scrutton, N. S. (1998) Electron transfer from flavin to iron in the *Pseudomonas oleovorans* rubredoxin reductase-rubredoxin electron transfer complex. *Biochemistry* **37**, 15513-15522.

Lee, H. J., Lian, L. Y. & Scrutton, N. S. (1996) Rubredoxin/rubredoxin reductase of *Pseudomonas oleovorans*: a model system for investigating interprotein electron transfer. *Biochemical Society Transactions* **24**, S447.

Lee, H. J., Lian, L. Y. & Scrutton, N. S. (1997) Recombinant two-iron rubredoxin of *Pseudomonas oleovorans*: overexpression, purification and characterization by optical, CD and ¹¹³Cd NMR spectroscopies. *Biochemical Journal* **328**, 131-136.

Lee, M. & Chandler, A. C. (1941) A study of the nature, growth and control of bacteria growth in cutting compounds. *Journal of Bacteriology* **41**, 373-386.

Lee, W. Y., Brune, D. C., Lobrutto, R. & Blankenship, R. E. (1995) Isolation, characterization, and primary structure of rubredoxin from the photosynthetic bacterium, *Heliobacillus mobilis*. *Archives of Biochemistry and Biophysics* **318**, 80-88.

Leonard, J. J. & Yonetani, T. (1974) Interaction of cytochrome *c* peroxidase with cytochrome *c*. *Biochemistry* **13**, 1465-1468.

Lode, E. T. & Coon, M. J. (1971) Enzymatic ω -Oxidation. V. Forms of *Pseudomonas oleovorans* rubredoxin containing one or two iron atoms: structure and function in ω -hydroxylation. *The Journal of Biological Chemistry* **246**, 791-802.

Lovenberg, W. & Sobel, B. E. (1965) Rubredoxin: a new electron transfer protein from *Clostridium pasteurianum*. *Proceedings of the National Academy of Sciences of the United States of America* **54**, 193-199.

Lowery, M. D., Guckert, J. A., Gebhard, M. S. & Solomon, E. I. (1993) Active-site electronic structure contributions to electron-transfer pathways in rubredoxin and plastocyanin: direct versus superexchange. *Journal of the American Chemical Society* **115**, 3012-3013.

SPECIAL NOTE

**ITEM SCANNED AS SUPPLIED
PAGINATION IS AS SEEN**

Yoshikawa, S., Shinzawa-Itoh, K., Nakashima, R., Yanon, R., Yamashita, E., Inoue, N., Yao, M., Fei, M. J., Libeu, C. P., Mizushima, T., Yamaguchi, H., Tomizaki, T. & Tsukihara, T. (1998) Redox-coupled crystal structural changes in bovine heart cytochrome *c* oxidase. *Science* **280**, 1723-1729.

Zeng, Q. D., Smith, E. T., Kurtz, D. M. & Scott, R. A. (1996) Protein determinants of metal site reduction potentials: site-directed mutagenesis studies of *Clostridium pasteurianum* rubredoxin. *Inorganica Chimica Acta* **242**, 245-251.

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

Zhang, Z., Huang, L., Shulmeister, V. M., Chi, Y.-I., Kim, K. K., Hung, L.-W., Crofts, A. R., Berry, E. A. & Kim, S.-H. (1998) Electron transfer by domain movement in cytochrome *bc*₁. *Nature* **392**, 677-684.

Zhou, J. S. & Hoffman, B. M. (1993) Cytochrome *c* peroxidase simultaneously binds cytochrome *c* at two different sites with strikingly different reactivities: titrating a "substrate" with an enzyme. *Journal of the American Chemical Society* **115**, 11008-11009.

Zhou, J. S. & Hoffman, B. M. (1994) Stern-Volmer in reverse: 2:1 stoichiometry of the cytochrome *c*-cytochrome *c* peroxidase electron-transfer complex. *Science* **265**, 1693-1696.

Zhou, J. S. & Kostic, N. M. (1992) Photoinduced electron transfer from zinc cytochrome *c* to plastocyanin is gated by surface diffusion within the metalloprotein complex. *Journal of the American Chemical Society* **114**, 3562-3563.

Zhou, J. S., Nocek, J. M., Devan, M. L. & Hoffman, B. M. (1995) Inhibitor-enhanced electron transfer: copper cytochrome *c* as a redox-inert probe of ternary complexes. *Science* **269**, 204-207.

SPECIAL NOTE

**ITEM SCANNED AS SUPPLIED
PAGINATION IS AS SEEN**

Lowry, O. H., Rosebrough, N. J., Farr, A. L. & Randall, R. J. (1951) Protein measurement with the folin phenol reagent. *Journal of Biological Chemistry* **193**.

Marcus, R. A. & Sutin, N. (1985) Electron transfer in chemistry and biology. *Biochimica et biophysica acta* **811**, 265-322.

Martinez, S. E., Huang, D., Szczepaniak, A., Cramer, W. A. & Smith, J. L. (1994) Crystal structure of chloroplast cytochrome *f* reveals a novel cytochrome fold and unexpected heme ligation. *Structure* **2**, 95-105.

Mauk, M. R., Ferrer, J. C. & Mauk, A. G. (1994) Proton linkage in formation of the cytochrome *c*-cytochrome *c* peroxidase complex: electrostatic properties of the high- and low-affinity cytochrome binding sites on the peroxidase. *Biochemistry* **33**, 12609-12614.

Mauro, J. M., Fishel, L. A., Hazzard, J. T., Meyer, T. E., Tollin, G., Cusanovich, M. A. & Kraut, J. (1988) Tryptophan-191→phenylalanine, a proximal-side mutation in yeast cytochrome *c* peroxidase that strongly affects the kinetics of ferrocyanochrome *c* oxidation. *Biochemistry* **27**, 6243-6256.

May, S. W., Lee, L. G., Katopodis, A. G., Kuo, J. Y., Wimalasena, K. & Thowsen, J. R. (1984) Rubredoxin from *Pseudomonas oleovorans*: effects of selective chemical modification and metal substitution. *Biochemistry* **23**, 2187-2192.

McIntire, W. S., Wemmer, D. E., Chistoserdov, A. & Lidstrom, M. E. (1991) A new cofactor in a prokaryotic enzyme: tryptophan tryptophylquinone as the redox prosthetic group in methylamine dehydrogenase. *Science* **252**, 817-824.

Mckenna, E. J. & Coon, M. J. (1970) Enzymatic ω -oxidation. IV. Purification and properties of the ω -hydroxylation of *Pseudomonas oleovorans*. *Journal of Biological Chemistry* **245**, 3882-3889.

Merli, A., Brodersen, D. E., Morini, B., Chen, Z. W., Durley, R. C. E., Mathews, F. S., Davidson, V. L. & Rossi, G. L. (1996) Enzymatic and electron transfer activities in crystalline protein complexes. *Journal of Biological Chemistry* **271**, 9177-9180.

Meyer, J., Gagnon, J., Sieker, L. C., Vandorsselaer, A. & Moulis, J. M. (1990) Rubredoxin From *Clostridium thermosaccharolyticum*: amino acid sequence, mass spectrometric and preliminary crystallographic data. *Biochemical Journal* **271**, 839-841.

Meyer, T. E., Zhao, Z. G., Cusanovich, M. A. & Tollin, G. (1993) Transient kinetics of electron transfer from a variety of *c*-type cytochromes to plastocyanin. *Biochemistry* **32**, 4552-4559.

Millar, J. R., Beitz, J. V. & Huddleston, R. K. (1984) Effect of free energy on rates of electron transfer between molecules. *Journal of the American Chemical Society* **106**, 5057-5068.

Miller, M. A., Liu, R. Q., Hahm, S., Geren, L., Hibdon, S., Kraut, J., Durham, B. & Millett, F. (1994) Interaction domain for the reaction of cytochrome *c* with the radical and the oxyferryl heme in cytochrome *c* peroxidase compound I. *Biochemistry* **33**, 8686-8693.

Misaki, S., Morimoto, Y., Ogata, M., Yagi, T., Higuchi, Y. & Yasuoka, N. (1999) Structure determination of rubredoxin from *Desulfovibrio vulgaris* Miyazaki F in two crystal forms. *Acta Crystallographica Section D-Biological Crystallography* **55**, 408-413.

Mochan, E. & Nicholls, P. (1971) Complex-formation between cytochrome *c* and cytochrome *c* peroxidase. Equilibrium and titration studies. *Biochemical Journal* **121**, 69-82.

Moore, J. M., Case, D. A., Chazin, W. J., Gippert, G. P., Havel, T. F., Powls, R. & Wright, P. E. (1988) Three-dimensional solution structure of plastocyanin from the green alga *Scenedesmus obliquus*. *Science* **240**, 314-317.

Morand, L. Z., Frame, M. K., Colvert, K. K., Johnson, D. A., Krogmann, D. W. & Davis, D. J. (1989) Plastocyanin cytochrome *f* interaction. *Biochemistry* **28**, 8039-8047.

Moser, C. C. & Dutton, P. L. (1996) Outline of theory of protein electron transfer, In: Protein Electron Transfer (Bendall, D. S., Ed.), Bios Scientific Publishers. pp. 1-18.

Moser, C. C., Keske, J. M., Warncke, K., Farid, R. S. & Dutton, P. L. (1992) Nature of biological electron transfer. *Nature* **355**, 796-802.

Nieboer, M., Kingma, J. & Witholt, B. (1993) The alkane oxidation system of *Pseudomonas oleovorans*: induction of the *alk* genes in *Escherichia coli* W3110(pGEC47) affects membrane biogenesis and results in overexpression of alkane hydroxylase in a distinct cytoplasmic membrane subfraction. *Molecular Microbiology* **8**, 1039-1051.

Nilges, M., Macias, M. J., Odonoghue, S. I. & Oschkinat, H. (1997) Automated NOESY interpretation with ambiguous distance restraints: the refined NMR solution structure of the pleckstrin homology domain from β -spectrin. *Journal of Molecular Biology* **269**, 408-422.

Nocek, J. M., Liang, N., Wallin, S. A., Mauk, A. G. & Hoffman, B. M. (1990) Low-temperature conformational transition within the [Zn-Cytochrome *c* peroxidase, cytochrome *c*] electron transfer complex. *Journal of the American Chemical Society* **112**, 1623-1625.

Nordling, M., Sigfridsson, K., Yong, S., Lundberg, L. G. & Hansson, O. (1991) Flash-photolysis studies of the electron transfer from genetically modified spinach plastocyanin to photosystem I. *FEBS letters* **291**, 327-330.

Northrup, S. H., Boles, J. O. & Reynolds, J. C. L. (1988) Brownian dynamics of cytochrome *c* and cytochrome *c* peroxidase association. *Science* **241**, 67-70.

Onuchic, J. N., Beratan, D. N., Winkler, J. R. & Gray, H. B. (1992) Pathway analysis of protein electron transfer reactions. *Annual Review of Biophysics and Biomolecular Structure* **21**, 349-377.

Page, C. C., Moser, C. C., Chen, X. & Dutton, P. L. (1999) Natural engineering principles of electron tunnelling in biological oxidation-reduction. *Nature* **402**, 47-52.

Papavassiliou, P. & Hatchikian, E. C. (1985) Isolation and characterization of a rubredoxin and a two (4Fe-4S) ferredoxin from *Thermodesulfobacterium commune*. *Biochimica Et Biophysica Acta* **810**, 1-11.

Pascual, J., Pfuhl, M., Walther, D., Saraste, M. & Nilges, M. (1997) Solution structure of the spectrin repeat: a left-handed antiparallel triple-helical coiled-coil. *Journal of Molecular Biology* **273**, 740-751.

Pearson, D. C., Gross, E. L. & David, E. S. (1996) Electrostatic properties of cytochrome *f*: implications for docking with plastocyanin. *Biophys.J.* **71**, 64-76.

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

Peterson, J. A., Basu, D. & Coon, M. J. (1966) Enzymatic ω -hydroxylation. I. Electron carriers in fatty acid and hydrocarbon hydroxylation. *Journal of Biological Chemistry* **241**, 5162-5164.

Peterson, J. A. & Coon, M. J. (1968) Enzymatic ω -oxidation. III. Purification and properties of rubredoxin, a component of the ω -hydroxylation system of *Pseudomonas oleovorans*. *The Journal of Biological Chemistry* **243**, 329-34.

Peterson, J. A., Kusunose, M., Kusunose, E. & Coon, M. J. (1967) Enzymatic ω -oxidation. II. Function of rubredoxin as the electron carrier in ω -hydroxylation. *The Journal of Biological Chemistry* **242**, 4334-4340.

Pettigrew, G. W. & Moore, G. R. (1987) Cytochromes *c*: biological aspects. Springer-Verlag, Berlin.

Pladziewicz, J. R. & Brenner, M. S. (1987) Partitioning of ferrocenium ions between multiple redox sites on spinach plastocyanin. *Inorganic chemistry* **26**, 3629-3634.

Poulos, T. L. & Kraut, J. (1980) A hypothetical model of the cytochrome *c* peroxidase•cytochrome *c* electron transfer complex. *Journal of Biological Chemistry* **255**, 10322-10330.

Pountney, D. L., Henehan, C. J. & Vasak, M. (1995) Establishing isostructural metal substitution in metalloproteins using ^1H NMR, circular-dichroism, and Fourier transform infrared spectroscopy. *Protein Science* **4**, 1571-1576.

Qin, L. & Kostic, N. M. (1992) Electron-transfer reactions of cytochrome *f* with flavin semiquinones and with plastocyanin. Importance of protein-protein electrostatic interactions and of donor-acceptor coupling. *Biochemistry* **31**, 5145-5150.

Qin, L. & Kostic, N. M. (1993) Importance of protein rearrangement in the electron-transfer reaction between the physiological partners cytochrome *f* and plastocyanin. *Biochemistry* **32**, 6073-6080.

Qin, L. & Kostic, N. M. (1996) Enforced interaction of one molecule of plastocyanin with two molecules of cytochrome *c* and an electron-transfer reaction involving the hydrophobic patch on the plastocyanin surface. *Biochemistry* **35**, 3379-3386.

Richie, K. A., Teng, Q., Elkin, C. J. & Kurtz, D. M. (1996) 2D ^1H and 3D ^1H - ^{15}N NMR of zinc-rubredoxins: contributions of the β -sheet to thermostability. *Protein Science* **5**, 883-894.

Roberts, V. A., Freeman, H. C., Olson, A. J., Tainer, J. A. & Getzoff, E. D. (1991) Electrostatic orientation of the electron-transfer complex between plastocyanin and cytochrome *c*. (*The*) *Journal of biological chemistry* **266**, 13431-13441.

Ruettinger, R. T., Griffith, G. R. & Coon, M. J. (1977) Characterisation of the ω -hydroxylase of *Pseudomonas oleovorans* as a nonheme iron protein. *Archive biochem biophys* **183**, 528-537.

Ruettinger, R. T., Olson, S. T., Boyer, R. F. & Coon, M. J. (1974) Identification of the ω -hydroxylase of *Pseudomonas oleovorans* as a nonheme iron protein requiring phospholipid for catalytic activity. *Biochemical and Biophysical Research Communications* **57**, 1011-1017.

Saeki, K., Yao, Y., Wakabayashi, S., Shen, G. J., Zeikus, J. G. & Matsubara, H. (1989) Ferredoxin and rubredoxin from *Butyrivibrio methylotrophicum*: complete primary structures and construction of phylogenetic trees. *Journal of Biochemistry* **106**, 656-662.

Salemme, F. R. (1976) A hypothetical structure for the intermolecular electron transfer complex of cytochrome *c* and *b₅*. *Journal of Molecular Biology* **102**, 563-568.

Seki, Y., Seki, S., Satoh, M., Ikeda, A. & Ishimoto, M. (1989) Rubredoxin from *Clostridium perfringens*: complete amino acid sequence and participation in nitrate reduction. *Journal of Biochemistry* **106**, 336-341.

Semenyuk, A. V. & Svergun, D. I. (1991) GNOM a program package for small-angle scattering data processing. *Journal of Applied Crystallography* **24**, 537-540.

Shanklin, J., Whittle, E. and Fox, B. G. (1994) 8 histidine residues are catalytically essential in a membrane associated iron enzyme, stearoyl-CoA desaturase, and are conserved in alkane hydroxylase and xylene monooxygenase. *Biochemistry* **33**, 12787-12794.

Shanklin, J., Achim, C., Schmidt, H., Fox, B. G. & Munck, E. (1997) Mössbauer studies of alkane ω -hydroxylase: evidence for a diiron cluster in an integral-membrane enzyme. *Proceedings of the National Academy of Sciences of the United States of America* **94**, 2981-2986.

Sieker, L. C., Stenkamp, R. E. & Legall, J. (1994) Rubredoxin in crystalline state. *Methods in Enzymology* **243**, 203-216.

Sivaraja, M., Goodin, D. B., Smith, M. & Hoffman, B. M. (1989) Identification byENDOR of Trp¹⁹¹ as the free-radical site in cytochrome *c* peroxidase compound ES. *Science* **245**, 738-740.

Smits, T. H. M., Rothlisberger, M., Witholt, B. & van Beilen, J. B. (1999) Molecular screening for alkane hydroxylase genes in gram-negative and gram-positive strains. *Environmental Microbiology* **1**, 307-317.

Stemp, E. D. A. & Hoffman, B. M. (1993) Cytochrome *c* peroxidase binds two molecules of cytochrome *c*: evidence for a low-affinity, electron-transfer-active site on cytochrome *c* peroxidase. *Biochemistry* **32**, 10848-10865.

Stenkamp, R. E., Sieker, L. C. & Jensen, L. H. (1990) The structure of rubredoxin from *Desulfovibrio desulfuricans* strain-27774 at 1.5 Å resolution. *Proteins-Structure Function and Genetics* **8**, 352-364.

Stuhrmann, H. B. (1970) Interpretation of small-angle scattering functions of dilute solutions and gases. A representation of the structure related to a one-particle-scattering function. *Acta Crystallographica Section A* **26**, 297-306.

Summers, M. F. (1988) ¹¹³Cd NMR spectroscopy of coordination compounds and proteins. *Coordination Chemistry Reviews* **86**, 43-134.

Svergun, D. I. & Stuhrmann, H. B. (1991) New developments in direct shape determination from small-angle scattering. 1. Theory and model calculations. *Acta Crystallographica Section A* **47**, 736-744.

Therien, M. J., Selman, J. M., Gray, H. B., Chang, I. J. & Winkler, J. R. (1990) Long-range electron transfer in ruthenium-modified cytochrome *c*: evaluation of porphyrin ruthenium electronic couplings in the *Candida krusei* and horse heart proteins. *Journal of the American Chemical Society* **112**, 2420-2422.

Ubbink, M., Ejdebäck, M., Karlsson, B. G. & Bendall, D. S. (1998) The structure of the complex of plastocyanin and cytochrome *f*, determined by paramagnetic NMR and restrained rigid-body molecular dynamics. *Structure* **6**, 323-335.

Ueda, T. & Coon, M. J. (1972) Enzymatic ω-oxidation. VII. Reduced diphosphopyridine nucleotide-rubredoxin reductase: properties and function as an electron carrier in ω hydroxylation. *Journal of Biological Chemistry* **247**, 5010-5016.

Ueda, T., Lode, E. T. & Coon, M. J. (1972) Enzymatic ω -oxidation. VI. Isolation of homogeneous reduced diphosphopyridine nucleotide-rubredoxin reductase. *Journal of Biological Chemistry* **247**, 2109-2116.

Ullmann, G. M., Knapp, E.-W. & Kostic, N. M. (1997) Computational simulation and analysis of dynamic association between plastocyanin and cytochrome *f*. Consequences for electron-transfer reaction. *Journal of the American Chemical Society* **119**, 42-52.

Ullmann, G. M. & Kostic, N. M. (1995) Electron-tunneling paths in various electrostatic complex between cytochrome *c* and plastocyanin. Anisotropy of the copper-ligand interactions and dependence of the iron-copper electronic couplings on the metalloprotein orientation. *Journal of the American Chemical Society* **117**, 4766-4774.

van Beilen, J. B. (1994) Alkane oxidation by *Pseudomonas oleovorans*: genes and proteins. Ph.D thesis. University of Groningen, the Netherlands.

van Beilen, J. B., Eggink, G., Enequist, H., Bos, R. & Witholt, B. (1992a) DNA sequence determination and functional characterization of the OCT-plasmid-encoded *alkJKL* genes of *Pseudomonas oleovorans*. *Molecular Microbiology* **6**, 3121-3136.

van Beilen, J. B., Kingma, J. & Witholt, B. (1994) Substrate-specificity of the alkane hydroxylase system of *Pseudomonas oleovorans* Gp01. *Enzyme and Microbial Technology* **16**, 904-911.

van Beilen, J. B., Penninga, D. & Witholt, B. (1992b) Topology of the membrane-bound alkane hydroxylase of *Pseudomonas oleovorans*. *Journal of Biological Chemistry* **267**, 9194-9201.

Vriend, G. & Sander, C. (1993) Quality control of protein models: directional atomic contact analysis. *Journal of Applied Crystallography* **26**, 47-60.

Wang, X. M. & Pielak, G. J. (1999) Equilibrium thermodynamics of a physiologically-relevant heme-protein complex. *Biochemistry* **38**, 16876-16881.

Wastl, J., Sticht, H., Maier, U. G., Rosch, P. & Hoffmann, S. (2000) Identification and characterization of a eukaryotically encoded rubredoxin in a cryptomonad alga. *Febs Letters* **471**, 191-196.

Watenpaugh, K. D., Sieker, L. C. & Jensen, L. H. (1979) The structure of rubredoxin at 1.2 Å resolution. *Journal of Molecular Biology* **131**, 509-522.

Willey, D. L., Auffret, A. D. & Gray, J. C. (1984) Structure and topology of cytochrome *f* in pea chloroplast membranes. *Cell (Cambridge)* **36**, 555-562.

Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*. John Wiley & Sons, New York.

Wuttke, D. S., Bjerrum, M. J., Winkler, J. R. & Gray, H. B. (1992) Electron tunneling pathways in cytochrome *c*. *Science* **256**, 1007-1009.

Yi, Q., Erman, J. E. & Satterlee, J. D. (1994) Studies of protein-protein association between yeast cytochrome *c* peroxidase and yeast iso-1 ferricytochrome *c* by hydrogen-deuterium exchange labeling and proton NMR spectroscopy. *Biochemistry* **33**, 12032-12041.

Yoon, K. S., Hille, R., Hemann, C. & Tabita, F. R. (1999) Rubredoxin from the green sulfur bacterium *Chlorobium tepidum* functions as an electron acceptor for pyruvate ferredoxin oxidoreductase. *Journal of Biological Chemistry* **274**, 29772-29778.

Yoshikawa, S., Shinzawa-Itoh, K., Nakashima, R., Yanon, R., Yamashita, E., Inoue, N., Yao, M., Fei, M. J., Libeu, C. P., Mizushima, T., Yamaguchi, H., Tomizaki, T. & Tsukihara, T. (1998) Redox-coupled crystal structural changes in bovine heart cytochrome *c* oxidase. *Science* **280**, 1723-1729.

Zeng, Q. D., Smith, E. T., Kurtz, D. M. & Scott, R. A. (1996) Protein determinants of metal site reduction potentials: site-directed mutagenesis studies of *Clostridium pasteurianum* rubredoxin. *Inorganica Chimica Acta* **242**, 245-251.

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

Zhang, Z., Huang, L., Shulmeister, V. M., Chi, Y.-I., Kim, K. K., Hung, L.-W., Crofts, A. R., Berry, E. A. & Kim, S.-H. (1998) Electron transfer by domain movement in cytochrome *bc*₁. *Nature* **392**, 677-684.

Zhou, J. S. & Hoffman, B. M. (1993) Cytochrome *c* peroxidase simultaneously binds cytochrome *c* at two different sites with strikingly different reactivities: titrating a "substrate" with an enzyme. *Journal of the American Chemical Society* **115**, 11008-11009.

Zhou, J. S. & Hoffman, B. M. (1994) Stern-Volmer in reverse: 2:1 stoichiometry of the cytochrome *c*-cytochrome *c* peroxidase electron-transfer complex. *Science* **265**, 1693-1696.

Zhou, J. S. & Kostic, N. M. (1992) Photoinduced electron transfer from zinc cytochrome *c* to plastocyanin is gated by surface diffusion within the metalloprotein complex. *Journal of the American Chemical Society* **114**, 3562-3563.

Zhou, J. S., Nocek, J. M., Devan, M. L. & Hoffman, B. M. (1995) Inhibitor-enhanced electron transfer: copper cytochrome *c* as a redox-inert probe of ternary complexes. *Science* **269**, 204-207.