CRYSTALLOGRAPHIC STUDIES OF IRON PROTEINS

Thesis by

Andrew P. Yeh

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosopy

California Institute of Technology

Pasadena, California

2002

(Defended June 3, 2002)

ACKNOWLEDGMENTS

As is the case with all scientific endeavors, the work presented here would not have been possible without the aid and support of many people. I would like to thank my advisor, Prof. Doug Rees, for his advice, support, and guidance in all of these projects. I also thank Profs. Jack Beauchamp, Sunney Chan, and Harry Gray, members of my thesis committee. I am grateful to all of the members of the Rees lab whom I've had the privilege of working with and getting to know through the years. The crystallographic study of SOR was greatly aided by Yonglin Hu, and investigations of ferredoxin were a collaborative effort with Xavier Ambroggio, Susana Andrade, and Oliver Einsle. I am very appreciative to former Rees group members (Clara Kielkopf, John Peters, Caroline Kisker, Hermann Schindelin, and Michael Stowell) for patiently teaching me a whole gamut of crystallographic techniques. Thanks also to Geoffrey Chang and Robert Spencer for their assistance when I was just a rookie first getting started in the lab. And thanks to Hsiu-Ju Chiu, Benedikt Schmid, and Pavel Strop for their help with the use of various programs along the way. Kind folks from various other labs that I would like to thank for their help in the structure determinations include Brian Crane, Hong Li, and Xiao Dong Su. I am also grateful to our many collaborators who have been instrumental to this work: Francis Jenney, Jr. and Michael Adams for providing the SOR; Claire Chatelet and Jacques Meyer for providing the ferredoxin; and Ed Abresch, Herb Axelrod, and George Feher for all of the amazing work that they have done with the photosynthetic reaction center. And last but certainly not least, I would like to thank my family, especially my parents and grandparents, for all of their loving support over the years. This thesis is dedicated to them.

ABSTRACT

The crystal structures of a number of iron proteins from various microbes have been determined in order to better understand the structure-function relationship of these proteins. Several of these iron proteins are as follows:

- Superoxide Reductase (SOR) from the hyperthermophile *Pyrococcus furiosus*. SOR is a non-heme mono-iron protein that functions in anaerobic microbes (e.g., *Pyrococcus furiosus*) as a defense mechanism against reactive oxygen species by catalyzing the reduction of superoxide to hydrogen peroxide. Crystal structures of SOR in both its oxidized and reduced states have been determined and suggest a possible mechanism by which superoxide accessibility may be regulated.

- [2Fe-2S] Ferredoxin 4 (Fd4) from the hyperthermophile *Aquifex aeolicus*. The crystal structure of this [2Fe-2S] ferredoxin has been determined and reveals a thioredoxin-like fold that is novel among iron-sulfur proteins. Protein sequence alignments show that this fold is present as components of more complex anaerobic and aerobic electron transfer systems (e.g., complex I of aerobic respiratory chains). The crystal structures of two variants of this protein in which one of the [2Fe-2S] cysteine ligands was substituted with a serine have also been determined. The structures of these variants provide metric details of unprecedented accuracy for serine-ligated iron-sulfur clusters in proteins.

- *The* Photosynthetic Reaction Center (RC) from the photosynthetic purple bacterium *Rhodobacter sphaeroides*. The primary process of bacterial photosynthesis, which is light-induced trans-membrane charge separation, occurs in the reaction center (RC), an integral membrane protein-pigment complex. We have obtained the crystal structures of the RC bound to the inhibitor stigmatellin in the presence and absence of light to determine any structural change(s) that may be associated with one of its light-induced charge-separated ($D^+Q_A^-$) states. In addition, we have determined the crystal structure of the RC complexed with its physiological electron donor, the soluble monoheme protein cytochrome c_2.

TABLE OF CONTENTS

LIST OF TABLES

Chapter 5

Chapter 6

LIST OF FIGURES

Chapter 1

Chapter 2

Chapter 3

Chapter 4

Chapter 5

Chapter 6

ABBREVIATIONS

C55S	Fd4 Cys55 to Ser55 variant
C59S	Fd4 Cys59 to Ser59 variant
CpFd	[2Fe-2S] ferredoxin from *Clostridium pasteurianum*
Cyt c_2	cytochrome c_2
D	primary electron donor in the RC
DFX	desulfoferrodoxin
DQ_AS	charge-neutral state of RC complexed with stigmatellin
$D^+Q_A^-S$	light-induced charge-separated state of RC complexed with stigmatellin
DQ_AQ_B	charge-neutral state of native RC (PDB entry 1AIJ)
$DQ_AQ_B^-$	light-induced charge-separated state of native RC (PDB entry 1AIG)
ET	electron transfer
F_c	calculated structure factor amplitudes
F_o	observed structure factor amplitudes
Fd4	[2Fe-2S] ferredoxin 4 from *Aquifex aeolicus*
LDAO	lauryl-*N*,*N*-dimethylamineoxide.
PEG 4000	polyethylene glycol 4000
Q_A	primary quinone electron acceptor in the RC
Q_B	secondary quinone electron acceptor in the RC
RC	reaction center
Rd	rubredoxin
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
SOD	superoxide dismutase

SOR	superoxide reductase
SOR-OxLT	superoxide reductase in the oxidized form at low temperature
SOR-OxRT	superoxide reductase in the oxidized form at room temperature
SOR-RedLT	superoxide reductase in the reduced form at low temperature
STG	stigmatellin
Tricine	tris(hydroxymethyl)methylglycine
Tris-HCl	tris(hydroxymethyl)aminomethane hydrochloride
UQ_0	2,3-dimethoxy-5-methylbenzoquinone
UQ_0H_2	2,3-dimethoxy-5-methylbenzoquinol.
WT	wild type

CHAPTER 1

Introduction to Iron Proteins

Introduction

In addition to the twenty naturally occurring amino acids that serve as the building blocks of proteins, the incorporation of other factors into proteins, such as metals, is widespread. It has been estimated that roughly one-third of all proteins purified to homogeneity necessitate metal ions for their proper biological function (*1*). A wide assortment of metals have been found in proteins, including cadmium, calcium, cobalt, copper, iron, magnesium, manganese, molybdenum, tungsten, vanadium, and zinc, among others. The incorporation of such metals into proteins greatly broadens the range of catalytic chemistry accessible to proteins. Catalytic reactions involving dioxygen chemistry or processes involving electron transfer, for example, would be extremely inefficient or perhaps even impossible in many cases without the aid of metals. Among the most extensively used of all metals in biological systems is iron, the prevalence of which has been postulated to be derived from several factors, including its large abundance during the early developmental times of terrestrial life as well as its versatile redox and coordination properties (*2*).

As the predominant redox metal found in biological systems, iron is essential to nearly all organisms, which in turn have developed complex systems to harvest iron from the environment. Aerobic bacteria, for example, secrete so-called siderophores, which are low molecular weight organic compounds that can ligate ferric iron with high affinities. Siderophores complexed with ferric iron are actively transported across the outer and inner membranes of the bacteria via a host of proteins whereupon the iron is released inside the cell. In contrast, in mammals, iron is extracted from food sources.

Once inside cells, iron can subsequently be stored or transported by transferrins, which are proteins that have high affinity for ferric iron. Through the interplay of iron-loaded transferrins and transferrin receptors on cell surfaces, vesicles carrying the iron-loaded transferrins are formed and transport iron to other cells. The lower pH of these vesicles eventually causes the iron to be released from the transferrin. Once obtained and released inside the cell, iron can subsequently be incorporated into a wide array of protein folds to yield iron proteins that will carry out a broad spectrum of different functions. Based primarily on the nature of the ligands to the iron, iron proteins can be broadly divided into three main categories: (1) heme proteins, (2) iron-sulfur proteins, and (3) iron-oxygen/nitrogen proteins, and these are described in further detail below.

Heme Proteins

One common form in which iron is incorporated into proteins is via coordination with a porphyrin cofactor to form iron porphyrins known as hemes (*3, 4*). As illustrated in Figure 1.1, several different types of naturally occurring hemes that are present in proteins include heme *a*, heme *b*, heme *c*, heme *d*, heme d_1, heme *o*, chloroheme, heme P460, and siroheme (*4*). The biosynthesis of such hemes is a multistep process involving multiple enzymes (*5-7*). While their basic structures are similar, these hemes differ in the various substituents attached to their porphyrin ring. Other differences between the various hemes include the way in which they are bound to proteins and the coordination number of the iron. Proteins containing heme *c*, for example, are covalently bound to the heme via two thioether bonds formed between the heme and two cysteine sulfhydryl groups. Similarly, in hydroxylamine oxidoreductase (HAO), heme P460 is covalently

attached to the protein via a carbon-carbon bond to a tyrosine residue, in addition to the thioether bonds to cysteine residues. Such covalent attachment of heme groups to the protein requires the assistance of enzymes (*8*). In contrast, proteins containing hemes other than types *c* or P460 are noncovalently bound to the heme via protein-heme interactions and do not require the aid of other proteins for proper heme incorporation. In addition to being ligated at the equatorial positions to the four nitrogens of the porphyrin ring, the heme iron is ligated to either one or two axial protein ligands resulting in an overall five- or six-coordination sphere, respectively. Residues which have thus far been found to be ligands to the heme iron include histidine, methionine, cysteine, tyrosine, and lysine, with histidine being the most prevalent ligand. Both the nature and number of axial ligands to the heme iron are important factors in determining such properties as the redox potential, spin state, and accessibility of the iron, which in turn ultimately determine the function of the heme protein. Based on these as well as other factors, including the protein environment and solvent accessibility of the heme, heme proteins have been found to possess a wide range of functions, from electron transfer, oxygen transport and storage, nitric oxide transport, O_2 and CO sensing, to catalysis (Table 1.1) (*9*).

Heme proteins that are involved in electron transfer processes are termed cytochromes, of which there are many different types, based on factors such as sequence similarity and the type of heme they possess. Cytochromes can be either soluble proteins, in which case they usually shuttle electrons between membrane protein complexes (e.g.,

those involved in photosynthesis or respiration), or they can be membrane-bound, in which case they are part of multidomain complexes that contain other redox centers.

Cytochrome c_2 [1] from *Rhodobacter sphaeroides*, for example, is a soluble monoheme cytochrome that donates electrons to cofactors in the photosynthetic reaction center during the initial light-driven steps of phosynthesis. Examples of membrane-bound cytochromes include cytochrome bc_1 (ubiquinol:ferricytochrome *c* oxidoreductase) and cytochrome b_6f (plastoquinol:plastocyanin reductase), which are complexes each consisting of three redox subunits. In both complexes, two of the redox subunits are cytochrome *b* and a [2Fe-2S] Rieske protein. The third redox subunit in cytochrome bc_1 and cytochrome b_6f is cytochrome c_1 and cytochrome *f*, repsectively.

Besides functioning as electron transfer proteins, heme proteins have also been found to bind and transport small molecules such as dioxygen or nitric oxide. The group that is involved in dioxygen binding and/or transport are known as globins (*10*) and are found in a wide range of organisms including bacteria, fungi, plants, invertebrates, and vertebrates. Examples include myoglobin and hemoglobin, which store and transport dioxygen, respectively (*11-13*). Instead of binding dioxygen, heme proteins from the insect *Rhodnius prolixus* have been found to bind and transport nitric oxide (NO) (*14, 15*).

---

[1] Cytochromes are usually named after the type of heme (e.g., *a*, *b*, *c*, etc.) that they contain. Subscripts are applied to further differentiate between the cytochromes. Subscripts usually denote either the historical order of discovery of the cytochrome or the maximum electronic absorption peak of the α band of the cytochrome in its reduced form.

These heme proteins, termed nitrophorins, carry NO from the salivary glands of the insect to the host, whereby the NO inhibits platelet aggregation in the host.

A growing number of heme proteins have also been found to act as biological sensors, of which the bacterial proteins FixL and CooA are examples. Via the interaction between an O_2 binding heme domain and a histidine kinase domain, FixL promotes the downregulation of the expression of certain genes in the organism under hypoxic conditions (*16-18*). CooA, on the other hand, acts as a transcription activator by binding to DNA in the presence of CO and thereby promoting the upregulation of gene expression of protein components that are a part of CO oxidation (*19*).

Finally, a number of heme proteins are involved in catalysis. The types of reactions that are catalyzed by heme proteins encompass a broad range, but in all cases, the heme iron coordination is penta-coordinate, allowing for the binding of substrate at the sixth ligand position. The particular reaction catalyzed by a heme protein is in large part due to both the nature of the fifth (proximal) iron ligand, which determines the redox potential of the heme iron, and the nature of the residues surrounding the fifth iron ligand. Among heme enzymes that have been discovered and studied to date are catalases, peroxidases, cytochromes P450, nitric oxide synthases, hydroxylamine oxidoreductases, nitrite reductases, and heme oxygenases, all of which exhibit different functions (Table 1.1).

Iron-Sulfur Proteins

Besides heme, another way in which iron has been incorporated into proteins is in the form of iron-sulfur (Fe-S) clusters. It has been postulated that Fe-S proteins arose during the times of early anaerobic life, when sulfide ions, ferrous iron, and organic thiols were readily available on Earth under its reducing atmosphere (*20-22*). The observation that Fe-S proteins such as ferredoxins are used in primitive organisms whereas NAD/NADP is substituted for the same function(s) in more advanced organisms is suggestive of such proposed early origins of Fe-S proteins (*23*).

As is the case for heme proteins, the variation on the iron-sulfur theme in proteins is vast, due, in part, to the intrinsically high affinity of iron ions for sulfur ligands. Fe-S proteins contain clusters that can range from one to eight irons, with each iron being either in the +2 or +3 oxidation state (*24*). Several of the most common types of Fe-S clusters found in proteins are illustrated in Figure 1.2. The simplest Fe-S center is that of rubredoxins, in which one iron is coordinated by four cysteine thiols (*25*). Next in increasing complexity are [2Fe-2S] clusters, in which two irons are bridged by two sulfide ions to form a rhombohedral framework. In addition, each iron is further coordinated by two cysteine thiols to form ultimately a $[(RS)_4[2Fe\text{-}2S]]^{2-/3-}$ cluster, which is commonly found in many ferredoxins. A slight variation of this cluster type is found in Rieske proteins, in which one of the irons is ligated by the imidazole Nδ's of two histidines rather than by cysteine thiols. Another common type of Fe-S cluster is the [4Fe-4S] cluster in which four irons ions are bridged by four sulfide ions to form a distorted cube termed a "cubane." Each iron is further ligated by a cysteine thiol to form

a resulting $[(RS)_4[4Fe\text{-}4S]]^{-/2-/3-/4-}$. A variation of this cluster is the [3Fe-4S] cluster in which one of the iron ions is missing. In most cases, the coordination of the iron in all of these clusters is approximately tetrahedral. While cluster types such as these are the most common in Fe-S proteins, more complex clusters that are formed with mixed ligands and/or mixed metals or that are coupled to other metal-containing clusters also exist. Examples include the P-cluster and iron-molybdenum cofactor (FeMoco) of nitrogenase (*26, 27*), the H cluster of hydrogenase (*28*), the [4Fe-4S]-siroheme of sulfite reductase (*29, 30*), and the nickel-iron cluster of carbon monoxide dehydrogenase (*31, 32*) (Figure 1.3). While many of these Fe-S clusters can spontaneously assemble under the proper conditions *in vitro*, such is usually not the case *in vivo*, where enzymes are often required to produce elemental sulfur and also to carry sulfur and iron to the site of cluster synthesis (*33-35*). In the nitrogenase system, for example, a number of nitrogen fixation (*nif*) genes are required for the regulation, assembly, and incorporation of metal clusters to form fully mature Fe and MoFe proteins (*36*).

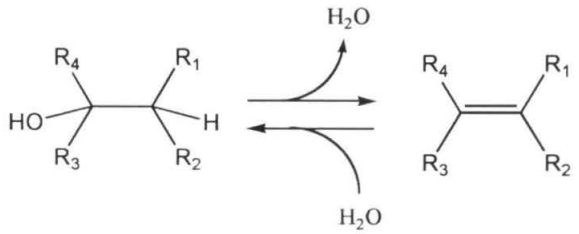
An interesting feature of Fe-S clusters, which attests to their structural versatility and modularity, is their ability to occasionally interconvert between various forms. Interconversion between the [3Fe-4S] and [4Fe-4S] cluster forms by the uptake or loss of one iron, for example, has been demonstrated to occur in ferredoxin II from *Desulfovibrio gigas* as well as in the enzyme aconitase (*37, 38*). The conversion of the [4Fe-4S] cluster to the [3Fe-4S] form in aconitase deactivates the enzyme while re-conversion to the [4Fe-4S] form by the addition of an iron at a substitutionally labile site in the cluster activates it. Another example is the conversion of the [4Fe-4S] cluster to a

[2Fe-2S] form in the FNR protein of *E. coli* upon exposure to dioxygen, thereby inactivating the protein (*39*). In addition to such interconversions, Fe-S clusters have also been found to reorganize via so-called "ligand swapping," where the conformational flexibility of particular protein regions surrounding a cluster allows for the substitution of one or more cluster ligand(s) with another suitable nearby ligand(s) (*40*). Examples of such ligand swapping have been found, again in aconitase (*41*) as well as in the [2Fe-2S] ferredoxin from *Clostridium pasteurianum* (*42*), *Azotobacter vinelandii* (*43*), and *Aquifex aeolicus* (*44*).

Fe-S proteins are quite versatile in terms of their biological functions, as can be seen in Table 1.2. A vast majority of known Fe-S proteins are involved in both intra- and intermolecular electron transfer processes. A feature which makes Fe-S proteins conducive to functioning in electron transfer reactions is the broad range of redox potentials attainable by the cluster (Figure 1.4). While the particular type of cluster plays a large role in determining the final redox potential, much also depends on the protein environment surrounding the cluster (i.e., the cluster ligands, solvent accessibility, and hydrogen bonding network formed between protein moieties and cluster atoms). In proteins of vastly different, unrelated folds that contain the same metal cluster, the range of attainable redox potentials can span hundreds of millivolts (Figure 1.4). In addition to tuning the redox potential to the necessary value for a particular electron transfer process to occur, the protein matrix around the Fe-S cluster also serves to recognize the physiological redox partner and to mediate the intramolecular electron transfer rate from the cluster to the protein surface. As can be seen in Table 1.2, Fe-S proteins function as

electron transfer agents in a broad range of processes such as nitrogen fixation, photosynthesis, and mitochondrial respiration.

While the vast majority of known Fe-S proteins are involved in electron transfer processes, many have been found to serve other functions as well, such as substrate binding and catalysis. For example, different families of Fe-S proteins have been found to function as hydrolyases, catalyzing the following generalized dehydration reaction:



Examples of such hydrolyases from different families include aconitase, methylcitrate dehydratase, isopropylmalate isomerase, fumarase A and B, and maleate hydratase (*45*). In addition to catalyzing dehydration reactions, Fe-S proteins have also been found to serve as primary electron donors in initiating reactions which occur by a free radical mechanism. Such proteins include anaerobic ribonucleotide reductase (*46*), pyruvate formate lyase (*47, 48*), and biotin synthase (*49*). Aside from substrate binding and catalysis, Fe-S proteins can also play a structural role in stabilizing protein structures. The [4Fe-4S] cluster of the endonuclease III family of DNA repair enzymes, for example, stabilizes a particular DNA binding site (*50*). In the absence of the [4Fe-4S] cluster, the endonuclease is inactive. Finally, Fe-S proteins have also been found to perform regulatory and sensing functions. [4Fe-4S] amidotransferase, for instance, senses

dioxygen and protects proteins from attack by intracellular proteases (*51*). [4Fe-4S] aconitase, upon sensing low iron availability, converts to the apo-form and regulates iron usage (*52*). And [2Fe-2S] SoxR, upon sensing superoxide, transcriptionally activates the superoxide detoxification system (*53*). At times, the signal produced by the sensing may require the destruction of the Fe-S cluster, as is the case with amidotransferase and aconitase.

Iron-Oxygen/Nitrogen Proteins

In addition to Fe-S proteins, another group of non-heme proteins are those in which the iron(s) is coordinated predominantly by nitrogen and/or oxygen ligands that come mainly from residues such as histidine, aspartate, glutamate, and tyrosine (these proteins will hereafter be referred to as Fe-O/N proteins). Unlike their Fe-S protein counterparts, which can contain clusters of up to as many as eight irons, Fe-O/N proteins contain only either one or two iron ion(s), usually in the high-spin ferric or ferrous state (*2*). The active form of a particular Fe-O/N protein is dependent on its iron oxidation state. While some Fe-O/N proteins are active only when their iron is in the ferrous state, others are active only when the iron is in its ferric state. As is the case with other types of iron proteins, the redox potential of the iron center is highly dependent on a number of factors, including the ligand type, the polarity of the surrounding environment, local hydrogen bonding network, and the pK_a values of the coordinating iron ligands (*2*). Irons in Fe-O/N proteins generally have a coordination number ranging from four to six, which results in tetrahedral, pyramidal, trigonal bipyramidal, or octahedral coordination geometry (Figure 1.5). However, these coordination geometries are often significantly

distorted from ideal symmetry. In addition, the coordination number of the iron can change quite readily upon a change in oxidation state of the iron and/or binding of substrate. This is in contrast to Fe-S proteins, where each iron of an Fe-S cluster generally maintains tetrahedral coordination geometry, even upon changes in the iron oxidation state. In Fe-O/N proteins, the relative flexibility of the glutamate carboxylate group allows for its ready dissociation from iron, and as a result, carboxylate groups can adopt multiple coordination modes. In diiron proteins such as ribonucleotide reductase, for example, a glutamate carboxylate ligand can convert between mono- and bidentate modes of iron binding, and in so doing, control substrate accessibility to the diiron site (*54, 55*). In contrast, due to their bulkiness and relative structural rigidity, histidine ligands generally maintain their coordination to the iron, providing a rigid coordination framework.

While the majority of known Fe-S proteins are involved in electron transfer processes, Fe-O/N proteins function mainly in a catalytic capacity, particularly in the promotion of reactions involving dioxygen chemistry. Fe-O/N proteins are well suited for such processes as they do not self-inactivate in presence of dioxygen, unlike many Fe-S proteins (*2*). The observation that Fe-O/N proteins catalyze dioxygen-related reactions, in contrast to the fact that Fe-S proteins often suffer deleterious effects in the presence of dioxygen, suggests that Fe-O/N proteins may have been evolutionarily subsequent to Fe-S proteins. Fe-O/N proteins can be classified into four general groups, based primarily on their function and number of irons: (1) mononuclear ferric Fe-O/N proteins that activate substrate, (2) mononuclear Fe-O/N proteins that activate O_2, (3) dinuclear Fe-

O/N proteins that activate/bind O_2, and (4) Fe-O/N proteins that act as hydrolases (*2*). Several examples of proteins from each group and their particular functions are listed in Table 1.3. As can be seen in this table, reactions that are catalyzed by Fe-O/N proteins utilizing dioxygen (groups 1-3) span a broad spectrum and include desaturation, oxidative cyclization, mono- and di-oxygenations, epoxidation, and electron transfer reactions, among others (*2*). While the utilization of dioxygen in the reaction scheme is a common denominator amongst proteins from the first three groups, there are significant differences in the way in which this oxidative power of dioxygen is harnessed. The most common way is for Fe(II) of the Fe-O/N protein to react with and activate dioxygen, thereby forming an iron-oxo species that can subsequently react with the organic substrate (*2*). Such is the method employed by groups 2 and 3. The method used by Fe-O/N proteins from group 1, on the other hand, differs in that Fe(III), rather than Fe(II), is the active species. Instead of activating dioxygen, Fe(III) activates the organic substrate, which subsequently is able to undergo dioxygen insertion.

Research Objectives

The iron proteins that are the subject of this thesis are representative of all three main groups described above. Crystallographic investigations of these iron proteins were carried out in the hope of better understanding their function and mechanism in light of their three-dimensional structures. In Chapter 2, the structure determination of superoxide reductase (SOR) from the hyperthermophilic marine bacterium *Pyrococcus furiosus* is described. SOR can be considered as a Fe-O/N protein, as its iron has been found to coordinate four histidines, a cysteine, and a variable glutamate. Like many other

Fe-O/N proteins, SOR is involved in dioxygen chemistry, as it utilizes ferrous iron to reduce superoxide anion to hydrogen peroxide. The structures of SOR in both its oxidized and reduced states are described and reveal a possible gating mechanism by which superoxide substrate accessibility is regulated.

The subject of Chapter 3 is the structure determination of an iron-sulfur protein, namely, the [2Fe-2S] ferredoxin 4 (Fd4) from the hyperthermophilic marine bacterium *Aquifex aeolicus*. The structure of this ferredoxin reveals a thioredoxin-like protein fold that has not been previously observed for other [2Fe-2S] ferredoxins of known structure. As an extension of the work described in Chapter 3, the crystal structure of wild type Fd4 was determined at a higher resolution, providing a much more detailed description of the [2Fe-2S] cluster site. In addition to the higher resolution wild type Fd4 structure, the structures of two single substitution variants of Fd4 in which a serine was substituted for one of the cysteine ligands to the [2Fe-2S] cluster were also investigated. These structures provide geometries of serine-ligated Fe-S active sites with unprecedented accuracy and reveal ways in which the [2Fe-2S] cluster and the surrounding protein environment accommodate the cysteine to serine substitutions. The results of these investigations are reported in Chapter 4.

Chapters 5 and 6 concern crystallographic studies on the photosynthetic reaction center (RC) from the purple photosynthetic bacterium *Rhodobacter sphaeroides*. While its precise function is still unclear, the iron in the RC, with its three histidine and one glutamate ligands, is yet another example of iron coordinated predominantly by oxygen

and/or nitrogen ligands. The aim of our research on the RC was twofold. One was to determine whether or not any structural changes occur upon formation of the light-induced $D^+Q_A^-$ charge-separated state of the RC. The presence or absence of any structural changes associated with this charge-separated state may yield a better understanding of the charge-transfer mechanism catalyzed by the RC. The results of this investigation are reported in Chapter 5.

The second aim of our research on the RC concerns a heme protein in addition to the RC. As mentioned previously, in *Rhodobacter sphaeroides*, the electron donor to the RC during light-driven trans-membrane charge separation is the soluble monoheme protein cytochrome c_2. Unlike photosynthetic RC complexes in other bacteria (e.g., *Bcl. viridis* and *T. tepidum*), in which a soluble tetraheme cytochrome is permanently bound to the membrane-spanning portion of the RC complex, in *Rb. sphaeroides*, cytochrome c_2 readily dissociates from the membrane-spanning RC. While the structures of each of these proteins alone had already previously been determined, that of the complex remained unsolved, leaving questions regarding the location and mode of binding of cytochrome c_2 with respect to the RC unanswered. Such questions were addressed by determining the crystal structure of the RC-cytochrome c_2 complex, which is presented in Chapter 6.

References

1. Holm, R. H., Kennepohl, P., and Solomon, E. I. (1996) *Chem. Rev. 96*, 2239-2314.
2. Nordlund, P. (2001) in *Handbook on Metalloproteins* (Bertini, I., Sigel, A., and Sigel, H., Eds.) pp 461-570, Marcel Dekker, Inc., New York.
3. Lever, A. B. P., and Gray, H. B. (Eds.) (1983) *Iron Porphyrins*, Addison-Wesley, Reading, MA.
4. Kadish, K. M., Smith, K. M., and Guilard, R. (Eds.) (2000) *The Porphyrin Handbook*, Academic Press, San Diego, CA.
5. Jordan, P. M. (1991) *Biosynthesis of Tetrapyrroles*, Elsevier, Amsterdam.
6. Dailey, H. A. (1997) *J. Biol. Inorg. Chem. 2*, 411-417.
7. Dailey, H. A. (1990) in *Biosynthesis of Heme and Chlorophylls* (Dailey, H. A., Ed.) pp 123-161, McGraw-Hill, New York.
8. Thoeny-Meyer, L. (1997) *Microbiol. Rev. 61*, 337-376.
9. Turano, P., and Lu, Y. (2001) in *Handbook on Metalloproteins* (Bertini, I., Sigel, A., and Sigel, H., Eds.) pp 269-356, Marcel Dekker, Inc., New York.
10. Jameson, G. B., and Ibers, J. A. (1994) in *Bioinorganic Chemistry* (Bertini, I., Lippard, S. J., Gray, H. B., and Valentine, J. S., Eds.), University Science Books, Mill Valley, CA.
11. Suzuki, T., and Imai, K. (1998) *Cellular and Molecular Life Sciences 54*, 979-1004.
12. Dickerson, R. E., and Geis, I. (1983) *Hemoglobin: Structure, Function, Evolution, and Pathology*, Benjamin/Cummings, Menlo Park, CA.
13. Hardison, R. C. (1996) *Proc. Natl. Acad. Sci. U. S. A. 93*, 5675-5679.

14. Champagne, D. E., Nussenzveig, R. H., and Ribeiro, J. M. C. (1995) *J. Biol. Chem. 270*, 8691-8695.

15. Valenzuela, J. G., and Ribeiro, J. M. C. (1998) *J. Exp. Biol. 201*, 2659-2664.

16. Pellequer, J. L., Brudler, R., and Getzoff, E. D. (1999) *Curr. Biol. 9*, R416-R418.

17. Rodgers, K. R. (1999) *Curr. Opin. Chem. Biol. 3*, 158-167.

18. Chan, M. K. (2001) *Curr. Opin. Chem. Biol. 5*, 216-222.

19. Aono, S., and Nakajima, H. (1999) *Coord. Chem. Rev. 192*, 267-282.

20. Grabowski, R., Hofmeister, A. E. M., and Buckel, W. (1993) *Trends Biochem. Sci. 18*, 297-300.

21. Wächtershäuser, G. (1992) *Prog. Biophys. Mol. Biol. 58*, 85-201.

22. Schauder, R., Widdel, F., and Fuchs, G. (1987) *Arch. Microbiol. 148*, 218-225.

23. Daniel, R. M., and Danson, M. J. (1995) *J. Mol. Evol. 40*, 559-563.

24. Beinert, H., Holm, R. H., and Münck, E. (1997) *Science 277*, 653-659.

25. Day, M. W., Hsu, B. T., Joshuator, L., Park, J. B., Zhou, Z. H., Adams, M. W. W., and Rees, D. C. (1992) *Protein Sci. 1*, 1494-1507.

26. Chan, M. K., Kim, J. S., and Rees, D. C. (1993) *Science 260*, 792-794.

27. Peters, J. W., Stowell, M. H. B., Soltis, S. M., Finnegan, M. G., Johnson, M. K., and Rees, D. C. (1997) *Biochemistry 36*, 1181-1187.

28. Adams, M. W. W. (1990) *Biochimica Et Biophysica Acta 1020*, 115-145.

29. Christner, J. A., Münck, E., Kent, T. A., Janick, P. A., Salerno, J. C., and Siegel, L. M. (1984) *J. Am. Chem. Soc. 106*, 6786-6794.

30. Crane, B. R., Siegel, L. M., and Getzoff, E. D. (1995) *Science 270*, 59-67.

31. Dobbek, H., Svetlitchnyi, V., Gremer, L., Huber, R., and Meyer, O. (2001) *Science 293*, 1281-1285.

32. Drennan, C. L., Heo, J. Y., Sintchak, M. D., Schreiter, E., and Ludden, P. W. (2001) *Proc. Natl. Acad. Sci. U. S. A. 98*, 11973-11978.

33. White, R. H. (1983) *Biochem. Biophys. Res. Commun. 112*, 66-72.

34. Takahashi, Y., Mitsui, A., Hase, T., and Matsubara, H. (1986) *Proc. Natl. Acad. Sci. U. S. A. 83*, 2434-2437.

35. Flint, D. H. (1996) *J. Biol. Chem. 271*, 16068-16074.

36. Dean, D. R., Bolin, J. T., and Zheng, L. M. (1993) *J. Bacteriol. 175*, 6737-6744.

37. Moura, J. J. G., Moura, I., Kent, T. A., Lipscomb, J. D., Huynh, B. H., Legall, J., Xavier, A. V., and Münck, E. (1982) *J. Biol. Chem. 257*, 6259-6267.

38. Kent, T. A., Dreyer, J. L., Kennedy, M. C., Huynh, B. H., Emptage, M. H., Beinert, H., and Münck, E. (1982) *Proc. Natl. Acad. Sci. U. S. A. 79*, 1096-1100.

39. Khoroshilova, N., Popescu, C., Münck, E., Beinert, H., and Kiley, P. J. (1997) *Proc. Natl. Acad. Sci. U. S. A. 94*, 6087-6092.

40. Golinelli, M. P., Akin, L. A., Crouse, B. R., Johnson, M. K., and Meyer, J. (1996) *Biochemistry 35*, 8995-9002.

41. Kennedy, M. C., Kent, T. A., Emptage, M., Merkle, H., Beinert, H., and Münck, E. (1984) *J. Biol. Chem. 259*, 4463-4471.

42. Golinelli, M. P., Chatelet, C., Duin, E. C., Johnson, M. K., and Meyer, J. (1998) *Biochemistry 37*, 10429-10437.

43. Chatelet, C., and Meyer, J. (1999) *J. Biol. Inorg. Chem. 4*, 311-317.

44. Chatelet, C., Gaillard, J., Petillot, Y., Louwagie, M., and Meyer, J. (1999) *Biochem. Biophys. Res. Commun. 261*, 885-889.

45. Flint, D. H., and Allen, R. M. (1996) *Chem. Rev. 96*, 2315-2334.

46. Ollagnier, S., Meier, C., Mulliez, E., Gaillard, J., Schuenemann, V., Trautwein, A., Mattioli, T., Lutz, M., and Fontecave, M. (1999) *J. Am. Chem. Soc. 121*, 6344-6350.

47. Broderick, J. B., Duderstadt, R. E., Fernandez, D. C., Wojtuszewski, K., Henshaw, T. F., and Johnson, M. K. (1997) *J. Am. Chem. Soc. 119*, 7396-7397.

48. Kulzer, R., Pils, T., Kappl, R., Huttermann, J., and Knappe, J. (1998) *J. Biol. Chem. 273*, 4897-4903.

49. Duin, E. C., Lafferty, M. E., Crouse, B. R., Allen, R. M., Sanyal, I., Flint, D. H., and Johnson, M. K. (1997) *Biochemistry 36*, 11811-11820.

50. Thayer, M. M., Ahern, H., Xing, D. X., Cunningham, R. P., and Tainer, J. A. (1995) *Embo J. 14*, 4108-4120.

51. Grandoni, J. A., Switzer, R. L., Makaroff, C. A., and Zalkin, H. (1989) *J. Biol. Chem. 264*, 6058-6064.

52. Gütlich, P., Link, R., and Trautwein, A. (1978) *Mössbauer Spectroscopy and Transition Metal Chemistry*, Spinger Verlag, Berlin.

53. Haile, D. J., Rouault, T. A., Harford, J. B., Kennedy, M. C., Blondin, G. A., Beinert, H., and Klausner, R. D. (1992) *Proc. Natl. Acad. Sci. U. S. A. 89*, 11735-11739.

54. Nordlund, P., Sjoberg, B. M., and Eklund, H. (1990) *Nature 345*, 593-598.

55. Nordlund, P., and Eklund, H. (1993) *J. Mol. Biol. 232*, 123-164.

56. Bentrop, D., Capozzi, F., and Luchinat, C. (2001) in *Handbook on Metalloproteins* (Bertini, I., Sigel, A., and Sigel, H., Eds.) pp 461-570, Marcel Dekker, Inc., New York.

57. Capozzi, F., Ciurli, S., and Luchinat, C. (1998) *Metal Sites in Proteins and Models 90*, 127-160.

Table 1.1. Classification of heme proteins, along with biological function, heme type, axial ligation of heme, and oxidation and spin states of the heme iron (Adapted from (*9*)).

Protein class/family	Function	Heme type	Axial ligation	Formal iron oxidation/spin states
Cytochromes	Electron transfer			
Cytochromes *c*:		*c*		
Class I			His Nε/Met Sδ	
Class II (beside *c'*)			His Nε/Met Sδ	
Class II cytochrome *c'*			His Nε	
Class III			His Nε/Met Sδ	Fe(II) (S=0)
Class IV			His Nε/Met Sδ + His Nε/ His Nε	Fe(III) (S=1/2) with the exception of Cyt *c'*, which has Fe(II) (S=2) Fe(III) (S=5/2)
Cytochromes *b*:		*b*		
Cytochrome b_5, b_2, b/b_6			His Nε/His Nε	
Cytochrome b_{562}			His Nε/Met Sδ	
Cytochrome *f*		*c*	His Nε/amino group of Tyr	
Globins:				
Myoglobins	O_2 storage	*b*		
Hemoglobins	O_2 transport	*b* or chloro-heme	His Nε	Fe(II) (S=2) deoxy Fe(II) (S=0) oxy
Flavohemoproteins	O_2 sensing; NO deoxygentaion	*b*		
Nitrophorins	Nitric oxide transport	*b*	His Nε	Fe(II) (S=2) Fe(III) (S=5/2)
Peroxidases	$XH_2 + H_2O_2 \rightarrow X + 2H_2O$	*b*	His Nε (+ H_2O)	Fe(III) (S=5/2)
Chloroperoxidases	$XH + Y^- + H^+ H_2O2 \rightarrow XY + H_2O$	*b*	Cys Sγ	Fe(III) (S=5/2)
Catalases	$2H_2O_2 \rightarrow 2H_2O + O_2$	*b*	Tyr Oη	Fe(III) (S=5/2)
P450 proteins	$RH + O_2 + 2H^+ + 2e^- \rightarrow ROH + H_2O$	*b*	Cys Sγ (+ H_2O or OH^-)	Fe(III) (S=5/2)

Table 1.1. continued

Heme oxygenase	Heme biodegration: Oxidation of heme to biliverdin, CO, and free iron	*b*	His Nε	Fe(III) (S=5/2)
CooA	CO-sensing transcription activator	*b*	His Nε (or Cys)/Pro	Fe(II) (S=0)
FixL	O_2 sensor	*b*	His Nε	Fe(II) (S=2) deoxy
Nitric oxide synthases	L-Arg + $2O_2$ + 3/2 NADPH → citrulline + NO + 3/2 $NADP^+$	*b*	Cys Sγ	Fe(III) (S=5/2)
Cytochrome *c* oxidase	O_2 + 4Fe(II)cyt *c* + $4e^-$ + $4H^+$ → $2H_2O$ + 4Fe(III)cyt *c*	a_3	His Nε	Fe(III) (S=5/2) Sping-coupled to Cu_B
Cytochrome *bo* quinol oxidase	O_2 + 4Fe(II)cyt *c* + $4e^-$ + $4H^+$ → $2H_2O$ + 4Fe(III)cyt *c*	*o*	His Nε	Fe(III) (S=5/2)
Cyt c_{554} Heme 1 Heme 3, heme 4 Heme 2	Oxidation of ammonia to nitrite	*c*	 His Nε/His Nδ His Nε/ Met Sδ His Nε	 Fe(II) (S=0) Fe(III) (S=1/2) Fe(III) (S=5/2)
Hydroxylamine oxidoreductase	$NH_2OH + H_2O$ → NO_2^- + $4e^-$ + $5H^+$	P460	His Nε	Fe(III) (S=5/2 or S=3/2)
Cytochrome cd_1 nitrite reductase	NO_2^- + $2H^+$ + e^- → NO + $2H_2O$ & O_2 + $4H^+$ + $4e^-$ → $2H_2O$	d_1	His Nε or His Nε + Tyr Oη	Fe(III) (S=1/2)
Cytochrome *c* nitrite reductase	NO_2^- + $2H^+$ + e^- → NO + $2H_2O$	*c*	Lys Nζ	Fe(III) (S=5/2)
Nitrite reductases	NO_2^- + $8H^+$ + $6e^-$ → NH_4^+ + $2H_2O$	Siro-heme	Cys Sγ	Fe(II) (S=1 or S=2)
Sulfite reductases	HSO_3^- + $6H^+$ + $6e^-$ → HS^- + $3H_2O$ $3HSO_3^-$ + $3H^+$ + $2e^-$ → $S_3O_6^{2-}$ + $3H_2O$ $2HSO_3^-$ + $4H^+$ + $4e^-$ → $S_2O_3^{2-}$ + $3H_2O$	Siro-heme	Cys Sγ	Fe(II) (S=1 or S=2)
Bacterioferritins	Iron uptake	*b*	Met Sδ/Met Sδ	Fe(II) (S=0) Fe(III) (S=1/2)
Guanylyl cyclase	Conversion of GTP to cGMP	*b*	His Nε	Fe(II) (S=2)
Cystathionine β-synthase	L-serine + L-homocysteine → cystathione + H_2O	*b*	Cys Sγ + His Nε	Fe(III) (S=1/2) Fe(II) (S=0)
Indoleamine 2,3-dioxygenase	L-tryptophan + O_2 → N-formylkynurenine	*b*	His	Fe(II) (S=2) deoxy
Tryptophan 2,3-dioxygenase	L-tryptophan + O_2 → N-formylkynurenine	*b*	His	Fe(II) (S=2) deoxy

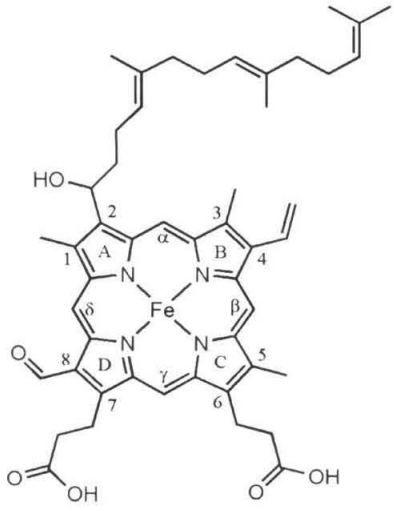
Table 1.2. Biological functions of representative iron-sulfur proteins (Adapted from (*56*)).

Function	Cluster	Protein
Electron transfer	$[Cys_4Fe]^{-/2-}$	Rubredoxin, desulforedoxin
	2Fe-2S	Rieske proteins
	2Fe-2S and/or 3Fe-4S and/or 4Fe-4S	Ferredoxins
		Iron-only hydrogenase
		Subunit B of fumarate reductase
	4Fe-4S	High-potential Fe-S protein
		Nitrogenase iron protein
		Trimethylamine dehydrogenase
		Pyruvate:ferredoxin oxidoreductase
Catalysis of a non-redox reaction	4Fe-4S	Aconitase
Catalysis of redox reactions	H cluster (4Fe-4S + 2Fe subcluster)	Fe-only hydrogenase
	4Fe-4S + siroheme	Sulfite reductase hemoprotein
Stabilization of protein structure for DNA repair	4Fe-4S	Endonuclease III, MutY
Sensing and regulation: (1) oxygen sensors: loss of original cluster and of activity	4Fe-4S	Glutamine PRPP amidotransferase
	4Fe-4S/2Fe-2S	FNR protein
	4Fe-4S/3Fe-4S	Aconitase
(2) sensor of O_2^- and NO: redox-regulated control of transcription	2Fe-2S	SoxR protein
(3) iron sensor: post-transcriptional regulation	Apoprotein/4Fe-4S	Iron regulatory protein/aconitase
Redox-mediated generation of free radicals	4Fe-4S	Anaerobic ribonucleotide reductase, pyruvate formate-lyase activating enzyme
		Biotin synthase
Stabilization of an intermediate in disulfide reduction	4Fe-4S	Ferredoxin:thioredoxin reductase

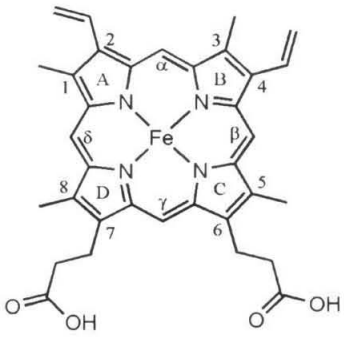
Table 1.3. Four main families of iron-oxygen/nitrogen proteins, with representative proteins from each family and their functions (Adapted from (*2*)).

Protein	Function	Secondary electon donor(s)
Substrate-activating mononuclear Fe(III) enzymes:		
Lipoxygenases	Hydroperoxidation	None
Intradiol dioxygenases	Oxidative catechol ring cleavage	None
O_2-activating mononuclear Fe(II) enzymes (2-His-1-Asp/Glu proteins):		
Pterin-dependent hydroxylases	Aromatic hydroxylation	Tetrahydrobiopterin
2-Oxoglutarate-dependent oxidases	Alkane hydroxylation / $2e^-$ oxidations	2-Oxoglutarate
Isopenicillin-N synthase	Double oxidative ring formation	None
Type I extradiol dioxygenases	Oxidative catechol ring cleavage	None
4-Hydroxyphenylpyruvate dioxygenase	Coupled decarboxylation and aromatic hydroxylation	None
Type II extradiol dioxygenases	Oxidative catechol ring cleavage	None
Rieske-type cis-dioxygenases	Aromatic / aliphatic cis-dihydroxylation	NADPH-flavin reductase → Rieske center
O_2-activating/binding diiron proteins (diiron carboxylate proteins):		
Hemerythrin	Reversible O_2 binding	
Ribonucleotide reductases R2	$1e^-$ Tyrosine oxidation	Fe(II) (?)
Methane monooxygenase	Alkane hydroxylation	NADP flavin reductase → [2Fe-2S] site
Stearoyl-carrier Δ-desaturase	Alkane desaturation	NADP flavin reductase → [2Fe-2S] ferredoxin
Rubrerythrin	O_2 / O_2^- / H_2O_2 scavenger (?)	NADPH, Fe(II) (?)
Ferritins / bacterioferritins	Ferroxidase	Fe(II)
Iron-dependent hydrolases:		
Purple acid phosphatases	Phosphoryl transfer	
Nitrile hydratase	Nitrile hydration	

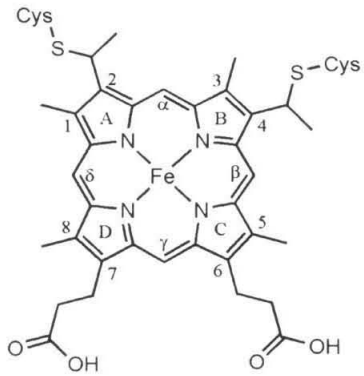
Figure 1.1. Naturally occurring iron porphyrins (adapted from (*9*)).



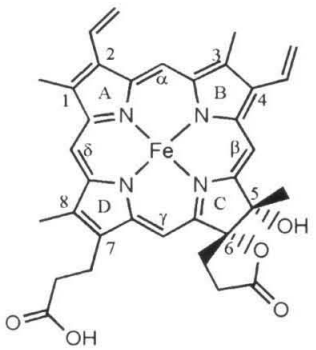
Heme *a*



Heme *b*

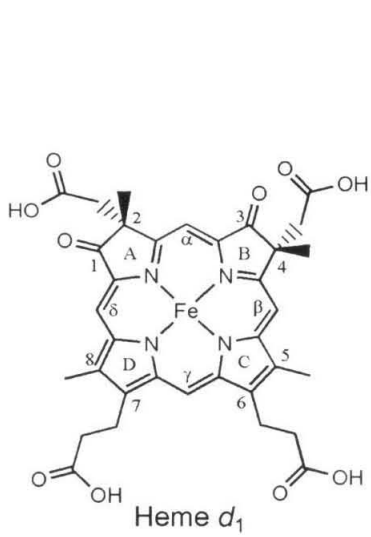


Heme *c*

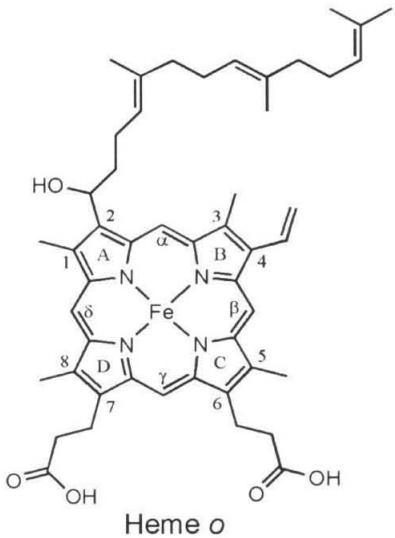


Heme *d*

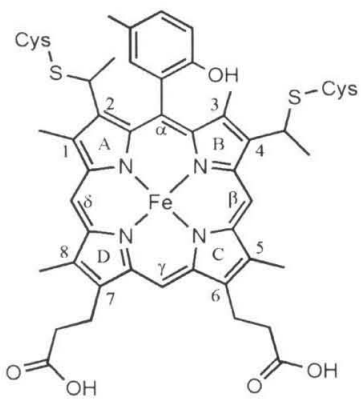
Figure 1.1. *continued*



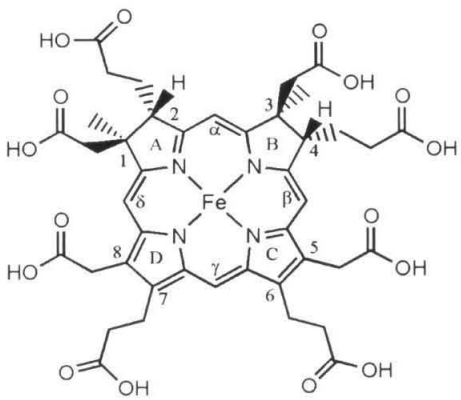
Heme d_1



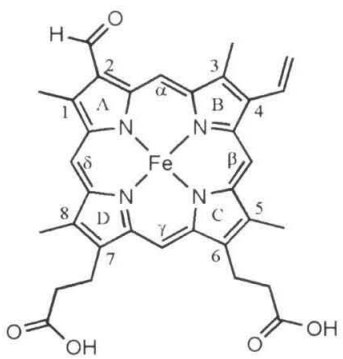
Heme *o*



Heme P460



Siroheme



Chloroheme

Figure 1.2. Common Fe-S cluster types found in Fe-S proteins.

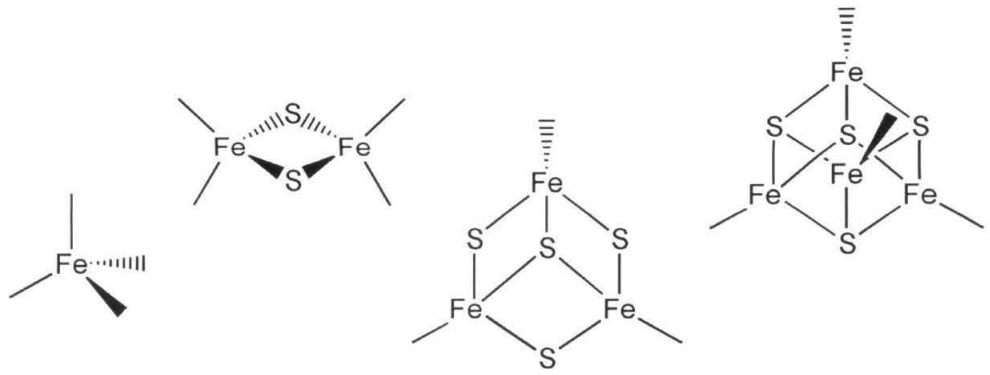


Figure 1.3. Examples of several complex metalloclusters: (a) the P-cluster and (b) FeMoco cluster of nitrogenase (*26, 27*), (c) [4Fe-4S] cluster bridged by Cys Sγ to a heme group in sulfite reductase (*30*), and (d) the [Ni-4Fe-5S] cluster of carbon monoxide dehydrogenase (*31*). Iron, nickel, molybdenum, sulfur, phosphorus, oxygen, nitrogen, and carbon atoms are colored green, cyan, magenta, yellow, lavender, red, blue, and gray, respectively. PDB entries 3MIN, 1AOP, and 1JJY were used for this figure.

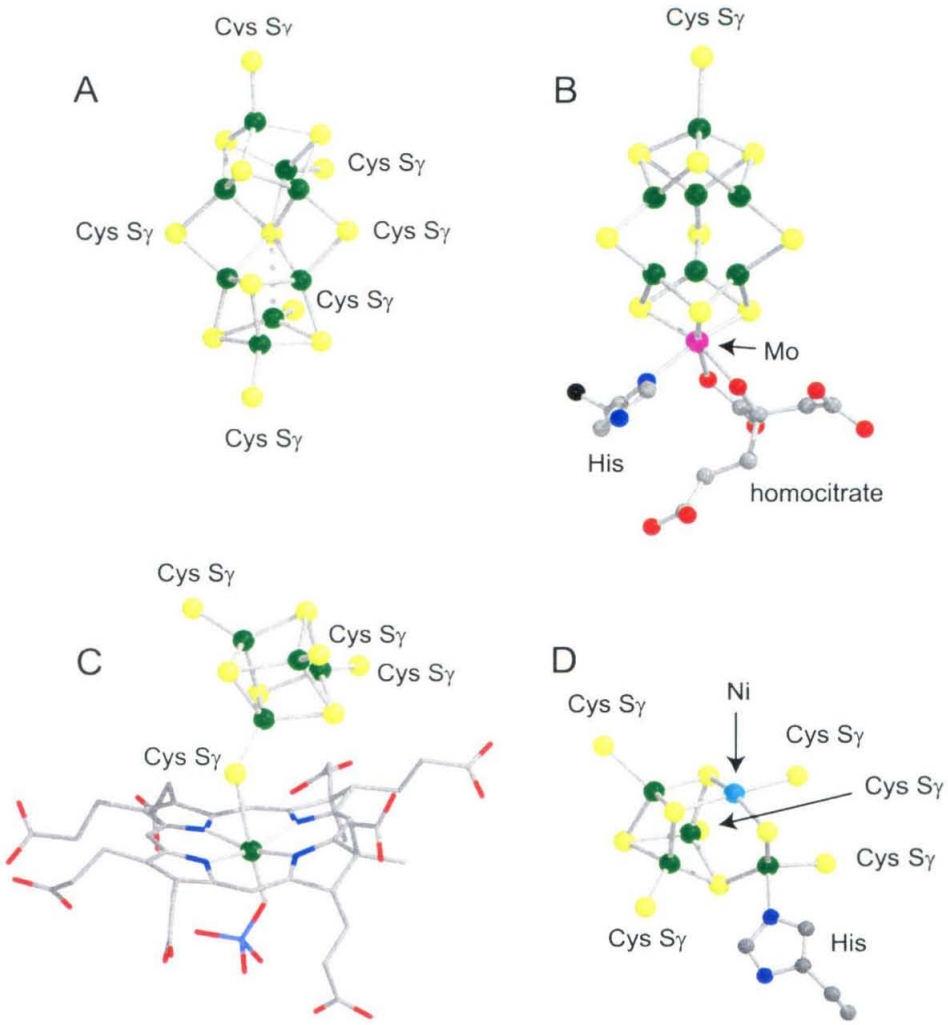


Figure 1.4. Spectrum of experimentally determined reduction potentials (vs. NHE) of various groups of Fe-S proteins, showing the broad range of reduction potentials attainable by Fe-S clusters (Adapted from (*57*)).

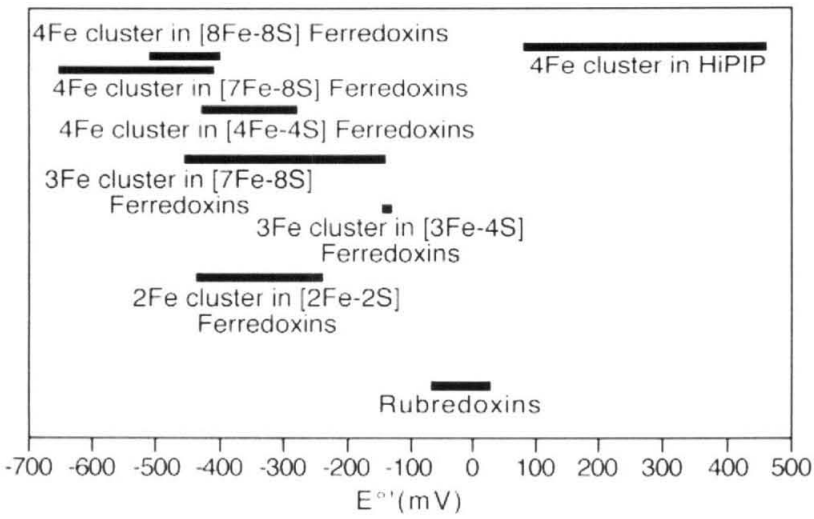
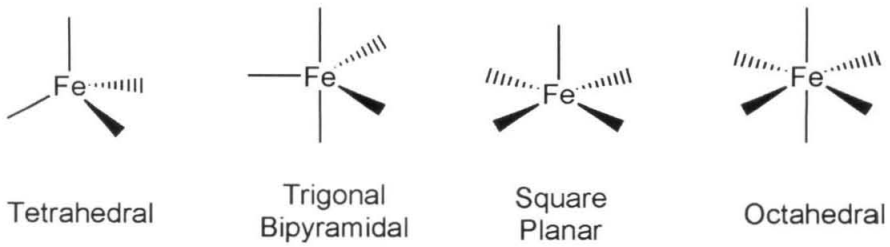


Figure 1.5. Common iron coordination geometries observed in Fe-O/N proteins.



CHAPTER 2

Structures of the Superoxide Reductase from *Pyrococcus furiosus* in the Oxidized and Reduced States *

* Adapted from Yeh, A. P., Hu, Y., Jenney, F. E. Jr, Adams, M. W. W., and Rees, D. C. (2000) *Biochemistry 39*, 2499-2508.

Structures of the Superoxide Reductase from *Pyrococcus furiosus* in the Oxidized and Reduced States[†,‡]

Andrew P. Yeh[§], Yonglin Hu[§,°], Francis E. Jenney Jr.[⊥], Michael W.W. Adams[⊥], and Douglas C. Rees[§,¶,*]

† This work was supported in part by grants from the Department of Energy (FG05-95ER20175) and the National Science Foundation (MCB 9809060) to M.W.W.A.. A.P.Y. was supported in part by a National Science Foundation Graduate Research Fellowship. The protein crystallographic facility at the Stanford Synchrotron Radiation Laboratory (SSRL) is funded by the Department of Energy, Office of Basic Energy Sciences, the National Institutes of Health, National Center for Research Resources, Biomedical Technology Program, and the Department of Energy, Office of Biological and Environmental Research.

‡ Coordinates have been deposited in the RCSB Protein Data Bank, entries 1DO6, 1DQI, and 1DQK.

§ Division of Chemistry and Chemical Engineering, 147-75CH, California Institute of Technology, Pasadena, CA 91125

° Present address: Howard Hughes Medical Institute, Department of Physiology and Microbiology and Molecular Genetics, University of California at Los Angeles, Los Angeles, CA 90095

⊥ Department of Biochemistry and Molecular Biology, University of Georgia, Athens, GA 30602

¶ Howard Hughes Medical Institute

* To whom correspondence should be addressed. Telephone: 626-395-8393; Fax: 626-744-9524; E-mail: dcrees@caltech.edu

Abstract

Superoxide reductase (SOR) is a blue non-heme iron protein that functions in anaerobic microbes as a defense mechanism against reactive oxygen species by catalyzing the reduction of superoxide to hydrogen peroxide [Jenney Jr., F. E., Verhagen, M. F. J. M., Cui, X., and Adams, M. W. W. (1999) *Science 286*, 306-309]. Crystal structures of SOR from the hyperthermophilic archaeon *Pyrococcus furiosus* have been determined in the oxidized and reduced forms to resolutions of 1.7 Å and 2.0 Å, respectively. SOR forms a homotetramer, with each subunit adopting an immunoglobulin-like beta barrel fold that coordinates a mononuclear, non-heme iron center. The protein fold and metal center are similar to those observed previously for the homologous protein desulfoferrodoxin from *Desulfovibrio desulfuricans* [Coelho, A. V., Matias, P., Fülöp, V., Thompson, A., Gonzalez, A., and Carrondo, M. A. (1997) *J. Bioinorg. Chem 2*, 680-689]. Each iron is coordinated to imidazole nitrogens of four histidines in a planar arrangement, with a cysteine ligand occupying an axial position normal to this plane. In two of the subunits of the oxidized structure, a glutamate carboxylate serves as the sixth ligand to form an overall six-coordinate, octahedral coordinate environment. In the remaining two subunits, the sixth coordination site is either vacant or occupied by solvent molecules. The iron centers in all four subunits of the reduced structure exhibit pentacoordination. The structures of the oxidized and reduced forms of SOR suggest a mechanism by which superoxide accessibility may be controlled and define a possible binding site for rubredoxin, the likely physiological electron donor to SOR.

Introduction

Superoxide reductases (SORs) are mononuclear, non-heme iron proteins that catalyze the one electron reduction of superoxide to peroxide (*1*):

$$1a \qquad SOR{\bullet}Fe^{+2} + O_2^{-} + 2H^{+} \rightarrow SOR{\bullet}Fe^{+3} + H_2O_2$$

$$1b \qquad D_{red} + SOR{\bullet}Fe^{+3} \rightarrow D_{ox} + SOR{\bullet}Fe^{+2}$$

The electron donor for reduction reaction 1b is likely a second non-heme iron protein, rubredoxin, which is reduced by an oxidoreductase using NAD(P)H (*1*). SORs are distinct from superoxide dismutases (SODs) since SORs do not couple the reduction of one superoxide molecule to the oxidation of a second superoxide to yield dioxygen. It has recently been proposed (*1*) that SORs participate in a novel oxygen detoxification system that is utilized by all anaerobic microorganisms that cannot tolerate the production of dioxygen generated from superoxide by the action of SOD.

The SOR from the hyperthermophilic archaeon *Pyrococcus furiosus* is the most extensively characterized member of this family (*1*); it is organized as a homotetramer composed of 14.3 kDa subunits that each contains a single mononuclear non-heme iron. The oxidized enzyme exhibits a characteristic blue color that prompted the name neelaredoxin (derived from *neela*, the Sanskrit word for blue) originally assigned to a homologous protein (50% identical to the *P. furiosus* protein) of undetermined function isolated from *Desulfovibrio gigas* (*2*). SOR also shares extensive sequence similarity

(40% identity) with the 93 residue C-terminal domain of *Desulfovibrio desulfuricans* desulfoferrodoxin (DFX), which is an electron transfer protein of unassigned function (*3*). A recent crystal structure analysis of DFX (*4*) demonstrated that this domain binds Fe in a square planar coordination geometry generated by four histidine rings in a plane, with a fifth ligand, cysteine, occupying an axial position.

In view of the biochemical and spectroscopic characterization of the SOR which established the SOR activity of this center, and the available structural analyses of the likely physiological electron donor to this center, *P. furiosus* rubredoxin (*5, 6*), we have determined the crystal structures of the *P. furiosus* SOR in the oxidized state (room temperature and low temperature to resolutions of 2.0 Å and 1.7 Å, respectively) and in the reduced state at low temperature to 2.0 Å resolution, to provide a structural framework for the interpretation of the spectroscopic properties and mechanistic features of this protein.

Results and Discussion

Structural organization of SOR

SOR exists as a homotetramer with 222 point group symmetry in the crystal structure (Figure 2.2). The overall dimensions of the tetramer are approximately 45 Å x 49 Å x 53 Å. The SOR subunit is organized as a seven stranded β-barrel that adopts an immunoglobulin-like fold, with one turn of 3_{10} helix (residues 1-4) connected to the barrel by a 15 residue long loop at the N-terminus (Figure 2.3A). As expected, this

structure is similar to the fold observed for the C-terminal domain of desulfoferrodoxin (DFX), with an RMSD of 1.3 Å between 72 Cα atoms of SOR-OxLT and DFX) (Figure 2.4). The topology of the seven β-strands corresponds to that of a C2-type Ig-like fold, which is characterized by the strand order A, B, E, D, C, F, and G (*7*) (Figure 2.3B). The three stranded antiparallel β-sheet contains β-strands A (residues 19-25), B (residues 28-34), and E (residues 92-98) while the four stranded β-sheet consists of β-strands C (residues 49-58), D (residues 65-71), F (residues 103-111), and G (residues 115-123). The main difference between SOR and the C-terminal domain of DFX is that the loop regions in SOR are longer. In SOR, loops AB, CD, DE, and EF are longer than they are in DFX by 4, 6, 13, and 3 residues, respectively (Figure 2.4).

In the SOR-OxRT model, subunits A and C are crystallographically equivalent, as are subunits B and D. Subunits A and B, related by noncrystallographic symmetry (as are subunits C and D) are nearly identical to each other, with two regions of notable exception (Figure 2.5A). One region spans Gly 9 to Lys 15, while the other includes Gly 36 to Pro 40. The difference between the two subunits in the Gly 9 to Lys 15 region adjacent to the iron site is quite substantial (Figure 2.5B), with a RMSD of 6.4 Å between corresponding Cα atoms. The differences between the region spanning Gly 36 to Pro 40 are less drastic, with a RMSD between corresponding Cα atoms of 1.2 Å. Excluding these two regions, the Cα r.m.s. fit between the remaining 113 residues in the two subunits is 0.1 Å. Similar relationships between these regions are observed in the SOR-OxLT structure.

Such drastic differences between the subunits in these two regions were not observed, however, in the SOR-RedLT structure. The average RMSDs between the Cα atoms in the different subunits of the regions spanning Gly 9 to Lys 15 and Gly 36 to Pro 40 in this structure are only 0.3 Å and 0.5 Å, respectively, making all four subunits in the asymmetric unit nearly identical to one another.

Inter-subunit interactions

Each subunit forms extensive contacts with the other three subunits, resulting in the burial at subunit-subunit interfaces of ~40% of the surface area of each isolated subunit, as calculated with GRASP (*8*). Two of the NCS-related subunits in SOR-OxRT, A and B (which are equivalent to C and D, respectively), form a tight dimer that is held together by many hydrogen bonds and electrostatic interactions. The interface area between these two subunits, calculated using GRASP, was found to be ~1340 $Å^2$ per subunit. The numerous interactions that are formed involve residues from the 3_{10} helix, the loop region connecting β-strands C and D, and β-strands C, D, F, and G. The two crystallographically related subunits (A and C) also interact extensively, with the surface area buried at this interface calculated to be ~1360 $Å^2$ per subunit. A notable feature of the subunits A and C pair is that the 4-stranded antiparallel β-sheets of the two subunits are juxtaposed such that β-strands D of subunits A and C pair up to form an extended 8-stranded antiparallel β-sheet (Figure 2.6A). A similar dimerization interaction between the 4-stranded β-sheets is observed in desulfoferrodoxin (*4*). In addition, in SOR,

residues 88-98 from β-strand E of subunit A form water-mediated as well as direct hydrogen bonds to the corresponding residues on β-strand E from subunit C (Figure 2.6B). The other NCS-related subunits A and D, equivalent to C and B, also interact with each other, but to a much lesser extent than the first pair of NCS-related subunits. The calculated dimer interface for subunits A and D (~590 $Å^2$ per subunit) is less than half that of either of the other two NCS-related subunit pairs. Finally, it is interesting to note that at the core of the homotetramer, near where the three molecular twofold axes intersect, are located a number of buried charged residues. Arginine 69 and glutamates 53 and 71 from each of the subunits are positioned near the twofold axes such that they form direct and water-mediate hydrogen bonds as well as electrostatic interactions with one another.

Iron center

Each SOR subunit in the homotetramer contains one mononuclear, non-heme iron center that resides on the periphery of the subunit between three loop regions—the loop region between the 3_{10} helix and β-strand A, the loop region between β-strands B and C, and the loop region between β-strands F and G (Figure 2.3A). The coordination geometry in subunits A and C of the oxidized structures may be described as octahedral, with four histidine ligands (His 16, His 41, His 47, and His 114) arranged in an equatorial plane, and a cysteine ligand (Cys 111) and a glutamate ligand (Glu 14) positioned axially (Figure 2.7A). Three of the four histidines (His 16, His 41, and His 47) coordinate the

iron via their Nε2 atoms while the fourth histidine, His 114, coordinates the iron via its Nδ2 atom.

Differences are observed in the sixth coordination position in subunits B and D of oxidized SOR. In SOR-OxRT, where Glu 14 is the sixth ligand in subunits A and C (which are identical due to crystallographic symmetry), no ligand is present at this position in subunits B and D (which are also crystallographically related). This pentacoordination geometry, as shown in Figure 2.7B, is similar to that observed for the iron center seen in domain II of DFX (*4*).

The iron centers seen in SOR-OxLT are indistinguishable from those seen in SOR-OxRT in most respects except for one. Where a sixth iron ligand was not observed in subunits B and D in the SOR-OxRT model, electron density for a solvent molecule in that position is now observed ~2.6 Å from the iron. The distances between the iron center and its ligands for all three SOR structures are listed in Table 2.3.

The differences in metal coordination observed between the iron centers of subunits A/C and B/D in both oxidized SOR structures reflect the substantial differences, as discussed earlier, in the polypeptide conformation for residues Gly 9 to Lys 15 (Figure 2.5B). In subunits B and D, this region is positioned such that Glu 14 is displaced from the iron center. The weak electron density, as well as the considerably higher temperature factors observed in this region of subunits B and D, indicate that it is probably somewhat disordered. As further evidence for the greater disorder of this region in subunits B/D

than in subunits A/C, the refined temperature factor for the iron atoms in subunits B/D (26 $Å^2$) of the SOR-OxRT model is higher than that of the iron atoms in subunits A/C (17$Å^2$). Whether or not the flexibility observed in this Gly 9-Lys 15 loop region is mechanistically relevant is unclear. It is interesting to note that the other loop region that differs between the subunits, Gly 36-Pro 40, is in close contact with the Gly 9-Lys 15 loop region. It seems plausible, therefore, that the interactions between these two regions of the subunit are correlated and could play a role in the catalytic mechanism of the enzyme.

The origin of the differences in the polypeptide conformations and the resulting differences in iron coordination between the different subunits of oxidized SOR is unclear. One possibility is that oxidized SOR may exist in an equilibrium between two different conformations and that the observed structures represent a trapped mixture of the two different oxidized conformations (with subunits A and C displaying one conformation and subunits B and D displaying the other). Yet another possibility is that the conformational differences observed in the oxidized structures arise from crystal packing. The molecules are packed in such a way that Trp 11 and Lys 12 of subunits A and C form lattice contacts with residues from an adjacent crystallographically related molecule while the same residues in subunits B and D are not in contact with any residues from another molecule. It might be argued then that the crystal packing interactions present in subunits A and C could be the cause for the greater stability observed in the Gly 9-Lys 15 loop region of these subunits. However, analysis of the reduced structure indicates otherwise. As will be discussed subsequently, all four subunits in the reduced

SOR structure adopt the same conformation in the flexible regions. The homogeneity of the subunits in the reduced structure suggests that crystal packing is unlikely to be the sole determinant of the conformation of the flexible regions. While crystal packing may still play a role, the reduced structure shows that the conformational differences observed in the oxidized structures are not necessarily enforced by crystal packing. A third possibility is that the iron sites of two of the subunits were photoreduced by x-rays during data collection, as has been previously observed for Cu,Zn SOD (*9, 10*), and the resulting structures represent a mixture of oxidized and reduced states. Such a scenario appears to be more likely than either of the first two possibilities. The conformations of the flexible regions in subunits B and D of the oxidized structures are the same as that of the corresponding regions in all four subunits of the reduced structure, suggesting that subunits B and D in the oxidized structures might actually be in the reduced state while subunits A and C are in the oxidized state.

Low temperature reduced form

The structure of dithionite-reduced SOR was determined to ascertain whether or not changes to the overall structure and/or iron coordination occur upon iron reduction. As discussed above, in contrast to the different polypeptide conformations observed in the oxidized structures, the structure of SOR in the reduced form shows that the conformations of the Gly 9-Lys 15 and Gly 36-Pro 40 loop regions are the same in all four subunits. These regions in the SOR-RedLT structure share very similar conformations to the corresponding regions in subunits B and D of the SOR-OxRT

structure, with the Cα RMSDs between these regions of the two structures being 0.4 Å and 0.6 Å.

The iron centers of reduced SOR display the same pentacoordinate configuration in all four subunits. In subunits A and C, Glu 14 is no longer coordinated to the iron, as it was in the oxidized structure; in addition, the solvent molecule that was observed as a sixth ligand to the iron in subunits B and D of the SOR-OxLT model is no longer observed. While the iron center coordination is the same among the four subunits of SOR-RedLT, there are some notable differences between the sites in the apparent occupancies of the iron atom at each site. In both subunits B and D, the electron densities of the iron and two of the histidines to which it is coordinated (histidines 16 and 114) are considerably weaker than those in subunits A and C. Assuming that this reflects partial loss of iron, rather than high mobility, the occupancies of Fe in the B and D subunits were determined to be ~0.2 through crystallographic refinement. Because of this low occupancy, the metal-ligand distances at these sites cannot be reliably determined, which likely contributes to the apparent changes in the values reported for SOR-RedLT subunits B and D in Table 2.3.

Mechanistic implications

The fact that the sixth iron ligand is variable between the subunits of the oxidized structure as well as between the oxidized and reduced structures suggests that this variability could play a role in the catalytic mechanism of SOR. One possible role of having a variable sixth ligand may be to regulate the accessibility of superoxide to the

iron site in the reduced and oxidized forms of SOR. According to equation 1a, superoxide should bind to SOR in the Fe(II) form. In the SOR-RedLT structure, the iron site is pentacoordinate, leaving the iron at the sixth ligand position exposed. One could speculate that superoxide may bind to the iron at this vacant sixth ligand position, forming a six-coordinate iron center; the superoxide would subsequently be reduced and released. In contrast, the iron centers in at least two of the subunits in the oxidized form of SOR are six-coordinate, with Glu 14 serving as the sixth ligand. In these subunits, the binding of superoxide to the iron would most likely be hindered, if not altogether prevented, by Glu 14. Such a gating mechanism could explain the binding of superoxide to the reduced, and not the oxidized, iron center.

It is not apparent from the structure, however, how binding of superoxide to the iron center is facilitated. In the case of both Cu,Zn SODs and also Mn and Fe SODs, it has been proposed that electrostatically positive residues around and along the substrate channel leading to the active site center serve to guide anions such as superoxide to the active site (*11-14*). This does not appear to be the case, however, for SOR. One difference between SOD and SOR is the apparent ease of accessibility of superoxide to the respective metal centers. In SOD, the active site is located at the bottom of a channel, whereas the iron center of SOR lies near the surface of the subunit. The seemingly more accessible active site of SOR may obviate the need for an electrostatic focusing effect such as that proposed for SOD. Lys 15, the closest positively charged residue to the iron center, could perhaps play some role in directing superoxide to the active site, but aside from this residue, the environment around the active site is mostly hydrophobic, as

reflected by the surface electrostatic potential map calculated with GRASP (*8*) (Figure 2.8).

Upon substrate reduction, the Fe(III) center of SOR must be re-reduced to the Fe(II) form. As discussed in (*1*), the likely electron donor for this electron transfer reaction is the non-heme iron containing protein rubredoxin (Rd). Since domain I of DFX contains a $FeCys_4$ rubredoxin-like center, the possibility that Rd is bound to SOR in a similar fashion as DFX domain I is attached to DFX domain II was explored. To test this possibility, one of the subunits of SOR was superimposed onto domain II of DFX while the structure of *P. furiosus* Rd (*5*) was superimposed onto domain I of DFX (RMSD of 1.3 Å between 21 Cα atoms). The superpositions revealed that *Pf* Rd cannot bind to SOR in the same position and orientation that domain I is attached to domain II of DFX, since this would result in *Pf* Rd directly overlapping with another subunit of SOR.

Inspection of the SOR structure points instead to a potential Rd binding site involving the immediate solvent-exposed residues surrounding the iron center. Such residues include three of the iron ligands, Glu 14, His 47, and His 114, as well as adjacent residues Trp 11, Ile 39, Pro 40, Pro 42, Thr 44, and Ile 113. One factor which suggests that Rd may bind at this site is the close proximity achievable between the SOR and Rd iron centers if Rd were indeed to dock at this site. Compared to other potential binding sites, the binding of Rd at this position would result in the shortest iron-to-iron distance (~8 Å), which perhaps would facilitate electron transfer between the two iron sites. In addition to the proximity between the iron centers, the hydrophobicity of the residues in

this area provides another indication that Rd may bind at this location, since the structure of *Pf* Rd (*5, 6*) reveals that the majority of the solvent-exposed residues surrounding the $FeCys_4$ center in Rd are also hydrophobic. Finally, many of the abovementioned residues surrounding the iron center in SOR are either a part of, or adjacent to, the flexible loop regions of Gly 9-Lys 15 and Gly 36-Pro 40. If these residues were indeed involved in binding Rd, the different conformations observed for the Gly 9-Lys 15 and Gly 36-Pro 40 regions between the oxidized and reduced structures could provide a mechanism by which Rd binding to SOR is regulated.

Materials and Methods

Crystallization and data collection

SOR was overexpressed and purified as described in (*15*). Crystals of SOR for data collection were prepared by the sitting drop vapor diffusion method, following a PEG screen, using 2 µl of ~100 mg/ml SOR in 50 mM Tris-HCl buffer at pH 8.0 and 2 µl of reservoir solution equilibrated against a reservoir solution containing 22% PEG 4000, 100 mM Tris-HCl at pH 8.0, 10% (v/v) glycerol, 200 mM NaCl, and 2% (v/v) ethanol. Crystals appeared within 4-5 days after setup and belong to the orthorhombic space group $P2_12_12$ (a = 50.31 Å, b = 94.02 Å, and c = 52.90 Å), with two SOR subunits in the asymmetric unit (solvent content ~44%).

Three native data sets were collected from SOR crystals: (a) SOR-OxRT: diffraction data to 2.0 Å resolution were collected from a capillary mounted, oxidized

crystal at room temperature on a Siemens multiwire area detector using monochromatized CuKα radiation produced by a Siemens rotating anode generator operated at 50 kV, 90 mA. The data set was processed and scaled using the XDS package (*16*). (b) SOR-OxLT: After soaking in a mother liquor solution containing 20% (v/v) PEG 400, SOR crystals could be flash-cooled. This development permitted collection of a second oxidized native data set at ~90 K to 1.7 Å resolution on beamline 9-1 at the Stanford Synchrotron Radiation Laboratory (λ=0.98 Å) on a 34.5 cm MAR Research imaging plate system. The data set was processed and scaled using DENZO and SCALEPACK (*17*). Upon flash-cooling of the crystal, the unit cell dimension along the c axis nearly doubled in length, increasing from 52.90 Å to 99.33 Å, and the space group changed from $P2_12_12$ to $P2_12_12_1$ (Table 2.1). This change in the space group resulted from a slight shift in the crystal packing upon flash-cooling, causing the crystallographic twofold axis to become noncrystallographic. (c) SOR-RedLT: SOR crystals were reduced by placing a few grains of sodium dithionite in the crystal-containing drop and were deemed fully reduced when the crystals became colorless. A 2.0 Å resolution data set of SOR in its reduced form was collected under cryogenic temperatures on an R-Axis IV imaging plate area detector using monochromatized CuKα radiation produced by a Rigaku RU 200 rotating anode generator operated at 50 kV, 100 mA. The data set was processed and scaled using DENZO and SCALEPACK (*17*).

Heavy-atom derivatives and phasing

Heavy-atom derivatives for the SOR-OxRT crystal form were prepared by soaking SOR crystal(s) in mother liquor solutions containing either 10 mM trimethyl lead acetate

(TMLA), 2 mM $Pt(NH_3)_2Cl_2$, or 2 mM $HgCl_2$. In addition to those three heavy-atom compounds, xenon gas was used for heavy-atom derivatization. Derivatization with xenon gas was carried out by mounting an SOR crystal in a quartz capillary and applying ~15 atmospheres of xenon gas to the capillary approximately 30 minutes before and throughout data collection (*18*). All heavy-atom derivative data sets were collected at room temperature on an R-Axis IIC imaging plate area detector using monochromatized CuKα radiation produced by a Rigaku RU 200 rotating anode generator operating at 50 kV, 100 mA and processed using the programs DENZO and SCALEPACK (*17*). Heavy-atom sites, including the native Fe sites, were determined by a combination of anomalous Patterson, difference Patterson, and cross difference Fourier maps generated with the CCP4 suite of programs (*19*). A total of nine heavy-atom sites (two TMLA sites, three $Pt(NH_3)_2Cl_2$ sites, three $HgCl_2$ sites, and one xenon site) were found. These sites, in combination with the two native iron sites, were used to calculate phases to 3.0 Å resolution using MLPHARE (*20*). A summary of the statistics for all data sets and phasing is listed in Table 2.1.

Model building and refinement

The initial MIR electron density map for the SOR-OxRT structure was improved through the use of twofold noncrystallographic symmetry (NCS) averaging and solvent flattening using the program DM (*21*). An approximate Cα trace of the model was built using the program XDLMAPMAN (*22*) and was used as a starting point for model building. Initial refinement employed strict NCS constraints in the simulated annealing protocol of the program X-PLOR (*23*), using a bulk solvent correction and all data

between ~30 Å and 2.0 Å. Refinement cycles were alternated with model rebuilding against $(2|F_o|-|F_c|)$ σ_A-weighted and $(|F_o|-|F_c|)$ σ_A-weighted maps (*24*) using the molecular graphics program O (*25*). After several rounds of refinement and model rebuilding, the R-factor and R-free converged to 16.0% and 18.9%, respectively. The final SOR-OxRT model contains two subunits (2 x 124 amino acid residues), two iron atoms, and 128 water molecules. The overall structures of the two NCS-related subunits are very similar, with the exception of the regions spanning Gly 9 through Lys 15 and Gly 36 through Pro 40. Of the 128 water molecules, 56 are NCS-related while 72 are not. A representative segment of the final σ_A-weighted $(2|F_o|-|F_c|)$ electron density map calculated at 2.0 Å is shown in Figure 2.1. The final model displays good stereochemistry, with an r.m.s. deviations in bond lengths and bond angles of 0.010 Å and 1.49 degrees, respectively. The average temperature factors of subunits A and B are 25.0 $Å^2$ and 27.9 $Å^2$, respectively, while the average temperature factor for the 128 water molecules is 34.3 $Å^2$. In the Ramachandran plot, as calculated with PROCHECK (*26*), 92.9% of the residues are in the most favored regions while 7.1% of the residues are in additionally allowed regions. There are no residues in disallowed regions.

Since the space group of the crystals changed from $P2_12_12$ to $P2_12_12_1$ upon flash cooling, the SOR-OxLT structure was determined by molecular replacement using the molecular replacement package AMoRe (*27*). Successive rounds of refinement and model rebuilding resulted in the convergence of the R-factor and R-free to 20.3% and 21.7%, respectively. The final SOR-OxLT model contains four subunits (4 x 124 amino acid residues), four iron atoms, and 416 water molecules. As was the case for the SOR-

OxRT model, the overall structures of the four subunits in the low temperature oxidized SOR model are very similar, except for the regions spanning Gly 9 to Lys 15 and Lys 37 to Pro 40. The average B values for subunits A, B, C and D are 17.1 $Å^2$, 18.7 $Å^2$, 16.6 $Å^2$ and 18.9 $Å^2$, respectively. The r.m.s. deviations in bond lengths and bond angles of SOR-OxLT are 0.004 Å and 1.24 degrees, respectively. Further analysis with PROCHECK (*26*) showed that 94.9% of the residues reside in the most favored regions, while 5.1% of the residues reside in the additionally allowed regions of the Ramachandran plot.

The SOR-OxLT structure, excluding the iron atoms and water molecules, was used as the starting model for the refinement of the SOR-RedLT structure. Successive rounds of simulated annealing, positional, and individual B-factor refinement in X-PLOR (*23*) using all data between 30 and 2.0 Å were alternated with model rebuilding. The occupancies of Lys 12, Gly 13, Glu 14, and Lys 15 were set to 0.10 due to the absence of electron density for those residues. The iron atoms, which were initially left out of the model, were placed back into all four subunits upon observing $(2|F_o|-|F_c|)$, $(|F_o|-|F_c|)$, and anomalous difference electron density at the iron sites in all four subunits. However, the electron densities at the iron sites in subunits A and C were considerably stronger than those in subunits B and D, suggesting that the iron atoms in subunits B and D were either at lower occupancy or more disordered. To distinguish between the two situations, two different refinements were performed and the resulting maps were compared. In one case, the occupancies of the iron atoms in subunits B and D were held constant at 1.0 while their B-factors were refined whereas in the other case, the B-factors of the two irons were fixed at their values in the SOR-OxLT model (~27 $Å^2$) while their occupancies were

refined. Comparison of the positive and negative difference density around the iron sites under each scenario suggested that the iron atoms in subunits B and D are present and ordered, but at low occupancy. The occupancies of the iron atoms in subunits B and D were ultimately refined to 0.22 and 0.24, respectively. Alternating rounds of refinement and rebuilding resulted in a final R-factor and R-free of 22.2% and 24.7%, respectively.

The final model of SOR in the reduced form at low temperature consists of four subunits (4 x 124 amino acid residues), four iron atoms, and 330 water molecules. The r.m.s. deviations from ideal bond lengths and bond angles are 0.006 Å and 1.58 degrees, respectively, while in the Ramachandran plot 92.2% of the residues are in the most favored regions and 7.8% of the residues are in additionally allowed regions. The average B-factors for each of the subunits are 26.6 $Å^2$, 27.6 $Å^2$, 26.2 $Å^2$, and 28.4 $Å^2$ and that for the water molecules is 40.7 $Å^2$. Refinement statistics for all structures are summarized in Table 2.2.

Acknowledgments

We greatly appreciate the contributions of Caroline Kisker, Hong Li, Hermann Schindelin, and Xiao-Dong Su to the structure determination.

References

1. Jenney Jr., F. E., Verhagen, M. F. J. M., Cui, X., and Adams, M. W. W. (1999) *Science 286*, 306-309.

2. Chen, L., Sharma, P., Le Gall, J., Mariano, A. M., Teixeira, M., and Xavier, A. V. (1994) *Eur. J. Biochem. 226*, 613-8.

3. Moura, I., Tavares, P., Moura, J. J. G., Ravi, N., Huynh, B.-H., Liu, M.-Y., and LeGall, J. (1990) *J. Biol. Chem. 265*, 21596-602.

4. Coelho, A. V., Matias, P., Fülöp, V., Thompson, A., Gonzalez, A., and Carrondo, M. A. (1997) *J. Bioinorg. Chem 2*, 680-689.

5. Day, M. W., Hsu, B. T., Joshua-Tor, L., Park, J.-B., Zhou, Z. H., Adams, M. W. W., and Rees, D. C. (1992) *Protein Sci. 1*, 1494-507.

6. Bau, R., Rees, D. C., Kurtz, D. M., Scott, R. A., Huang, H. S., Adams, M. W. W., and Eidsness, M. K. (1998) *J. Biol. Inorg. Chem. 3*, 484-493.

7. Vaughn, D. E., and Bjorkman, P. J. (1996) *Neuron 16*, 261-73.

8. Nicholls, A., Sharp, K. A., and Honig, B. (1991) *Proteins 11*, 281-96.

9. Stroppolo, M. E., Nuzzo, S., Pesce, A., Rosano, C., Battistoni, A., Bolognesi, M., Mobilio, S., and Desideri, A. (1998) *Biochem. Biophys. Res. Commun. 249*, 579-582.

10. Ferraroni, M., Rypniewski, W., Wilson, K. S., Viezzoli, M. S., Banci, L., Bertini, I., and Mangani, S. (1999) *J. Mol. Biol. 288*, 413-426.

11. Getzoff, E. D., Tainer, J. A., Weiner, P. K., Kollman, P. A., Richardson, J. S., and Richardson, D. C. (1983) *Nature 306*, 287-290.

12. *Klapper, I.*, Hagstrom, R., Fine, R., Sharp, K., and Honig, B. (1986) *Proteins: Structure, Function, and Genetics 1*, 47-59.

13. Sines, J., Allison, S., Wierzbicki, A., and McCammon, J. A. (1990) *J. Phys. Chem. 94*, 959-961.

14. Ludwig, M. L., Metzger, A. L., Pattridge, K. A., and Stallings, W. C. (1991) *J. Mol. Biol. 219*, 335-358.

15. Jenney Jr., F. E., Cui, X., Verhagen, M. F. J. M., and Adams, M. W. W. (1999) *(manuscript in preparation).*

16. Kabsch, W. (1988) *J. Appl. Crystallogr. 21*, 916-924.

17. Otwinowski, Z., and Minor, W. (1997) *Methods Enzymol. 276*, 307-326.

18. Soltis, S. M., Stowell, M. H. B., Wiener, M. C., Phillips Jr., G. N., and Rees, D. C. (1997) *J. Appl. Crystallogr. 30*, 190-194.

19. Bailey, S. (1994) *Acta Crystallogr. D50*, 760-763.

20. Otwinowski, Z. (1991) in *CCP4 Study Weekend Isomorphous Replacement and Anomalous Scattering* (Wolf, W., Evans., P., and Leslie, A., Eds.) pp 80-86, SERC Daresbury Laboratory, Warrington, UK.

21. Cowtan, K. (1994) *Joint CCP4 and ESF-EACBM Newsletter on Protein Crystallography 31*, 34-38.

22. Kleywegt, G. J., and Jones, T. A. (1996) *Acta Crystallogr. D52*, 826-828.

23. Brünger, A. T. (1992) *X-PLOR Manual Version 3.1--A System for X-ray Crystallography and NMR*, Yale University, New Haven, CT.

24. Read, R. J. (1986) *Acta Crystallogr. A42*, 140-149.

25. Jones, T. A., Zou, J. Y., Cowan, S. W., and Kjeldgaard, M. (1991) *Acta*

Crystallogr. A47, 110-119.

26. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) *J. Appl. Crystallogr. 26*, 283-291.

27. Navaza, J. (1994) *Acta Crystallogr. A50*, 157-163.

28. Esnouf, R. M. (1997) *J. Molec. Graphics 15*, 133-138.

29. Merritt, E. A., and Bacon, D. J. (1997) *Methods Enzymol. 277*, 505-524.

Table 2.1. Data sets and heavy-atom phasing statistics. The native anomalous signal and TMLA and $Pt(NH_3)_2Cl_2$ derivatives were phased to 3.0 Å, the $HgCl_2$ derivative was phased to 3.5 Å, while the xenon derivative was phased to 4.0 Å.

Data set	**SOR-OxRT**	**SOR-OxLT**	**SOR-RedLT**	**TMLA**	**$Pt(NH_3)_2Cl_2$**	**$HgCl_2$**	**Xe**
Source	CuKα	SSRL BL9-1	CuKα	CuKα	CuKα	CuKα	CuKα
Wavelength	1.54	0.98	1.54	1.54	1.54	1.54	1.54
Temperature	RT	90 K	90 K	RT	RT	RT	RT
Space Group	$P2_12_12$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12$	$P2_12_12$	$P2_12_12$	$P2_12_12$
Unit Cell Dimensions							
a	50.31	49.54	49.84	50.25	50.58	50.10	49.94
b	94.02	93.01	93.21	93.78	93.85	93.93	93.46
c	52.90	99.32	99.97	52.49	52.63	52.68	52.66
Maximum Resolution	2.0	1.7	2.0	2.5	3.0	2.8	3.0
No. of measurements	65376	222107	92775	24656	21012	23723	17184
Unique reflections	16334	51080	31825	8285	5358	6483	5243
Completeness (%) [b]	92.9 (84.2)	99.5 (99.8)	98.7 (95.7)	91.5 (79.9)	99.8 (98.4)	99.6 (97.0)	98.7 (96.8)
I/σ(I)	23.8 (4.4)	21.7 (5.5)	19.2 (4.7)	15.0 (6.8)	15.1 (8.3)	13.4 (4.8)	17.8 (8.3)
R_{sym} (%) [c]	6.2 (17.8)	6.3 (25.1)	6.0 (24.2)	7.9 (11.7)	8.7 (13.8)	10.1 (19.2)	6.5 (14.0)
Concentration of heavy-atom compound (mM)	-	-	-	10	2	2	~ 15 atm.
No. of sites	2	-	-	2	3	3	1
Phasing power [d]	-	-	-	1.04	0.72	0.96	0.86
R_{cullis} [e]	0.85	-	-	0.83	0.89	0.85	0.88
Overall FOM					0.515 (to 3.0 Å resolution)		

Table 2.1. continued

[a] Room temperature oxidized SOR data set was used for the initial MIR phasing.

[b] Numbers in parentheses correspond to values in the highest resolution shell.

[c] $R_{sym} = (\Sigma_{hkl} \Sigma_i | I_i(hkl) - \langle I(hkl) \rangle |) / (\Sigma_{hkl} \Sigma_i I(hkl))$.

[d] Phasing Power = $\langle F_H \rangle / \Sigma || F_{PH} - F_P | - F_H |$, where $\langle F_H \rangle$ is the r.m.s. calculated heavy-atom structure factor.

[e] $R_{cullis} = \Sigma || F_{PH} - F_P | - F_H | / \Sigma | F_{PH} - F_P |$, where F_H is the calculated heavy-atom structure-factor amplitude, and F_P and F_{PH} are the native and derivative structure-factor amplitudes, respectively.

Table 2.2. Final refinement statistics for the SOR models in the oxidized (SOR-OxRT and SOR-OxLT) and reduced (SOR-RedLT) forms.

	SOR-OxRT	**SOR-OxLT**	**SOR-RedLT**
Resolution limits (Å)	28.8 - 2.0	29.6 - 1.7	22.7 - 2.0
R-factor [a]	0.160	0.203	0.222
R-free	0.189	0.217	0.247
No. of monomers / asymmetric unit (a.u.)	2	4	4
No. of nonhydrogen atoms in a.u.			
Protein	2028	4056	4056
Iron	2	4	4
Water	128	416	330
RMS deviations from ideal values			
Bond lengths (Å)	0.010	0.004	0.006
Bond angles (°)	1.49	1.24	1.58
Dihedral angles (°)	29.9	29.6	29.3
Improper torsion angles (°)	0.73	0.55	0.68
Average temperature factor (Å^2)			
main-chain atoms	23.9	16.5	26.2
side-chain atoms	28.9	19.1	28.1
water molecules	34.3	31.3	40.7
Ramachandran plot			
residues in most favored regions (%)	92.9	94.9	92.2
residues in additional allowed regions (%)	7.1	5.1	7.8
residues in generously allowed regions (%)	0.0	0.0	0.0
residues in disallowed regions (%)	0.0	0.0	0.0

[a] R-factor = $\Sigma(|F_{obs}|-|F_{calc}|)/\Sigma|F_{obs}|$

Table 2.3. Bond distances between the iron atom and its ligands (X) in the different subunits of the oxidized (SOR-OxRT and SOR-OxLT) and reduced (SOR-RedLT) forms of SOR. The bond distances in subunits C and D of the SOR-OxRT model are not listed, as they are the same as those for subunits A and B, due to crystallographic symmetry. The bond distances reported for subunits B and D of the SOR-RedLT structure are less accurate due to the lower Fe occupancies at these sites and are therefore reported with lower precision values.

	SOR-OxRT		**SOR-OxLT**				**SOR-RedLT**			
X	**A**	**B**	**A**	**B**	**C**	**D**	**A**	**B**	**C**	**D**
His 16 Nε2	2.14	2.14	2.10	2.13	2.09	2.09	1.97	2.7	2.08	2.4
His 41 Nε2	2.15	2.21	2.17	2.21	2.20	2.24	2.29	2.7	2.26	3.1
His 47 Nε2	2.19	2.20	2.15	1.99	2.16	2.02	2.13	1.9	2.08	2.1
His 114 Nδ2	2.09	2.06	2.18	2.16	2.20	2.12	2.08	2.2	2.10	2.0
Cys 111 Sγ	2.42	2.33	2.46	2.68	2.46	2.67	2.44	2.6	2.40	2.7
Glu 14 Oε2	2.02	--	2.15	--	2.15	--	--	--	--	--
Solvent	--	--	--	2.57	--	2.59	--	--	--	--

Figure 2.1. Stereoview of the final σ_A-weighted ($2|F_o|-|F_c|$) electron density map around residues 48-53 of the SOR-OxRT structure. The map was calculated to 2.0 Å resolution and contoured at 1.0 σ. This and all subsequent figures, unless indicated otherwise, were produced using the programs BOBSCRIPT (*28*) and RASTER3D (*29*).

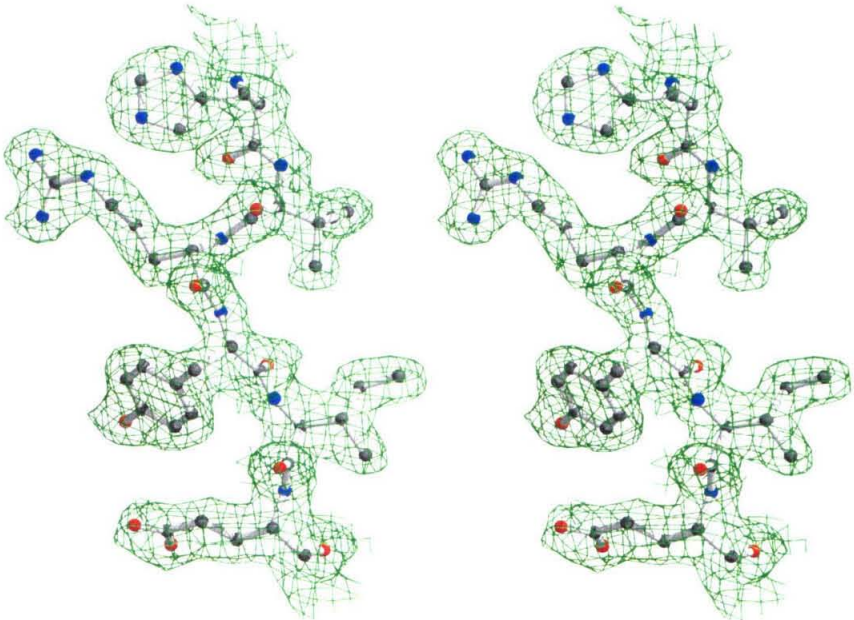


Figure 2.2. Ribbons diagram representation of the homotetrameric arrangement of SOR. Subunits A, B, C, and D are depicted in yellow, green, blue, and red, respectively, while iron atoms are depicted as gold spheres.

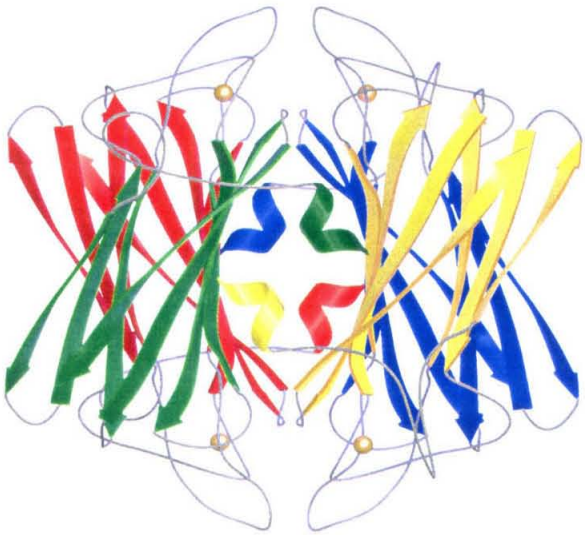


Figure 2.3. (a) Stereoview of the structure of a SOR monomer. β-strands in the 3-stranded and 4-stranded β-sheet are shown in light and dark blue, respectively. The iron atom is depicted as a green sphere. (b) Topology diagram of a SOR monomer. Secondary structures are depicted using the same color scheme as in (a). β-strands are lettered A through G, with the order of the strands being such that the topology is that of a C2 Ig-like fold. β-strands A, B, and E form the 3-stranded β-sheet while β-strands C, D, F, and G form the 4-stranded β-sheet. The iron center is shown as a green sphere and is connected by dotted lines to the general locations of its five, potentially six, ligands.

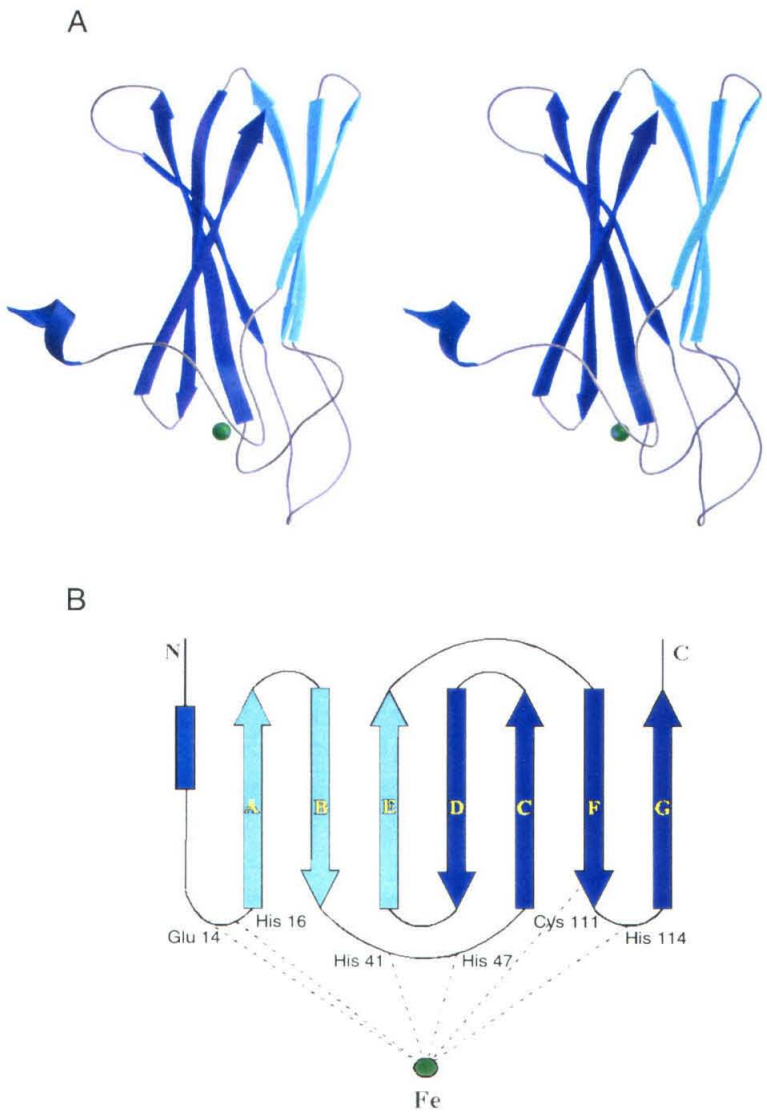


Figure 2.4. *Superposition* of SOR (yellow) and DFX domain II (darker purple) showing the similarity between the folds (RMSD between 72 Cα atoms of SOR-OxLT and DFX domain II is 1.3 Å). Domain I of DFX, which contains a rubredoxin-like $FeCys_4$ center, is shown in light purple.

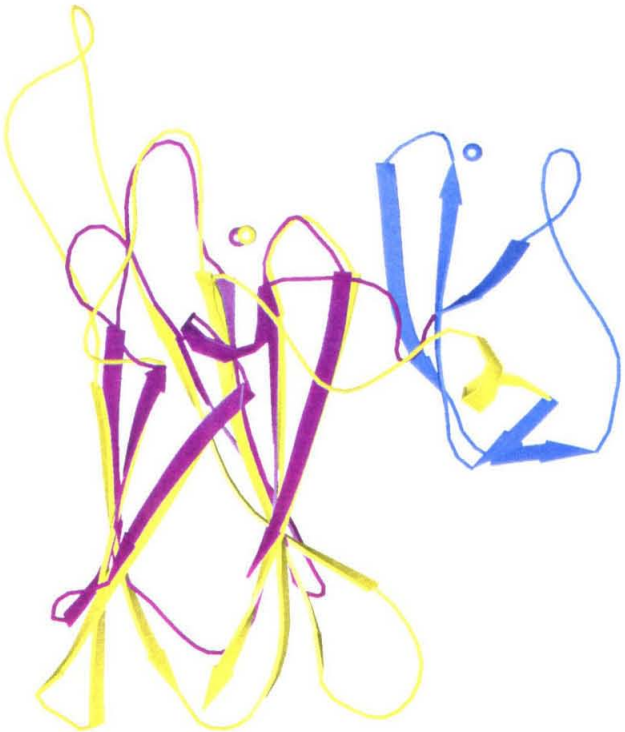


Figure 2.5. (a) Superposition of subunits A (yellow) and B (green) of SOR in the oxidized form. The two regions (Gly 9-Lys 15 and Gly 36-Pro 40) that differ between the subunits are shown in a darker shade of their respective colors. (b) Stereoview of the superposition of the iron centers of subunits A and B in oxidized SOR using the same color scheme as in (a). Glu 14 is located on one of the flexible loop regions (shown in a darker shade) which differ in conformation between subunits A and B. One result of this conformational difference is that in subunit A, Glu 14 can act as a sixth ligand to the iron whereas in subunit B, Glu 14 has moved ~5 Å away and as a result, no longer can act as a sixth ligand.

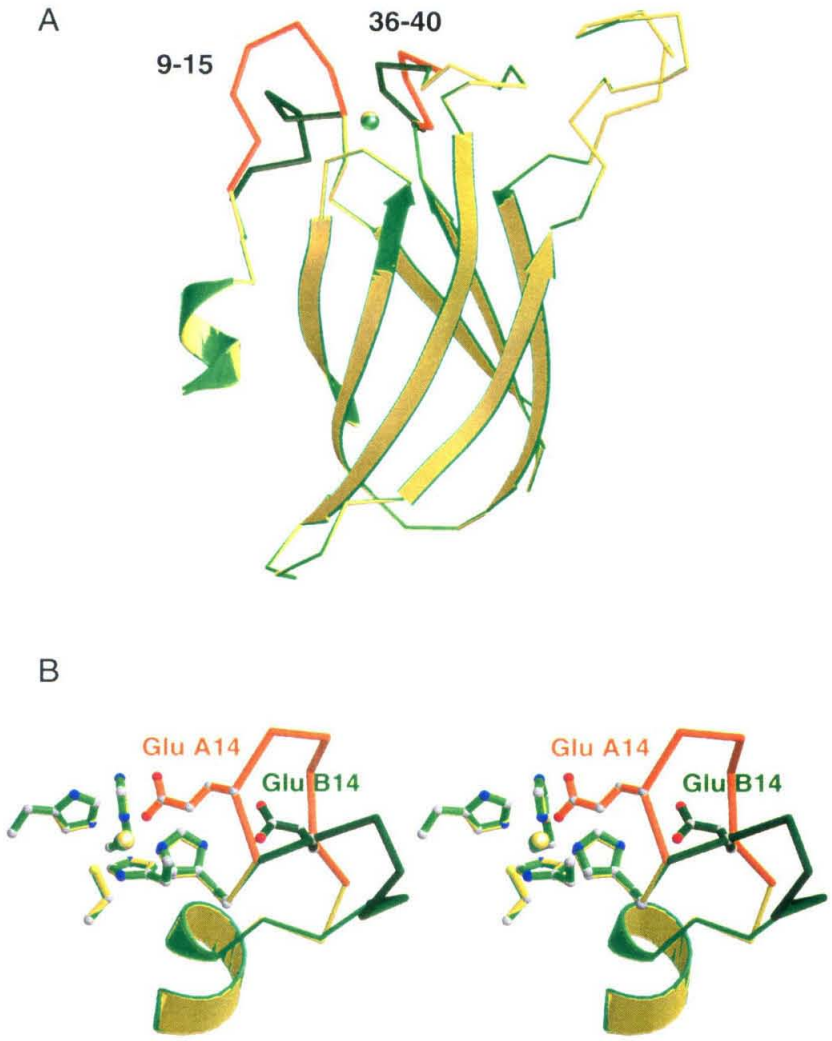


Figure 2.6. Dimerization interactions between subunits A (yellow) and C (blue) (the same interactions are observed for subunits B and D). As in Figure 3, lighter colored β-strands correspond to the 3-stranded β-sheet whereas darker colored β-strands correspond to the 4-stranded β-sheet. (a) View of subunits A and C showing hydrogen bonds formed between β-strands D of the two subunits, resulting in the formation of an extended 8-stranded antiparallel β-sheet. (b) View upon rotation of the molecule by 180° about the vertical axis in the plane of the paper, showing the direct and water-mediated hydrogen bonds that are formed between β-strands E of the two subunits.

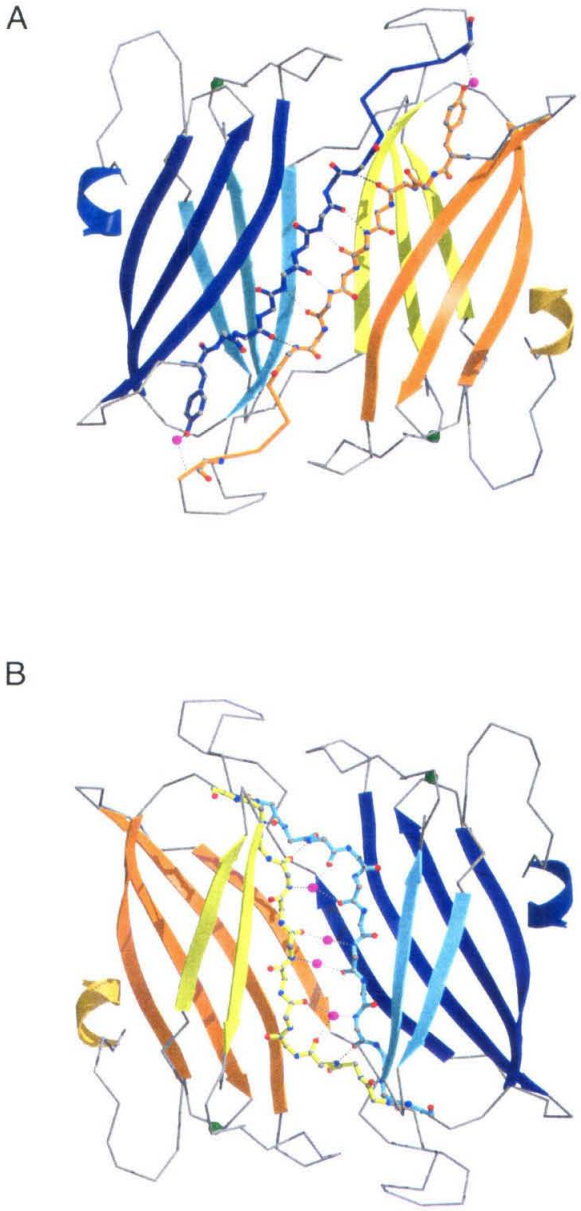


Figure 2.7. (a) Stereoview of the octahedral six-coordinate iron center observed in two (subunits A and C) of the four subunits of oxidized SOR. The iron, depicted as a green sphere, is coordinated in the equatorial positions by His 16, His 41, His 47, and His 114, and in the axial positions by Cys 111 and Glu 14. (b) Stereoview of the square pyramidal five-coordinate iron center observed in two (subunits B and D) of the four subunits of oxidized SOR as well as in all four subunits of reduced SOR. Glu 14 is no longer a ligand to the iron as a result of the Gly 9-Lys 15 loop region being in a different conformation in these subunits.

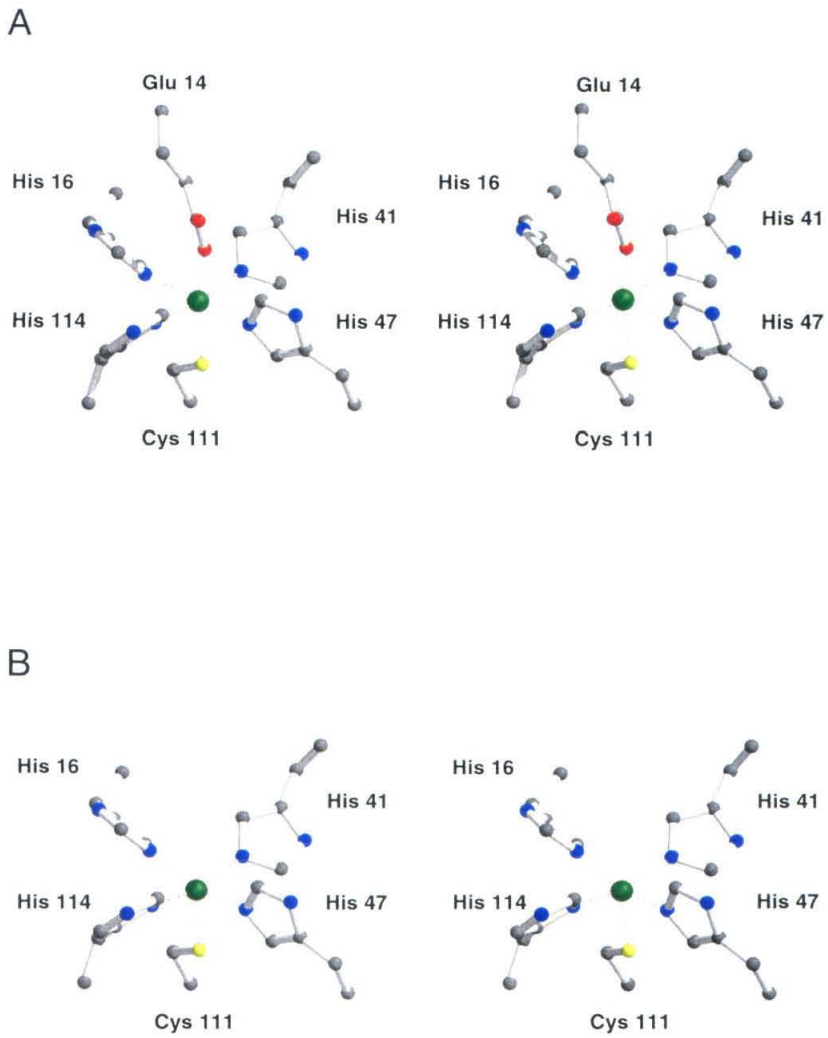
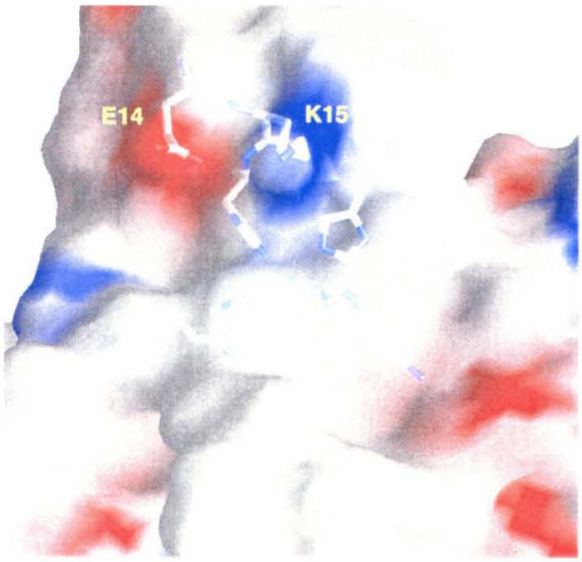


Figure 2.8. Electrostatic potential map contoured from -15 k_BT (red) to +15 k_BT (blue) showing the predominantly uncharged and relatively exposed environment of the iron center in reduced SOR. Glu 14 and Lys 15 are the closest charged residues to the iron, depicted as a green sphere. This figure was created using GRASP (*8*).



CHAPTER 3

Structure of a Thioredoxin-like [2Fe-2S] Ferredoxin from *Aquifex aeolicus* *

* Adapted from Yeh, A. P., Chatelet, C., Soltis, S. M., Kuhn, P., Meyer, J., and Rees, D. C. (2000) *J. Mol. Biol. 300*, 587-595.

Structure of a Thioredoxin-like [2Fe-2S] Ferredoxin from *Aquifex aeolicus*

Andrew P. Yeh[1], Claire Chatelet[2], S. Michael Soltis[3], Peter Kuhn[3],

Jacques Meyer[2], and Douglas C. Rees[1,4*]

1 Division of Chemistry and Chemical Engineering, 147-75CH, California Institute of Technology, Pasadena, California 91125

2 Département de Biologie Moléculaire et Structurale, CEA-Grenoble, F-38054, Grenoble, France

3 Stanford Synchrotron Radiation Laboratory, P.O. Box 4349, MS 69, Stanford, California 94309

4 Howard Hughes Medical Institute, California Institute of Technology, Pasadena, California 91125

* Corresponding author

E-mail address of the corresponding author: dcrees@caltech.edu

Abstract

The 2.3 Å resolution crystal structure of a [2Fe-2S] cluster-containing ferredoxin from *Aquifex aeolicus* reveals a thioredoxin-like fold that is novel among iron-sulfur proteins. The [2Fe-2S] cluster is located near the surface of the protein, at a site corresponding to that of the active site disulfide bridge in thioredoxin. The four cysteine ligands are located near the ends of two surface loops. Two of these ligands can be substituted by non-native cysteines introduced throughout a stretch of the polypeptide chain that forms a protruding loop extending away from the cluster. The presence of homologs of this ferredoxin as components of more complex anaerobic and aerobic electron transfer systems indicates that this is a versatile fold for biological redox processes.

Introduction

Proteins containing [2Fe-2S] clusters (*1-3*) participate in many important biological processes associated with oxidation-reduction reactions of both aerobic and anaerobic metabolism. [2Fe-2S] clusters are commonly incorporated into proteins or domains homologous to either the plant-type (*4, 5*) or the mammalian-type (*6*) ferredoxins. Other protein folds are also known to accommodate binuclear iron-sulfur clusters (*1, 2*); two of these have recently been structurally characterized in the Rieske-type proteins (*7*) and in the aldehyde oxidoreductase from *Desulfovibrio gigas* (*8*).

The [2Fe-2S] ferredoxin from the anaerobe *Clostridium pasteurianum* (CpFd) (*9*) has been recognized as distinct from the protein families outlined above on the basis of its amino acid sequence (*10*) and the spectroscopic properties of its cluster (*11*). CpFd is also distinguished by the ability of two of the cysteine ligands of the [2Fe-2S] cluster to undergo extensive ligand swapping (*11-13*), and the consequences of serine substitution for the cysteine ligands have been detailed in this system (*14-17*). While its biological function is still unclear, CpFd has recently been shown to interact specifically with the nitrogenase molybdenum-iron (MoFe) protein (*18*). Furthermore, an analysis of genomic sequences (see, e.g., http://www.tigr.org/tdb) has revealed that homologs of CpFd are pervasive throughout the currently sequenced genomes (Figure 3.1). The products of two of these genes, one from the hyperthermophilic bacterium *Aquifex aeolicus* (Fd4; the product of the *fdx4* gene) (*19*) and the other from the nitrogen fixing bacterium *Azotobacter vinelandii* (*20*), have recently been characterized (*21, 22*). Homologs have

also been identified as components of complex redox systems, including subunits or domains of hydrogenases (*23, 24*) and NADH-ubiquinone oxidoreductases (complex I of respiratory chains) (*23*).

In view of the widespread occurrence of this protein in diverse organisms, including the apparent involvement in nitrogen fixation, we have determined the crystal structure of the Fd4 [2Fe-2S] ferredoxin from *A. aeolicus*, a homolog to CpFd, to provide a structural framework for addressing the properties of this unusual class of iron-sulfur proteins.

Results and Discussion

Overall structure

Fd4 exists as a homodimer in the crystal structure (Figure 3.2ab), in agreement with previous biochemical studies (*21, 22, 25*). The structure of the Fd4 monomer is distinct from other structurally characterized [2Fe-2S] proteins (*26*). Consisting of five β-strands, two relatively long α-helices, and several additional short α- or 3_{10} helices, Fd4 bears an intriguing similarity to the overall α/β architecture of thioredoxin (*27-29*) (Figure 3.3). The rmsd between 76 equivalent Cα atoms is 2.6 Å, despite only ~7% sequence identity. The first four β-strands and the helices of Fd4 adopt a variation of the thioredoxin fold motif, which is characterized as having an N-terminal βαβ motif and a C-terminal ββα motif connected to each other by a third helix. Fd4 also contains a fifth β-strand that extends one edge of the thioredoxin fold, so that the overall β-strand order

from N- to C-terminus is 2-1-3-4-5, with strands 3 and 5 antiparallel to the other three strands (Figure 3.2c).

The two noncrystallographic symmetry related subunits of Fd4 are nearly identical (rmsd of 0.81 Å between 109 Cα atoms), with the notable exception of two regions. One region which differs between the two subunits is the extended surface loop connecting β-strand 1 to α-helix 1. This region, which displays an rmsd of 1.9 Å in Fd4, was previously found through site-directed mutagenesis experiments to be deletable in CpFd. Deletions of variable lengths (3 to 14 residues) in the region spanning residues 16-30 in CpFd can be tolerated without significant destabilization of the protein (*12*). It is worth noting that the length of this surface loop (~11 residues) is considerably longer than that of the corresponding loops (~1-5 residues) in other thioredoxin-fold containing proteins (*30*). Another difference between the two NCS-related subunits is the region spanning residues 35 through 47, which display an rmsd of 1.1 Å. As a result of this small difference, the secondary structure of this region in subunit A is a 3_{10} helix while in subunit B it is an α-helix.

The dimer interactions between the two subunits are mediated by twofold symmetric hydrogen bonds between the β2 strands as well as by predominantly apolar contacts involving the sidechains of Phe 3, His 5, Phe 7, Val 49, Pro 52, Gly 54, Tyr 68, Gly 104, Lys 105, and Pro 106. Approximately 810 $Å^2$ surface area per subunit is buried at this interface, as calculated with the program GRASP (*31*). The center-to-center

distance between the two [2Fe-2S] clusters in the dimer is ~21 Å, which is consistent with the absence of magnetic interactions observed in the EPR and Mössbauer spectra of the reduced protein (*11, 14, 15*).

[2Fe-2S] cluster

As illustrated in Figure 3.2, the [2Fe-2S] cluster resides on the periphery of each subunit between two loop regions: the loop between β-strand 1 and α-helix 1 and the loop between β-strands 2 and 3. The cluster is coordinated by the sidechains of Cys 9 and Cys 22 (liganded to Fe-1 of the cluster) and Cys 55 and Cys 59 (liganded to Fe-2). These observations confirm the ligand assignments previously established for CpFd (*11*). Fe-2 is the more solvent accessible of the two irons, which is consistent with the suggestion that the iron coordinated by Cys 55 and Cys 59 is more reducible (*11, 14, 16*) and hence is likely to be near the surface, as in plant-type ferredoxins (*5*). With average N-S distances of < 3.8 Å, possible hydrogen bonds to the cluster sulfides are between the Nδ1 atom of Arg 13 and S1, between the amide hydrogen of Cys 22 and S1, and between S2 and the amide hydrogens of Met 56, Asn 57, Ala 58, and Cys 59 (Figure 3.4a). Spectroscopic differences between plant-type and Fd4-type ferredoxins (*25*), including the valence-delocalized S=9/2 (*14, 15*) and S=1/2 (*17*) ground spin states of the reduced C56S and C60S variants of CpFd, likely reflect structural differences that cannot be distinguished at the present resolution of the structure.

Similarity to thioredoxin

Aside from the similarity in overall fold between Fd4 and thioredoxin, it is also interesting to note that two of the cysteine ligands of Fd4 (Cys9 and Cys22) occur in positions similar to those of the active site cysteines in thioredoxin (*27-29*) (Figure 3.3). This particular placement of cysteines in the region between β-strand 1 and α-helix 1 of the thioredoxin fold appears to be a relatively common theme, as other previously determined structures containing thioredoxin folds also have catalytically important residues located in these positions (*30*). Interestingly, thioredoxin has been adopted as a protein design platform for incorporating novel metallocenters into proteins (*32-35*), and one of these designs has resulted in the incorporation of a mononuclear iron-sulfur center at the active site disulfide bridge of thioredoxin (*36*).

In addition to the similar positioning of cysteines within the thioredoxin fold, another feature which Fd4 shares with a number of proteins in the thioredoxin superfamily is the existence of a cis-proline (Pro 63 in Fd4) between the second helix and β-strand 3. Analyses of the structures of the complexes of glutaredoxin and glutathione S-transferase and their substrate, glutathione, have shown that this conserved cis-proline forms a number of significant interactions with the glutathione (*37-39*). Although a cis-proline residue is not strictly conserved in the ferredoxin family (Figure 3.1), whether cis-Pro 63 plays a significant role in protein or substrate binding in Fd4 still remains to be determined.

The thioredoxin-like structure reported here is a novelty among iron-sulfur proteins. The addition of Fd4 to the set of proteins assuming the thioredoxin fold, which previously encompassed five classes (*30*), plus the recent addition of peroxiredoxin (*40*) and calsequestrin (*41*), raises new issues regarding the evolution of the thioredoxin fold and also reveals that thioredoxins and iron-sulfur proteins are not only functionally (*42*), but also, to at least some extent, structurally related.

Ligand substitutions at the [2Fe-2S] cluster

A variety of molecular variants were prepared to establish the substitution patterns of ligands to the [2Fe-2S] active site of CpFd (*11-13*) which can now be interpreted in light of the crystal structure (Figure 3.4b). Cys 24 (Cys 22 in Fd4), when mutated into serine or alanine, can be replaced as a cluster ligand by introducing a Cys at position 14 (Asp 12 in Fd4), which is spatially adjacent. Furthermore, the ability of Cys 24 to be replaced as a ligand of the [2Fe-2S] cluster by another cysteine introduced in positions 16 to 26 of CpFd is consistent with the location of these residues in the protruding loop and by the high affinity of thiolate ligands for Fe-S clusters (*2, 43*). Replacement of the Cys 60 ligand (Cys 59 in Fd4), when mutated into alanine, by a cysteine introduced in position 21 (19 in Fd4) was initially more puzzling (*13*), but this behavior is consistent with the relative proximity of the relevant residues.

It is also of interest to revisit the various consequences of the cysteine to serine mutations in CpFd (*11-13, 16*). In summary, mutations to serines of cysteine ligands which are located on the short loop region (i.e., Cys 55 and 59 in Fd4) were found not to

perturb protein stability whereas mutations of the other two cysteine ligands resulted in instability of the protein. The C56S and C60S variants of CpFd (Cys 55 and 59 in Fd4) were relatively stable, and had similar properties indicative of serine ligation. Both ligands occur on a short, solvent exposed loop that is presumably constrained enough to stabilize the less favorable serine ligation, but of sufficient flexibility to allow the displacements required by the shortening (from 2.3 to 1.8-2.0 Å) (*44, 45*) of the iron-ligand bond. The C11S variant (Cys 9 in Fd4) displayed features suggestive of serine ligation, but was too unstable to be satisfactorily characterized (*11*). The structure shows that this cysteine is the most buried inside the protein and hence most likely constrained. The instability of the C11S variant may reflect the increased constraints that make it difficult to accommodate the sulfur-to-oxygen substitution. In contrast, attempts to introduce serine ligation at the Cys24 (Cys22 in Fd4) site have remained unsuccessful (*11-13*), probably because the polypeptide chain in this region is too flexible to enforce serine coordination. As a result, serine is either replaced as a ligand by a nearby cysteine (in a similar fashion as in *A. vinelandii* ferredoxin I, where cysteine ligand swapping is also preferred over serine ligation (*46*)) or possibly by a solvent molecule if no cysteine is available.

Interaction with nitrogenase MoFe protein

CpFd has been shown to interact specifically with the nitrogenase MoFe-protein (*18*). The functional meaning of this interaction has not yet been elucidated, and since *A. aeolicus* is not a nitrogen fixer, Fd4 cannot be expected to undergo a similar interaction. The structural model discussed here may, nevertheless, be used for a better understanding

of the interaction between CpFd and the MoFe protein, and even more so since residues involved in this interaction have been identified on CpFd by site-directed mutagenesis, namely Glu 31, Glu 34 and Glu 38 (corresponding to Glu 29, Gln 32, and Glu 36, respectively, in Fd4). These residues are located on three successive turns on the external face of helix 26-41. The twofold related residues are positioned ~20-25 Å apart, along the base of the saddle (Fig. 2b) formed between the protruding loops of each subunit in the dimer. A complete mapping of the surface of the CpFd is currently being performed and will be presented later.

Similarity with the NuoE subunit of NADH-ubiquinone oxidoreductase

The thioredoxin-like protein fold exhibited by Fd4 occurs in numerous proteins and complexes, as shown by the growing body of sequence data. In addition to the homologs from *C. pasteurianum* and *A. vinelandii* (*22*), similar domains or subunits occur in hydrogenases (*23, 24*) and in complex I (NADH-ubiquinone oxidoreductase) of aerobic respiratory chains (*23, 47*). Because the sequences of the former are all similar to one another, the following discussion will be restricted to the NuoE subunit of complex I (*47*).

The most extensively characterized NuoE subunit is from *Paracoccus denitrificans* (*47*), whose sequence has been included in the alignment shown in Figure 3.1, together with the sequence of the putative NuoE from *A. aeolicus* (*19*). Both NuoE subunits contain 80- to 90-residue N-terminal extensions, and in the case of the *P. denitrificans* NuoE subunit, a C-terminal extension is also present. The most extensive

regions of sequence similarity occur in sequence blocks that encompass the cluster ligands and the core of the thioredoxin fold. The sequence alignment also highlights structural differences between the two types of proteins, particularly in the loop connecting the first two cysteine ligands, in the C-terminal regions, and in the long N-terminal extension in the NuoE subunits. These alterations may reflect the requirement for NuoE to be integrated into complex I in association with a number of other subunits.

The thioredoxin-like protein fold described here for the [2Fe-2S] Fd4 occurs in a fairly large group of proteins which may be split into two subfamilies. One of these includes short (~100 residues) dimeric proteins (from *C. pasteurianum*, *A. vinelandii*, *A. aeolicus*, and probably others, see ref. 20) that have a protruding surface loop near the cluster site. The second subfamily is characterized by longer polypeptide chains (160-240 residues) that are domains of larger proteins or subunits of complexes, in particular hydrogenases and complex I (*23, 24*). Aside from the NuoE subunit, other subunits of complex I have also been found to be homologous to proteins of known structure. For instance, the NuoI subunit of complex I is homologous to bacterial [4Fe-4S] ferredoxins (*48*), and the N-terminal domain of the NuoG subunit (*19*) (containing one [2Fe-2S] and either two or three [4Fe-4S] clusters) is homologous to the N-terminal domain of *C. pasteurianum* hydrogenase I (*49*). Models are thus available for the polypeptide folds around the two [2Fe-2S] and the four to five [4Fe-4S] clusters of complex I, which represent a sizable proportion of the eight to nine iron-sulfur clusters present in this complex (*23*). These structural models may be anticipated to be helpful in mapping the positions of subunits in forthcoming lower resolution structures of complex I (*50*), in a

similar fashion as recently reported for the ribosome structure (*51*). More generally, the increasing number of structurally characterized protein subunits or domains that are common to aerobic respiratory chains and to anaerobic enzymes confirms the existence of extensive bridges connecting aerobic and anaerobic electron transfer systems, and are expected to shed light on the evolution of energy transducing complexes.

Materials and Methods

Crystallization

The [2Fe-2S] ferredoxin from *A. aeolicus* (product of the *fdx4* gene) (*19*) was purified as described previously (*21*). Crystals of Fd4 for data collection were prepared under anaerobic conditions by the sitting drop vapor diffusion method, using 2 µl of ~11 mg/ml Fd4 in 10 mM Tris-HCl buffer at pH 8.0 and 0.2 M NaCl and 2 µl of reservoir solution equilibrated against a reservoir solution containing 1.6 M $(NH_4)_2SO_4$, 4-5% p-dioxane, and 100 mM 2-(*N*-morpholino)-ethane sulfonic acid at pH 6.5. Crystals belong to the space group $P4_12_12$ (a = b = 68.1 Å, and c = 102.5 Å), with two Fd4 subunits in the asymmetric unit.

Data collection, structure determination, and refinement

Crystals were cryoprotected by transferring them to a precipitant solution, whereupon glycerol was gradually added in 5% increments to the solution until the glycerol concentration reached 20% (v/v). Crystals were then flash-cooled at liquid nitrogen temperatures. Multi-wavelength anomalous dispersion (MAD) diffraction data

were collected under cryogenic temperatures at beamline 9-2 at the Stanford Synchrotron Radiation Laboratory on an Area Detector Systems Corp. (ADSC) Quantum-4 CCD detector controlled by the distributed control system software BLU-ICE (Beam Line Unification in an Integrated Control Environment) (Table 3.1). Based on the x-ray fluorescence spectrum measured from a single crystal, one data set to 2.3 Å resolution and three data sets to 2.6 Å resolution were collected at four different wavelengths near the iron *K* absorption edge (Table 3.1). Data sets were processed and scaled using DENZO and SCALEPACK (*52*). The positions of the two iron-sulfur clusters in the asymmetric unit were determined using SOLVE (*53*) and used to calculate phases to 3.0 Å resolution using MLPHARE (*54*) (Table 3.1). The initial electron density map allowed the tracing of residues 3-107 and the [2Fe-2S] cluster of subunit A and all 110 residues and the [2Fe-2S] cluster of subunit B. Geometrical parameters of the [2Fe-2S] cluster were restrained throughout refinement. Multiple rounds of refinement using CNS (*55*) and model rebuilding using O (*56*) resulted in a final R-factor and R-free (calculated from a test set with 7.5% of all reflections) of 22.5% and 27.0%, respectively. The final Fd4 model contains two subunits (residues 1-109 for subunit A and residues 1-110 for subunit B containing 1693 protein atoms), two [2Fe-2S] clusters, 70 water molecules, and two glycerol molecules, for a total of 1783 atoms. The model displays good stereochemistry, with r.m.s. deviations in bond lengths and angles of 0.013 Å and 1.53 degrees, respectively. The average temperature factors for the protein main-chain, side-chain, solvent, and glycerol atoms are 52.7, 54.1, 62.6, and 81.2 $Å^2$, respectively. In the Ramachandran plot, as calculated with PROCHECK (*57*), 87.4% of the residues are in

the most favored regions, 12.0% of the residues are in additionally allowed regions, and 0.6% of the residues are in generously allowed regions.

RCSB Protein Data Bank accession code

The atomic coordinates of Fd4 have been deposited with the Protein Data Bank at the Research Collaboratory for Structural Bioinformatics under the accession code 1F37.

Acknowledgments

We thank the National Institutes of Health (D.C.R.) for support. A.P.Y. was supported in part by a National Science Foundation Graduate Research Fellowship. This work is based upon research conducted at the Stanford Synchrotron Radiation Laboratory (SSRL), which is funded by the Department of Energy, Office of Basic Energy Sciences. The Biotechnology Program is supported by the National Institutes of Health, National Center for Research Resources, Biomedical Technology Program and the Department of Energy, and Office of Biological and Environmental Research.

References

1. Johnson, M. K. (1994) *Encyclopedia of Inorganic Chemistry*, Vol. 4, Wiley, Chichester, New York.
2. Beinert, H., Holm, R. H., and Münck, E. (1997) *Science 277*, 653-659.
3. Beinert, H. (1999) *J. Inorg. Biochem. 74*, 3.
4. Ikemizu, S., Bando, M., Sato, T., Morimoto, Y., Tsukihara, T., and Fukuyama, K. (1994) *Acta Crystallogr. D50*, 167-174.
5. Morales, R., Chron, M. H., Hudry-Clergeon, G., Petillot, Y., Norager, S., Medina, M., and Frey, M. (1999) *Biochemistry 38*, 15764-15773.
6. Müller, A., Müller, J. J., Muller, Y. A., Uhlmann, H., Bernhardt, R., and Heinemann, U. (1998) *Structure 6*, 269-280.
7. Iwata, S., Saynovits, M., Link, T. A., and Michel, H. (1996) *Structure 4*, 567-579.
8. Romão, M. J., Archer, M., Moura, I., Moura, J. J. G., LeGall, J., Engh, R., Schneider, M., Hof, P., and Huber, R. (1995) *Science 270*, 1170-1176.
9. Hardy, R. W. F., Knight Jr., E., McDonald, C. C., and D'Eustachio, A. J. (1965) *Non-heme iron proteins: role in energy conversion*, Antioch Press, Yellow Springs, Ohio.
10. Meyer, J., Bruschi, M. H., Bonicel, J. J., and Bovier-Lapierre, G. E. (1986) *Biochemistry 25*, 6054-6061.
11. Meyer, J., Fujinaga, J., Gaillard, J., and Lutz, M. (1994) *Biochemistry 33*, 13642-13650.

12. *Golinelli*, M.-P., Akin, L. A., Crouse, B. R., Johnson, M. K., and Meyer, J. (1996) *Biochemistry 35*, 8995-9002.

13. Golinelli, M.-P., Chatelet, C., Duin, E. C., Johnson, M. K., and Meyer, J. (1998) *Biochemistry 37*, 10429-10437.

14. Crouse, B. R., Meyer, J., and Johnson, M. K. (1995) *J. Am. Chem. Soc. 117*, 9612-9613.

15. Achim, C., Golinelli, M.-P., Bominaar, E. L., Meyer, J., and Münck, E. (1996) *J. Am. Chem. Soc. 118*, 8168-8169.

16. Fujinaga, J., Gaillard, J., and Meyer, J. (1993) *Biochem. Biophys. Res. Commun. 194*, 104-111.

17. Achim, C., Bominaar, E. L., Meyer, J., Peterson, J., and Münck, E. (1999) *J. Am. Chem. Soc. 121*, 3704-3714.

18. Golinelli, M.-P., Gagnon, J., and Meyer, J. (1997) *Biochemistry 36*, 11797-11803.

19. Deckert, G., Warren, P. V., Gaasterland, T., Young, W. G., Lenox, A. L., Graham, D. E., Overbeek, R., Snead, M. A., Keller, M., Aujay, M., Huber, R., Feldman, R. A., Short, J. M., Olsen, G. J., and Swanson, R. V. (1998) *Nature 392*, 353-358.

20. Jacobson, M. R., Brigle, K. E., Bennett, L. T., Setterquist, R. A., Wilson, M. S., Cash, V. L., Beynon, J., Newton, W. E., and Dean, D. R. (1989) *J. Bacteriol. 171*, 1017-1027.

21. Chatelet, C., Gaillard, J., Pétillot, Y., Louwagie, M., and Meyer, J. (1999) *Biochem. Biophys. Res. Commun. 261*, 885-889.

22. Chatelet, C., and Meyer, J. (1999) *J. Biol. Inorg. Chem. 4*, 311-317.

23. Albracht, S. P. J., Mariette, A., and deJong, P. (1997) *Biochim. Biophys. Acta 1318*, 92-106.

24. Verhagen, M. F. J. M., O'Rourke, T., and Adams, M. W. W. (1999) *Biochim. Biophys. Acta 1412*, 212-229.

25. Meyer, J., Moulis, J. M., and Lutz, M. (1984) *Biochem. Biophys. Res. Commun. 119*, 828-835.

26. Müller, J. J., Müller, A., Rottmann, M., Bernhardt, R., and Heinemann, U. (1999) *J. Mol. Biol. 294*, 501-513.

27. Holmgren, A., Söderberg, B.-O., Eklund, H., and Brändén, C.-I. (1975) *Proc. Natl. Acad. Sci. U. S. A. 72*, 2305-2309.

28. Katti, S. K., LeMaster, D. M., and Eklund, H. (1990) *J. Mol. Biol. 212*, 167-184.

29. Saarinen, M., Gleason, F. K., and Eklund, H. (1995) *Structure 3*, 1097-1108.

30. Martin, J. L. (1995) *Structure 3*, 245-250.

31. Nicholls, A., Sharp, K. A., and Honig, B. (1991) *Proteins 11*, 281-296.

32. Coldren, C. D., Hellinga, H. W., and Caradonna, J. P. (1997) *Proc. Natl. Acad. Sci. U. S. A. 94*, 6635-6640.

33. Pinto, A. L., Hellinga, H. W., and Caradonna, J. P. (1997) *Proc. Natl. Acad. Sci. U. S. A. 94*, 5562-5567.

34. Wisz, M. S., Garrett, C. Z., and Hellinga, H. W. (1998) *Biochemistry 37*, 8269-8277.

35. Hellinga, H. W. (1998) *J. Am. Chem. Soc. 120*, 10055-10066.

36. Benson, D. E., Wisz, M. S., Liu, W., and Hellinga, H. W. (1998) *Biochemistry 37*, 7070-7076.

37. Bushweller, J. H., Billeter, M., Holmgren, A., and Wüthrich, K. (1994) *J. Mol. Biol. 235*, 1585-1597.

38. Reinemer, P., Dirr, H. W., Ladenstein, R., Schaffer, J., Gallay, O., and Huber, R. (1991) *EMBO J. 10*, 1997-2005.

39. Ji, X., Zhang, P., Armstrong, R. N., and Gilliland, G. L. (1992) *Biochemistry 31*, 10169-10184.

40. Choi, H.-J., Kang, S. W., Yang, C. H., Rhee, S. G., and Ryu, S. E. (1998) *Nature Struct. Biol. 5*, 400-406.

41. Wang, S. R., Trumble, W. R., Liao, H., Wesson, C. R., Dunker, A. K., and Kang, C. H. (1998) *Nature Struct. Biol. 5*, 476-483.

42. Dai, S. D., Schwendtmayer, C., Schürmann, P., Ramaswamy, S., and Eklund, H. (2000) *Science 287*, 655-658.

43. Zhou, C. Y., and Holm, R. H. (1997) *Inorg. Chem. 36*, 4066-4077.

44. Hurley, J. K., Weber-Main, A. M., Hodges, A. E., Stankovich, M. T., Benning, M. M., Holden, H. M., Cheng, H., Xia, B., Markley, J. L., Genzor, C., Gomez-Moreno, C., Hafezi, R., and Tollin, G. (1997) *Biochemistry 36*, 15109-15117.

45. Xiao, Z. G., Lavery, M. J., Ayhan, M., Scrofani, S. D. B., Wilce, M. C. J., Guss, J. M., Tregloan, P. A., George, G. N., and Wedd, A. G. (1998) *J. Am. Chem. Soc. 120*, 4135-4150.

46. Shen, B., Jollie, D. R., Diller, T. C., Stout, C. D., Stephens, P. J., and Burgess, B. K. (1995) *Proc. Natl. Acad. Sci. U. S. A. 92*, 10064-10068.

47. Yano, T., Sled, V. D., Ohnishi, T., and Yagi, T. (1994) *FEBS Lett. 354*, 160-164.

48. Dauter, Z., Wilson, K. S., Sieker, L. C., Meyer, J., and Moulis, J. M. (1997) *Biochemistry 36*, 16065-16073.

49. Peters, J. W., Lanzilotta, W. N., Lemon, B. J., and Seefeldt, L. C. (1998) *Science 282*, 1853-1858.

50. Grigorieff, N. (1999) *Curr. Opin. Struct. Biol. 9*, 476-483.

51. Ban, N., Nissen, P., Hansen, J., Capel, M., Moore, P. B., and Steitz, T. A. (1999) *Nature 400*, 841-847.

52. Otwinowski, Z., and Minor, W. (1997) *Methods Enzymol. 276*, 307-326.

53. Terwilliger, T. C., and Berendzen, J. (1999) *Acta Crystallogr. D55*, 849-861.

54. Otwinowski, Z. (1991) in *CCP4 Study Weekend Isomorphous Replacement and Anomalous Scattering* (Wolf, W., Evans., P. R., and Leslie, A. G. W., Eds.) pp 80-86, SERC Daresbury Laboratory, Warrington, UK.

55. Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., J.-S., J., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) *Acta Crystallogr. D54*, 905-921.

56. Jones, T. A., Zou, J. Y., Cowan, S. W., and Kjeldgaard, M. (1991) *Acta Crystallogr. A47*, 110-119.

57. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) *J. Appl. Crystallogr. 26*, 283-291.

58. Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994) *Nucleic Acids Res. 22*, 4673-4680.

Table 3.1. Summary of data collection and phasing statistics.

Data set	remote	inflection	peak	reference
Wavelength (Å)	1.7968	1.7415	1.7340	1.5001
Resolution range (Å)	50-2.6	50-2.6	50-2.6	50-2.3
Total reflections	51,711	52,219	47,760	74,834
Unique reflections	7,839	7,795	7,843	11,218
Completeness (%) [a]	99.3 (99.7)	99.2 (100.0)	99.3 (100.0)	99.3 (99.8)
$I/\sigma(I)$	32.5 (7.0)	26.4 (7.2)	25.5 (7.2)	24.8 (5.3)
R_{sym} (%) [b]	4.6 (22.2)	5.7 (20.3)	5.4 (19.4)	5.5 (27.7)
Phasing power [c] (centric / acentric)	0.82 / 1.30	1.44 / 2.07	1.41 / 1.95	
R_{cullis} [d] (centric / acentric / anomalous)	0.76 / 0.76 / 0.94	0.57 / 0.59 / 0.55	0.58 / 0.62 / 0.58	-- / -- / 0.57
Overall FOM		0.703 (to 3.0 Å resolution)		

[a] Numbers in parentheses correspond to values in the highest resolution shell.

[b] $R_{sym} = (\Sigma_{hkl} \Sigma_i | I_i(hkl) - < I(hkl) > |) / (\Sigma_{hkl} \Sigma_i I(hkl))$.

[c] Phasing Power $= < F_H > / \Sigma || F_{PH} - F_P | - F_H |$, where $< F_H >$ is the r.m.s. calculated heavy-atom structure factor.

[d] $R_{cullis} = \Sigma || F_{PH} - F_P | - F_H | / \Sigma | F_{PH} - F_P |$, where F_H is the calculated heavy-atom structure-factor amplitude, and F_P and F_{PH} are the native and derivative structure-factor amplitudes, respectively.

Figure 3.1. Sequence alignment of Fd4 with [2Fe-2S] ferredoxins from *C. pasteurianum* (*10*) and *A. vinelandii* (*22*), and the NuoE subunits of complex I from *P. denitrificans* (*47*) and *A. aeolicus* (*19*). For simplicity, the sequences of only portions of the NuoE subunits from *P. denitrificans* (239 residues) and *A. aeolicus* (160 residues) that are homologous to the [2Fe-2S] ferredoxins are shown. Sequences were aligned using CLUSTALW (*58*). The secondary structure elements of Fd4, determined using PROCHECK (*57*), are indicated above the alignment. Cysteine ligands of the [2Fe-2S] clusters are highlighted in blue while other conserved residues between Fd4 and the other proteins are highlighted in orange.

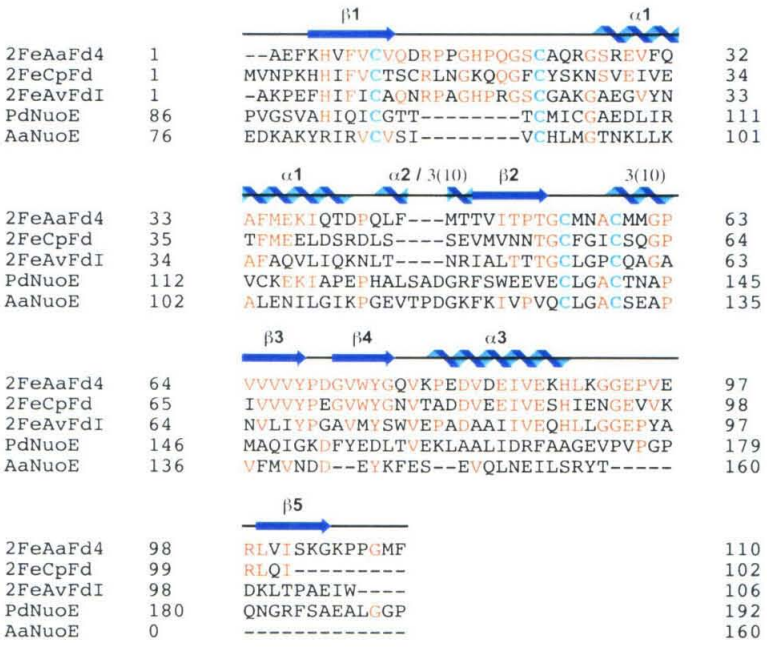


Figure 3.2. (a) Ribbons diagram representation of the homodimeric arrangement of Fd4. Subunits A (green) and B (purple) are shown with their respective [2Fe-2S] clusters represented by ball-and-stick models. Iron and sulfur atoms of the clusters are colored green and yellow, respectively. (b) View upon rotation of the Fd4 homodimer by 90^{O} about the vertical axis in the plane of the paper, highlighting the separation between the two [2Fe-2S] clusters and also illustrating the saddle formed by the two long surface loops. (c) Stereoview of a monomer of Fd4 in gradient coloration from the N- (blue) to the C-terminus (red). The [2Fe-2S] cysteine ligands are shown in magenta.

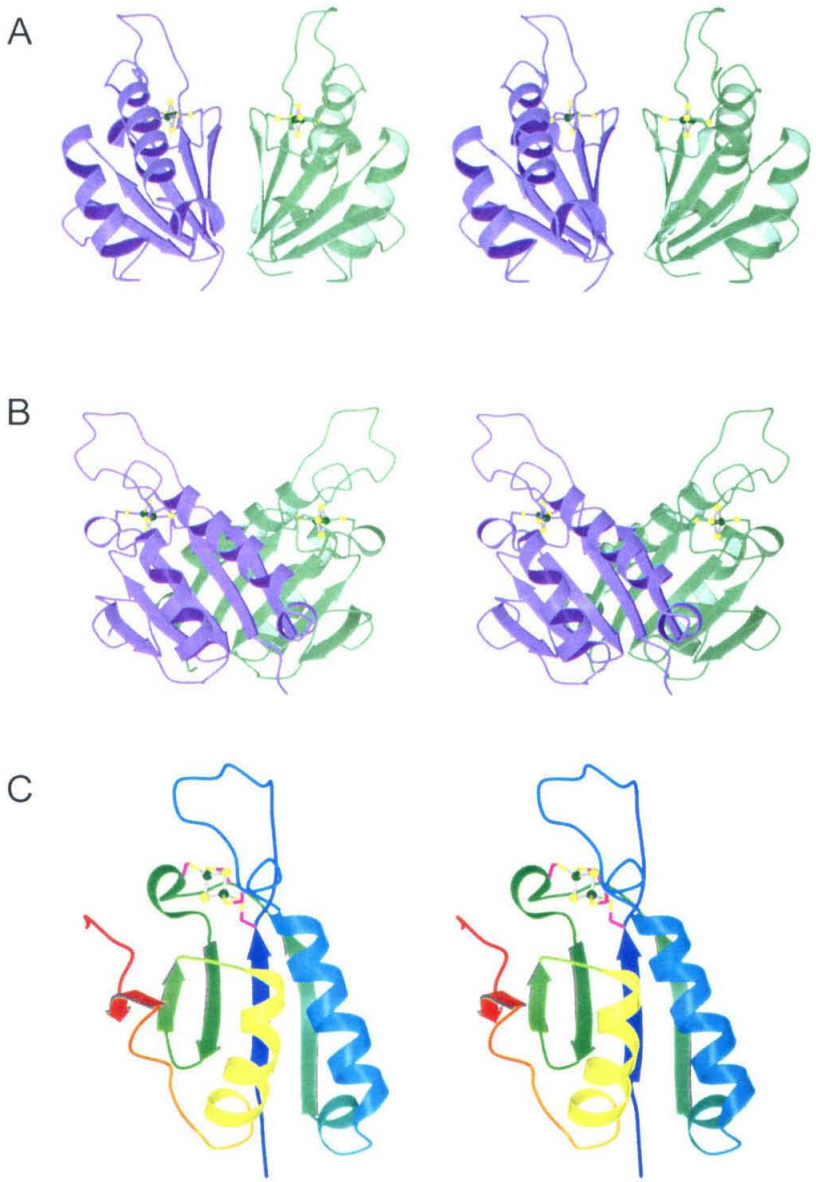


Figure 3.3. Stereoview of the superposition of *Anabaena* thioredoxin-2 (orange) (29) and a monomer of Fd4 (green), showing the similarity between the folds. In addition to the similarity of the overall folds, two of the [2Fe-2S] cysteine ligands, Cys 9 and Cys 22 (shown in a darker shade of green), in Fd4 are positioned near the two active site cysteines of thioredoxin-2, Cys 32 and Cys 35 (shown in a darker shade of orange).

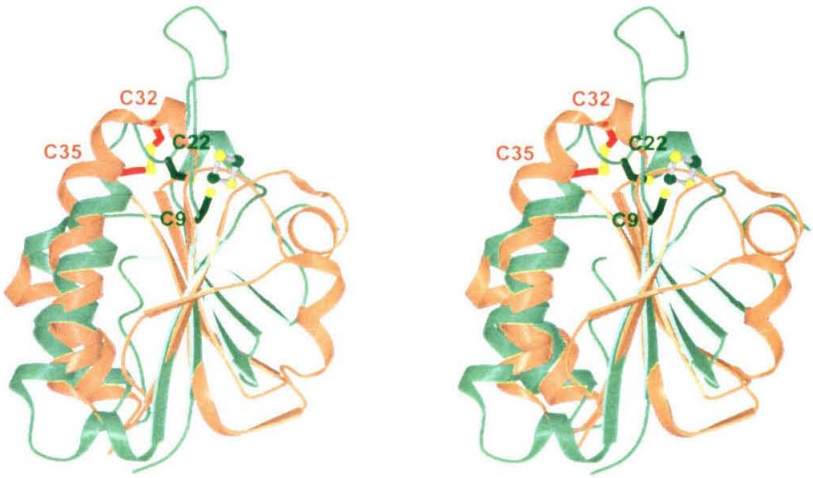
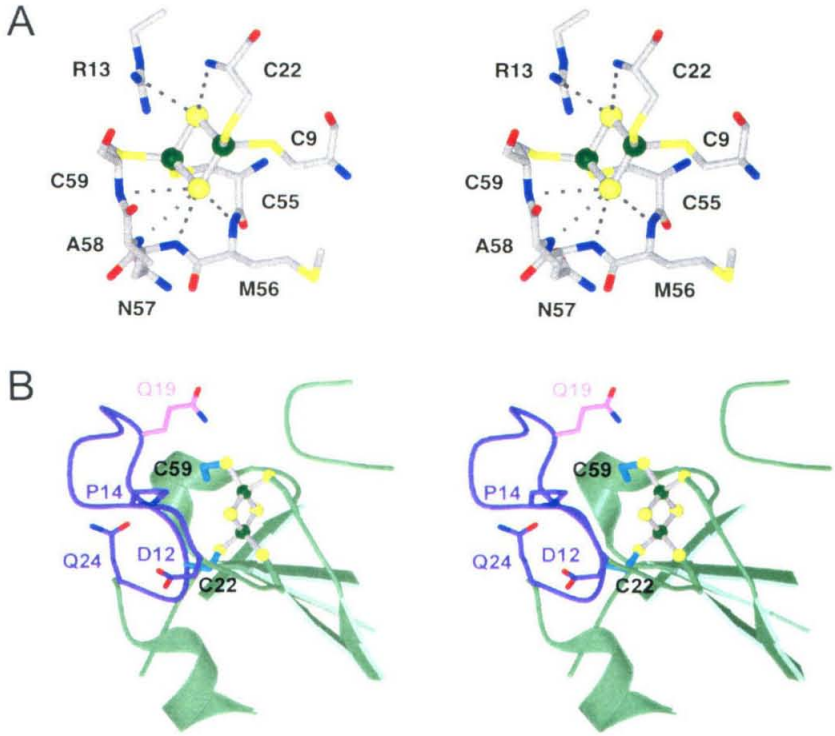


Figure 3.4. (a) Stereoview of the immediate environment around the [2Fe-2S] cluster. The color scheme for the [2Fe-2S] cluster is the same as in Fig. 2. Potential hydrogen bonds between the cluster sulfide S1 atom and the amide nitrogen of Cys 22 and the guanidinium group of Arg 13 and between the cluster sulfide S2 atom and the amide nitrogens of Met 56, Asn 57, Ala 58, and Cys 59 are shown as dotted lines. (b) Stereoview of the [2Fe-2S] cluster and nearby residues corresponding to ones in CpFd which, when replaced by cysteine, can serve as alternative ligands to the [2Fe-2S] cluster. Cys 22 and Cys 59, the two cysteine ligands which can be replaced (Cys 24 and Cys 60 in CpFd), are shown in cyan. Cysteines introduced into the polypeptide chain between residues 14 to 24 (represented in purple) can serve as a cluster ligand in place of Cys 24, while a cysteine substituted for Gln 19 (pink) can replace Cys 60 as a ligand.



CHAPTER 4

High-resolution Crystal Structures of the Wild Type and Cys55Ser and Cys59Ser Variants of the Thioredoxin-like [2Fe-2S] Ferredoxin from *Aquifex aeolicus*

High-resolution Crystal Structures of the Wild Type and Cys55Ser and Cys59Ser Variants of the Thioredoxin-like [2Fe-2S] Ferredoxin from *Aquifex aeolicus*

Andrew P. Yeh[1], Xavier Ambroggio[2], Susana Andrade[1,3], Oliver Einsle[1,3], Claire Chatelet[4], Jacques Meyer[4], and Douglas C. Rees[1,3†]

1 Division of Chemistry and Chemical Engineering, 147-75CH, California Institute of Technology, Pasadena, California 91125

2 Division of Biology, California Institute of Technology, Pasadena, California 91125

3 Howard Hughes Medical Institute, California Institute of Technology, Pasadena, California 91125

4 Département de Biologie Moléculaire et Structurale, CEA-Grenoble, F-38054, Grenoble, France

† Corresponding author

E-mail address of the corresponding author: dcrees@caltech.edu

Abstract

The [2Fe-2S] ferredoxin (Fd4) from *Aquifex aeolicus* adopts a thioredoxin-like polypeptide fold that is distinct from other [2Fe-2S] ferredoxins. Crystal structures of the Cys55Ser (C55S) and Cys59Ser (C59S) variants of this protein have been determined to 1.25 Å and 1.05 Å resolution, respectively, while the resolution of the wild type (WT) has been extended to 1.5 Å. The improved WT structure provides a detailed description of the [2Fe-2S] cluster, and in particular, two features that have not been previously noted in any [2Fe-2S] cluster containing protein, namely, a pronounced distortion in the cysteine coordination and a Cα-H—Sγ hydrogen bond between cluster ligands Cys55 and Cys9. These features may contribute to the unusual electronic and magnetic properties of the [2Fe-2S] clusters in WT and variants of this ferredoxin. The polypeptide chains surrounding the cysteine ligands are located at the base of a prominently protruding loop and are endowed with very different flexibilities. The structures of the two variants of Fd4, in which single cysteine ligands to the [2Fe-2S] cluster are replaced by serine, establish the metric details of serine-ligated Fe-S active sites with unprecedented accuracy. Both the cluster and its surrounding protein matrix change in subtle ways to accommodate this ligand substitution, particularly in terms of distortions of the Fe_2S_2 inorganic core from planarity and displacements of the polypeptide chain. These high-resolution structures illustrate how the interactions between polypeptide chains and Fe-S active sites reflect combinations of flexibility and rigidity on the part of both partners; these themes are also evident in more complex systems, as exemplified by changes associated with serine ligation of the nitrogenase P-cluster.

Introduction

The low potential iron-sulfur (Fe-S) electron carriers known as ferredoxins (Fds) are composed of three distinct classes, the [3Fe-4S]/[4Fe-4S] bacterial-type Fds, the plant- and mammalian-type [2Fe-2S] Fds, and the thioredoxin-like [2Fe-2S] Fds (*1*). The first two classes were discovered nearly forty years ago (*2, 3*) and have since been characterized in considerable detail, including by high-resolution X-ray crystallography (*4, 5*). The third class of Fd is more sparsely distributed and, therefore, has not been as thoroughly investigated (*1*). The best characterized members of that group are [2Fe-2S] Fds from the bacteria *Clostridium pasteurianum* (*6*), *Azotobacter vinelandii* (*7*), and *Aquifex aeolicus* (*8*). The high similarity between these proteins allows for easy transfer of structural information among them. For instance, many properties of molecular variants of *C. pasteurianum* [2Fe-2S] Fd (*6, 9, 10*) could be rationalized from the crystal structure of *A. aeolicus* Fd4 (*11*). The latter structure revealed the unexpected thioredoxin-like fold of these Fds and confirmed that they are distinct from the other two ferredoxin classes (*1*). While that structure uncovered a novel fold among Fe-S proteins, the resolution (2.3 Å) was not among the highest currently reported (ca. 1 Å) for metalloenzymes. Indeed, such high-resolution structures are of utmost interest as they bring forth precise geometries of metal sites (*4*) and may allow description of redox transitions (*5*). Additional efforts have therefore been made, both on wild type thioredoxin-like Fds and on several of the molecular variants that were produced over the years (*6, 10, 12*), with the aim of improving the crystallographic data.

In some modified forms of *C. pasteurianum* Fd, cysteine ligands of the Fe-S cluster were replaced by serine (*9, 13*). In their reduced $[2Fe\text{-}2S]^{+}$ level, these serine-ligated active sites were found to assume a delocalized mixed-valence state having a ground spin state of 9/2 (*14, 15*). This unprecedented occurrence in binuclear iron-sulfur clusters has stimulated experimental and theoretical work (*16, 17*). The development of these investigations has been hampered, however, by the absence of structural data on serine-ligated [2Fe-2S] active sites and their environment. Since only the Fd from *A. aeolicus* has been crystallized, we have repeated in that protein amino-acid substitutions previously performed on *C. pasteurianum* Fd, and produced the Cys55Ser and Cys59Ser variants that contain serine-ligated [2Fe-2S] clusters. We report here, for both of these variants as well as wild type protein, high-resolution structures that provide a wealth of information on the protein fold and metal site of thioredoxin-like Fds, as well as metric details of unprecedented accuracy for serine-ligated Fe-S clusters in proteins.

Results and Discussion

Wild type Fd4, overall structure

The WT and serine substituted forms of Fd4 analyzed in this study were crystallized in a monoclinic space group, *C2*, that is distinct from the original tetragonal form solved at 2.3 Å resolution (*11*). All three structures reported in this paper crystallize isomorphously in this new space group, although three different crystallization conditions were employed (see Materials and Methods), which are also distinct from the conditions used in the initial structure determination.

As observed originally, the current structure of WT Fd4 determined at 1.5 Å resolution exists as a homodimer, with each monomer adopting a thioredoxin-like fold (Figure 4.1). The two noncrystallographic symmetry related subunits are nearly identical, with an rmsd of 0.27 Å between 101 Cα atoms. As a consequence of the differences in crystal packing between the original and present WT structures, the two monomers in the Fd4 dimer undergo a slight rigid body shift relative to each other (Figure 4.1). With the A subunits of both WT forms superimposed, the shift in subunit B of the new WT form relative to that of the old form can be quantitatively characterized as a 3.8° rotation about an axis oriented ~74° from the twofold rotation axis that relates subunits in a dimer. The axis about which this 3.8° rotation occurs passes near residue Thr B53, which along with Pro B52, Gly B54, and the corresponding residues in subunit A, form a short stretch of antiparallel β-sheet that stabilizes the dimer interface. As a result of this change in dimer packing, the hydrogen bonding geometries of residues in this antiparallel arrangement of β-strands are slightly modified. Since cluster ligand Cys 55 is adjacent to this region, it is possible that these changes in subunit-subunit packing could provide a mechanism for coupling the oligomeric conformation of the protein to changes in cluster environment. The differences in crystal interactions are also reflected in changes in conformations in two flexible loop regions spanning residues 13-20 and residues 39-46 (Figure 4.1), which are also near cluster ligands (Cys 9 and 22). Finally, several residues at the polypeptide chain ends (residues B1-2 and 104-110 of both subunits), which were ordered in the lower resolution structure, are disordered in the high-resolution structure.

Wild type Fd4, iron-sulfur cluster

The [2Fe-2S] cluster, which resides near the surface of each monomer, is coordinated by four cysteines, with Cys9 and Cys22 ligating Fe1 and the more solvent-exposed Cys55 and Cys59 ligating Fe2. At the resolution of the current study, it was possible to conduct the refinement without restraining the cluster geometry, resulting in geometrical values for the cluster geometry that are both more accurate and also less biased than in the previous WT model (Table 4.3).

As an idealized reference point, a [2Fe-2S] cluster coordinated by four sulfhydryl groups may be regarded as a framework that consists of two edge-sharing tetrahedral Fe sites and that exhibits overall D_{2h} symmetry. The central part of this structure is a planar Fe_2S_2 inorganic core. The geometries of real clusters, observed in both model compounds and in proteins, generally reflect this expectation, although deviations from this idealization are evident (*18*). Table 4.4 lists the average stereochemical parameters (bond distances and angles) observed in [2Fe-2S] clusters of proteins refined at high-resolution. A notable deviation from the idealized symmetry is the nonplanarity of the Fe_2S_2 core in these proteins structures, which can be characterized by the average absolute value of ~175° for the Fe-S-S-Fe torsion angle, where 180° would correspond to exact planarity. For further comparison, the Fe_2S_2 unit present in [4Fe-4S] clusters is significantly more distorted than in [2Fe-2S] clusters, with an average value for this torsion angle of only 162.0° (*19*). The pronounced puckering of this unit in [4Fe-4S] clusters primarily reflects a more compressed Fe-S-Fe angle (71.7°), although the S-Fe-S angle is also slightly enlarged to 105.6° in [4Fe-4S] clusters relative to [2Fe-2S] clusters.

Bond distances and angles involving the [2Fe-2S] core of WT Fd4 are generally in good agreement with those observed in well-refined structures (Tables 4.1 and 4.2), with the exception of two pronounced outliers: (i) the relatively small Cys55 Sγ–Fe2–Cys59 Sγ bond angle and (ii) the long Cys55 Sγ–Fe2 bond. These peculiarities of the [2Fe-2S] active site of *A. aeolicus* Fd4 may contribute to the spectroscopic idiosyncrasies of this class of [2Fe-2S] proteins (*1, 9*). The Cys55 Sγ–Fe2–Cys59 Sγ angle averages 90.5° in the two crystallographically independent subunits of WT Fd4, which is significantly smaller than the average Cys9 Sγ–Fe2–Cys22 Sγ angle of 104.6° observed in the same structure and the consensus value of 105.1° observed in other well-refined [2Fe-2S] protein structures (Table 4.4). These latter values are close to the value of 109.5° expected for ideal tetrahedral coordination. The more compressed Cys55 Sγ–Fe2–Cys59 Sγ angle has not previously been observed in other well-refined ferredoxins with Cys ligands, but is similar to the His Nδ–Fe–His Nδ bond angles of ~94° observed in Rieske type [2Fe-2S] clusters (*20, 21*).

The other salient difference in cluster geometry involves the Cys55 Sγ–Fe2 bond, which is longer on average by ~0.06 Å than the other three Fe-Sγ bonds (Table 4.3). The residues around Cys55 are well defined and appear to be relatively rigid, as reflected by the lower average temperature factors in this region (14.9 $Å^2$ for residues 53-57 vs. 21.4 $Å^2$ for all protein atoms). This, in turn, may reflect the varying degree to which the cluster ligands are hydrogen bonded (Fig. 2). As indicated in Table 4.5, the Sγ atom of Cys55 forms hydrogen bonds with both the side chain of Arg13 and the backbone amide

nitrogen of Asn57. In contrast, the Sγ atoms of the other three cysteine ligands form longer hydrogen bonds with either only one residue and/or solvent molecules. The greater structural rigidity surrounding Cys55 may also be reflected in the distortion of the peptide bond torsion angles (ω) of the residues in this region, particularly Thr53, Met56, and Ala58, whose ω torsion angles are on average 166.0°, 171.8°, and 189.1°, respectively (2.4, 1.4, and 1.6 standard deviations from the ideal value of 180°). With an average Fe1-S1-S2-Fe2 torsion angle of –173°, the extent of nonplanarity of the [2Fe-2S] cluster in WT Fd4 is comparable to those observed in other well-refined [2Fe-2S] protein structures.

An unusual interaction between the cluster ligands in Fd4 is the probable formation of a Cα-H—Sγ hydrogen bond between the Cys55 Cα-H and Cys9 Sγ (Figure 4.3). Such an interaction has not been previously described to our knowledge in [2Fe-2S] cluster containing proteins. This interaction is identified on the basis of the Cys55 Cα - Cys 9 Sγ and the Cys55 Cα-H—Cys 9 Sγ distances of 3.6 Å and 2.7 Å, respectively, with a Cα-H—Sγ angle of 149°. For the purposes of this calculation, hydrogen positions were generated with the CCP4 program HGEN (*22*). Identification of this interaction as a hydrogen bond is consistent with the criteria used to identify potential Cα-H—O hydrogen bonds (a Cα-H—O distance of 2.7 Å (*23*) which does not take into account the increased van der Waals radius of S relative to O). For this interaction to occur, the Cys55 Cα must be positioned over the [2Fe-2S] cluster, with the hydrogen directed towards the Sγ ligand (Cys 9) of the adjacent iron. Examination of the [2Fe-2S] cluster containing proteins used in the analysis for Table 4.6, as well as of [4Fe-4S] cluster

containing proteins described in (*19*), suggests that there are two related side chain conformations that can potentially achieve this interaction. These may be defined in terms of three torsion angles: χ_1, the N-Cα-Cβ-Sγ angle, χ_2, which describes the Cα-Cβ-Sγ-Fe angle, and χ_3, which describes the Cβ-Sγ-Fe-S angle. Since there are either two or three cluster sulfides that can be used to define this latter angle, for [2Fe-2S] and [4Fe-4S] clusters, respectively, the convention that will be used is to adopt the angle whose absolute value is closest to 0°. With these definitions, the side chain conformations that place the ligand Cα over the ring are (χ_1, χ_2, χ_3) ~(180°, 60°, 30°) and (-60°, -60°, -30°); the positions of the Cα, Cβ and Sγ atoms for these two conformations are related by a mirror plane that passes through the two Fe sites, perpendicular to the Fe_2S_2 plane. The first solution corresponds to that observed for Cys55 of Fd4, while the latter is observed in [4Fe-4S] cluster containing proteins such as HiPIPs (Figure 4.3b). In the latter case, however, the Cα-H—Sγ bond cannot be formed because the significant pucker in the Fe_2S_2 moiety of [4Fe-4S] clusters leads to an increase of the Cα–Sγ distance to ~5.3 Å, a separation which is too great for this interaction to occur.

With the exception of the Cα-H—Sγ interaction just described, the hydrogen-bonding network in the [2Fe-2S] cluster environment of Fd4 (Table 4.6) is altogether not unlike that occurring in plant-type Fds (*5, 24*), which assume otherwise totally different protein folds. It is interesting to note that in both cases the Sγ atoms of the cysteine ligands of the more reducible Fe are collectively involved in a larger number of hydrogen bonds than the cysteine ligands of the less reducible Fe. The ligands of the more reducible Fe are those closest to the N-terminus in plant-type Fds (Cys41 and Cys46; (*5*))

while they are those closest to the C-terminus in *A. aeolicus* Fd4 (C55 and C59; (*9, 14*)). In both cases, however, the more reducible Fe is the one nearer to the solvent. Also, as noted in the *Anabaena* [2Fe-2S] ferredoxin (*5*), the Fe-S bond lengths involving the more solvent exposed Fe tend to be slightly longer than those to the buried site (Table 4.4). In addition, the hydrogen bonding network in Fd4 does not appear to remain static. Arg13 in one of the subunits in the current WT Fd4 structure, for example, has been modeled as adopting two different conformations: one in which its guanidino moiety can form a hydrogen to Cys59 Sγ and another in which the group is far enough removed such that no hydrogen bond to Cys59 Sγ occurs. Moreover, Arg13 in the former conformation appears to shield the [2Fe-2S] cluster from the solvent. Indeed, in the C55S and C59S structures, Arg13 has been modeled in the second conformation, and a water molecule appears to hydrogen bond to Cys59 Sγ instead of the guanidino nitrogen of Arg13. It is conceivable, therefore, that the apparent flexibility of the Arg13 side chain may play a role in the physiological function of Fd4.

Structural effects of the cysteine to serine substitutions

Previous spectroscopic characterizations of the C56S and C60S variants of the homologous [2Fe-2S] ferredoxin from *Clostridium pasteurianum* revealed that substitution of the cysteine with a serine at either of these positions resulted in stable protein with serine-coordinated [2Fe-2S] clusters (*9, 13*). The high-resolution crystal structures of the corresponding C55S (1.25 Å) and C59S (1.05 Å) mutated forms in *Aquifex aeolicus* confirms this serine coordination, and furthermore, allows the

consequences of this serine substitution on the [2Fe-2S] cluster geometry to be accurately assessed.

The overall structures of both C55S and C59S Fd4 are nearly identical to that of the WT structure of Fd4 at 1.5 Å resolution (Figure 4.1). The rmsd's between the corresponding subunits A and B of WT and C55S are 0.17 Å and 0.16 Å and of WT and C59S are 0.18 Å and 0.23 Å. While there are no gross perturbations between the mutant and WT structures, there are subtle differences in the cluster geometry and its local environment resulting from the cysteine to serine substitutions. For all three structures, the distances and angles pertaining to the [2Fe-2S] cluster and hydrogen bonds involving the cluster are provided in Tables 4.1 and 4.3.

As expected, the overall geometry of the cluster in both mutated forms is very similar to that of WT Fd4, with the most pronounced differences being the shorter length of the Fe2-Oγ bond in both variants compared to that of the Fe-Sγ bond in WT (Figure 4.4a,b). The average Fe2-Oγ distance is 1.99 Å in C55S and 1.94 Å in C59S, whereas the average length of the Fe-Sγ bonds in WT Fd4 is 2.31 Å. The Fe-O bond lengths in these structures are comparable to those seen in the structures of an *Anabaena* [2Fe-2S] ferredoxin in which one of the cysteine ligands, Cys49, was substituted with a serine (*25*) (PDB ID 1QOA; avg. Fe-O bond length of 1.91 Å) as well as of *C. pasteurianum* rubredoxin in which one of the iron ligands, Cys42, was substituted with a serine (*26*) (PDB ID 1BE7; Fe-O bond length of 1.94 Å (*27*)). These bonds, however, are shorter than the value of 2.16 Å reported for the Fe-O bond observed in the structure of a

Cys77Ser ligand exchanged form of the *C. vinosum* [4Fe-4S] high potential iron protein (HiPIP) (*28*). The relatively longer Fe-O bond length observed in that structure was speculated to result from the structural rigidity of the polypeptide backbone. Aside from the shorter Fe-O bond length, also associated with the substitution of Ser Oγ for Cys Sγ is an increase of ~6-9° in the Oγ-Fe-Sγ angle relative to the compressed values observed for the corresponding Cys55 Sγ-Fe-Cys59 Sγ angle in WT (Table 4.4). In the Cys49Ser variant of the *Anabaena* ferredoxin (*25*), an increase in the Oγ-Fe-Sγ angle of ~+7° on average was observed relative to the native Sγ-Fe-Sγ value (*5*), while changes from ~–4° to +5° in the three Oγ-Fe-Sγ angles relative to the corresponding Sγ-Fe-Sγ angles were observed (*27*) in the Cys42Ser variant of *C. pasteurianum* rubredoxin (*26*).

Accommodation of the cysteine to serine substitution in both C55S and C59S is facilitated by structural changes in the protein backbone/side chain, as well as the inorganic iron-sulfur core itself. The main displacements are those of the Oγ and Fe2 atoms towards each other, as well as a somewhat lesser movement of the Sγ atom of the other Fe2 ligand, which follows the movement of the Fe. The details of these structural changes in each case (C55S and C59S) differ, however, most likely as a result of the different flexibility of the polypeptide chain near each of these cysteine residues. In the case of C55S, adaptation to the substitution occurs primarily through movement of the Ser55 Oγ atom towards the iron-sulfur core (Figure 4.5a,b). As seen in Table 4.6, which lists the positional shifts of cluster and ligand atoms between the structures, the Ser55 Oγ atom shifts by ~0.64 Å relative to the position of the corresponding Cys55 Sγ atom in the

WT structure. In contrast, there is little movement ($\leq$ ~0.12 Å) of the Cβ, Cα, and other backbone atoms of Ser55 relative to those of the native Cys55. There is also a ~0.2 Å shift of Fe2 towards the Ser 55 Oγ atom to which it is coordinated. As a result, the distortion from planarity of the [2Fe-2S] core is even more pronounced than in WT (Figure 4.5), as reflected by the average Fe1-S1-S2-Fe2 torsion angle of -171°, compared to the WT value of -173°. Other changes in the cluster geometry include a decrease of ~0.04 Å in the Fe1—Fe2 distance that is associated with changes of ~+1° and –2° in the average S-Fe-S and Fe-S-Fe angles, respectively (Tables 4.1 and 4.2). The third structural change that arises from the Cys55Ser mutation involves Cys 59, the other Fe2 ligand. As a result of the slight movement of Fe2 towards Ser55, Cys59 is pulled towards Fe2 in order to maintain an Fe2-Sγ bond length of ~2.3 Å. It is interesting to note, however, that unlike Ser55, where only the side chain atoms are primarily shifted, movement of Cys59 involves both its side chain as well as its backbone atoms (Figure 4.5a,b, Table 4.6). Relative to the corresponding atoms in the WT structure, the Cα and Sγ atoms of Cys59 in the C55S form have shifted 0.23 Å and 0.25 Å, respectively, towards Fe2. The greater degree to which Cys59 is structurally perturbed compared to Ser55, particularly in terms of movement of the backbone atoms, highlights the apparently greater structural rigidity in the region surrounding residue 55.

The main structural perturbations resulting from the C59S mutation, as with C55S, also involve residues 55, 59, and Fe2 of the inorganic core. The shorter Fe-O bond in the C59S structure is accommodated by the movement of Ser59 and Fe2 towards each other (Figure 4.4a,b). Relative to the corresponding atoms in the WT structure (i.e., Cys59 Cα

and Sγ), the Cα and Oγ atoms of Ser59 have shifted 0.42 Å and 0.70 Å towards Fe2, while Fe2 has shifted by ~0.17 Å towards Ser59 (Table 4.6). As a result, the iron-sulfur core becomes more planar (torsion angle of -177°), as opposed to the increase in distortion observed in C55S (Table 4.3, Figure 4.4b). Aside from a shift in the Fe2 position, the only other discernable difference in the inorganic core is a decrease in the Fe1—Fe2 distance by ~0.04 Å, similar to what was observed in the C55S structure. And as in C55S, this compression is associated with changes of ~+1° and -1° in the values of the S-Fe-S and Fe-S-Fe angles, respectively. Reflecting the same trends as seen for the Cys55 Sγ-Fe and Cys59 Sγ-Fe bonds in the WT protein, the Fe-O bonds in C59S are shorter than those in C55S by an average of ~0.05 Å (Table 4.3). Again, a likely explanation for this observation is the greater flexibility in the region surrounding position 59 compared to that surrounding position 55, as discussed previously.

Conclusions

The crystal structure of the WT form of Fd4 at the higher resolution of 1.5 Å reveals metric details of the [2Fe-2S] cluster that could not be confidently assigned in the initial study at 2.3 Å resolution. Some of these features are shared with other structurally characterized [2Fe-2S] proteins (*5, 24, 29*), even though the polypeptide folds are distinct. For instance, in all known high-resolution structures, the Fe_2S_2 inorganic core of the active site is distorted from planarity by ~5°–10°. It is also worth noting that the cysteinyl sulfur ligands of one of the iron atoms (Fe2 in the case of *A. aeolicus* Fd4) are involved in a larger number of hydrogen bonds than those of the otheriron atom. In all

cases, and expectably so, the former iron atom is the more reducible one (*9, 14, 25*). It is also the one that is closer to the protein surface.

The [2Fe-2S] active site of *A. aeolicus* Fd4 is remarkable for features that have not been previously observed in any [2Fe-2S] protein: the Cys55 Sγ–Fe2 bond is unusually long, the Cys55 Sγ–Fe2–Sγ Cys59 angle is exceptionally small and constrained, and the Cys55 Cα-H and Cys9 Sγ atoms are sufficiently close to indicate a hydrogen bonding interaction between these ligands. These unique features are at least in part consequences of the rigid protein environment around residue 55, and are most probably the root of the spectroscopic idiosyncrasies of thioredoxin-like [2Fe-2S] Fds, particularly in the ways in which they differ from the plant-type [2Fe-2S] Fds (*1*).

The structures of the Cys55Ser and Cys59Ser variants have been obtained at even higher resolution (1.25 Å and 1.05 Å, respectively). These are to our knowledge the first atomic-resolution structures of serine-ligated iron-sulfur clusters in proteins. Furthermore, the resolution of the structure of the Cys59Ser variant is the highest to date for any [2Fe-2S] protein. These structures confirm in each case the presence of a Ser Oγ–Fe2 bond that had been previously inferred from spectroscopic data (*9, 13*). More importantly, they reveal in considerable detail the conformational changes, in both the inorganic core and the polypeptide chain, that take place in order to accommodate the shortening of the Fe–Sγ bond upon replacement of sulfur by oxygen. The main structural perturbations observed in each case involve positional shifts of both Fe2 ligands as well as adjustments to the nonplanarity of the iron-sulfur core. Interestingly, the structural

rearrangements in the C55S and C59S variants differ in significant ways that are in keeping with the higher rigidity of the polypeptide chain around Cys55.

The types of structural accommodations associated with changes observed for cluster ligands in Fd4 are also evident in more complex systems, such as nitrogenase. The P-cluster of nitrogenase is a [8Fe-7S] metallocenter that exhibits structurally distinct oxidation states (*30, 31*). In the dithionite reduced form assigned to the P^N oxidation state, the P-cluster may be considered as two [4Fe-4S] clusters that share a common, hexacoordinate, sulfur. This assembly is coordinated to the nitrogenase MoFe-protein through six cysteine ligands, four of which coordinate a single iron, while the remaining two cysteines bridge two irons. In an oxidized form identified as the P^{OX} state, two of the irons move away from the central hexacoordinate sulfur, and these interactions are replaced with protein ligands, an amide nitrogen of one of the cluster cysteines and the side chain hydroxyl of Serβ188. The structural rearrangements associated with the switch between these two forms of the P-cluster are primarily restricted to an increased planarity of the Fe_2S_2 core (Figure 4.6) as the relevant Fe change positions; these correspond to an increase in the magnitude of the Fe-S-S-Fe torsion angles from ~145° in P^N (near that of [4Fe-4S] clusters) to ~175° in P^{OX} (near that of [2Fe-2S] clusters). As these transitions are associated with little change in positions of the coordinating residues, the P^N to P^{OX} conversion more closely resembles the consequences for the [2Fe-2S] cluster geometry of serine ligation at residue Cys55.

The high-resolution structures of the C55S and C59S variants of *A. aeolicus* Fd4 may also be expected to shed light on a unique and puzzling property of the counterpart C56S and C60S variants of the homologous protein from *C. pasteurianum*. In the one-electron reduced $[2Fe\text{-}2S]^{+}$ level, these mutated proteins, but not the WT, assume a delocalized mixed-valence level resulting in a ground spin state S = 9/2, whereas in all other known cases, $[2Fe\text{-}2S]^{+}$ clusters display localized mixed-valence states with an S = 1/2 ground spin state (*14, 15*). Although the structures reported here are those of the $[2Fe\text{-}2S]^{2+}$ redox level, they may nevertheless be used, with due caution, in the present discussion. Indeed, high-resolution structures of both redox levels of a plant-type Fd have shown that no major structural reorganization of the [2Fe-2S] cluster occurs upon reduction (*5*). The structural features favoring the appearance of the delocalized mixed-valence pair may therefore be present, at least incipiently, in the $[2Fe\text{-}2S]^{2+}$ structures reported here. The distortion of the Fe_2S_2 inorganic core from planarity is unlikely to play a role because it is larger in C55S, and smaller in C59S, as compared to WT (Figure 4.4). In contrast, the shortening of the Fe – Fe distance, albeit small (ca. 0.04 Å), is a unique feature of these serine-ligated [2Fe-2S] clusters. It should be noted that this shortening of the Fe–Fe distance is not a universal consequence of the substitution of Ser for Cys, however, as indicated by the slight increase in this distance in the Ser49 variant of the *Anabaena* ferredoxin (*25*), and the absence of any significant change in the series of model compounds prepared by Coucouvanis and co-workers (*32*). It is feasible that this slight decrease in the Fe–Fe distance may favor the occurrence of the delocalized mixed-valence state, especially since it is consistent with the prediction that transition from the localized- to the delocalized-valence state is determined by subtle structural

modifications (*15*). Another potentially relevant feature occurs in both the WT and serine-ligated structures, namely the unique distortion of the Cys55 Sγ–Fe2–Sγ Cys59 moiety. While this strain is likely to increase the difference between the Fe1 and Fe2 sites in the WT protein, the Sγ/Oγ substitution may perhaps rebalance the electronic properties of Fe1 and Fe2, and thus favor the setup of double exchange and valence delocalization. These questions clearly beg for structural data on the reduced levels of both the WT and serine-ligated forms of *A. aeolicus* Fd4.

More generally, the crystal structure of *A. aeolicus* Fd4 at the high-resolution of 1.5 Å confirms the previously described, at the lower resolution of 2.3 Å (*11*), novel fold of this class of [2Fe-2S] Fds (*1*). The subunits of the homodimeric molecule are in strong interaction, and each assumes a thioredoxin-like fold. They are also remarkable by the presence of a protruding loop in the vicinity of the [2Fe-2S] cluster (Figure 4.1). Sequence similarities indicate that several large redox enzymes, in particular hydrogenases and complex I of respiratory chains (*1*), contain subunits or domains that are predicted to assume structures similar to that of *A. aeolicus* Fd4. These subunits or domains presumably function as electron transfer agents and differ in at least two ways from *A. aeolicus* Fd4: (i) a single Fd-like module appears to be present, unlike the dimeric structure of Fd4, and (ii) the protruding loop is absent. The latter observation suggests that this loop in *A. aeolicus* Fd4 (and homologues in other bacteria, e.g., *C. pasteurianum* and *A. vinelandii*) may serve a function possibly more sophisticated than just electron transfer. In that respect, it should be emphasized that very rigid (near Cys55) and more flexible (near Cys59 or Cys22) regions of the polypeptide chain, as well as a

structurally constrained [2Fe-2S] metal site, are combined at the base of the protruding loop. This enhances the likelihood of tight interactions between the conformation of the polypeptide chain and the electronic structure (e.g., redox level) of the metal site. The role of these structural idiosyncrasies in the yet mysterious function of the thioredoxin-like [2Fe-2S] Fds will be the aim of future research.

Materials and Methods

Protein samples

Fd4 from *Aquifex aeolicus* was purified as described (*8*). The C55S and C59S variants were prepared by site-directed mutagenesis as described for the C56S and C60S counterparts from *C. pasteurianum* (*13*). The mutagenic oligonucleotides were 5'cacgcgttcatgGaaccggtgggag3' (hybridizing to the coding strand, mutated base in underlined upper case) and 5'ggttgcatgaacgcgtCtatgatgggaccg3' (hybridizing to the non-coding strand), for C55S and C59S, respectively. The mutated genes were overexpressed in *E. coli* and the C55S and C59S proteins were purified as described for the WT (*8*).

Crystallization

Crystals of oxidized WT, C55S, and C59S Fd4 were prepared by the sitting drop vapor diffusion method. Although anaerobic conditions were employed to minimize degradation of the cluster by exposure to atmospheric oxygen, no reductants were present during the crystallizations, so that the proteins should remain in the oxidized state. In the case of C55S, crystals were obtained by equilibrating 2 µl of reservoir solution and 2 µl of ~83 mg/ml Fd4 C55S (in 20 mM Tris-HCl buffer at pH 8.0 and 0.2 M NaCl) against a

reservoir solution containing 30% (w/v) PEG 4000, 0.2 M ammonium acetate, and 0.1 M sodium acetate at pH 4.6. Crystals of Fd4 C59S were obtained by equilibrating 2 μl of reservoir solution and 2 μl of ~67 mg/ml Fd4 C59S (in 20 mM Tris-HCl buffer at pH 8.0 and 0.2 M NaCl) against a reservoir solution containing 1.0 M 1,6 hexanediol, 0.01 M cobalt chloride, and 0.1 M sodium acetate at pH 4.6. WT Fd4 was crystallized by equilibrating 2 μl of ~10 mg/ml Fd4 (in 10 mM Tris-HCl buffer at pH 8.0 and 0.2 M NaCl) and 2 μl of reservoir solution against a reservoir solution containing 0.01 M zinc sulfate heptahydrate, 0.1 M MES buffer at pH 6.5, and 25% polyethylene glycol monomethyl ether 550. Despite the different crystallization conditions, in all three cases, crystals were of space group *C2* (C55S: a = 67.3 Å, b = 59.8 Å, c = 46.9 Å, $\beta = 109.8^{o}$; C59S: a = 67.3 Å, b = 59.8 Å, c = 46.8 Å, $\beta = 109.3^{o}$; WT: a = 67.2 Å, b = 59.8 Å, c = 47.2 Å, $\beta = 110.3^{o}$), with one dimeric Fd4 molecule per asymmetric unit (Table 4.1).

Data collection, structure determination, and refinement

Diffraction data to 1.25 Å resolution for C55S and 1.05 Å resolution for C59S were collected under cryogenic conditions on beamline 9-2 at the Stanford Synchrotron Radiation Laboratory (SSRL) (C55S: $\lambda = 0.8860$ Å; C59S: $\lambda = 0.9583$ Å) on an Area Detector Systems Corp. (ADSC) Quantum-4 CCD detector controlled by the distributed control system software BLU-ICE. C55S and C59S data sets were processed and scaled using DENZO and SCALEPACK (*33*). Diffraction data to 1.5 Å resolution for WT Fd4 was collected under cryogenic conditions at SSRL beamline 9-2 on an ADSC Quantum 315 CCD detector and was processed and scaled using MOSFLM and SCALA (*22*). A summary of the data collection statistics is listed in Table 4.1.

The Fd4 C55S structure was solved by molecular replacement using EPMR (*34*) with the original wild type Fd4 structure determined at 2.3 Å resolution (*11*) as the search model. Multiple rounds of positional refinement and individual isotropic B-factor refinement with the program CNS (*35*) were alternated with model rebuilding in the molecular graphics program O (*36*) against $2|F_o|-|F_c|$ σ_a-weighted and $|F_o|-|F_c|$ σ_a-weighted maps (*37*). The [2Fe-2S] cluster geometry was not restrained during refinement. Upon solvent addition and completion of refinement with CNS, positional and anisotropic B-factor refinements of the model were done using the programs SHELX97 (*38*) and REFMAC5 (*39*), which resulted in a final R-factor and R-free of 14.4% and 19.6%, respectively. A final round of refinement in SHELX97 yielded the standard uncertainties in atomic coordinates and bond lengths and angles. The final Fd4 C55S model comprises two subunits (2 x 101 residues), two [2Fe-2S] clusters, and 215 water molecules. Due to the absence of electron density for the first two residues at the N-terminus and the last seven residues at the C-terminus, these residues were not modeled.

The Fd4 C55S model without water molecules was used as the starting model for the Fd4 C59S model. The Fd4 C59S model was refined using a protocol similar to that outlined above for Fd4 C55S to an R-factor and R-free of 13.8% and 16.2%, respectively. As with Fd4 C55S, electron density for the first two residues at the N-terminus and the last seven residues at the C-terminus was absent. The final Fd4 C59S model consists of two subunits (2 x 101 residues), two [2Fe-2S] clusters, and 198 water molecules.

The Fd4 C55S model without water molecules was also used as the starting model for the high-resolution WT structure. The model was refined using CNS (*35*) with a protocol similar to that described above. The cluster geometry was not restrained during refinement. During the solvent addition process, in addition to water molecules, four zinc ions and one sulfate anion were modeled based on difference electron density peaks that were significantly higher than those corresponding to the water molecules. Unlike for the models of the mutated forms, the temperature factors of the WT model were not refined anisotropically due to the lower resolution of the data. The final model comprises two subunits (103 residues in subunit A, 101 residues in subunit B), two [2Fe-2S] clusters, 187 water molecules, four zinc ions, and one sulfate anion. Final refinement statistics for all three models are listed in Table 4.2.

References

1. Meyer, J. (2001) *FEBS Lett. 509*, 1-5.
2. Mortenson, L. E., Valentine, R. C., and Carnahan, J. E. (1962) *Biochem. Biophys. Res. Commun. 7*, 448-452.
3. Tagawa, K., and Arnon, D. I. (1962) *Nature 195*, 537-543.
4. Dauter, Z., Wilson, K. S., Sieker, L. C., Meyer, J., and Moulis, J. M. (1997) *Biochemistry 36*, 16065-16073.
5. Morales, R., Charon, M. H., Hudry-Clergeon, G., Pétillot, Y., Nørager, S., Medina, M., and Frey, M. (1999) *Biochemistry 38*, 15764-15773.
6. Golinelli, M. P., Chatelet, C., Duin, E. C., Johnson, M. K., and Meyer, J. (1998) *Biochemistry 37*, 10429-10437.
7. Chatelet, C., and Meyer, J. (1999) *J. Biol. Inorg. Chem. 4*, 311-317.
8. Chatelet, C., Gaillard, J., Pétillot, Y., Louwagie, M., and Meyer, J. (1999) *Biochem. Biophys. Res. Commun. 261*, 885-889.
9. Meyer, J., Fujinaga, J., Gaillard, J., and Lutz, M. (1994) *Biochemistry 33*, 13642-13650.
10. Golinelli, M. P., Akin, L. A., Crouse, B. R., Johnson, M. K., and Meyer, J. (1996) *Biochemistry 35*, 8995-9002.
11. Yeh, A. P., Chatelet, C., Soltis, S. M., Kuhn, P., Meyer, J., and Rees, D. C. (2000) *J. Mol. Biol. 300*, 587-595.
12. Chatelet, C., and Meyer, J. (2001) *Biochim. Biophys. Acta 1549*, 32-36.
13. Fujinaga, J., Gaillard, J., and Meyer, J. (1993) *Biochem. Biophys. Res. Commun. 194*, 104-111.

14. Crouse, B. R., Meyer, J., and Johnson, M. K. (1995) *J. Am. Chem. Soc. 117*, 9612-9613.

15. Achim, C., Golinelli, M. P., Bominaar, E. L., Meyer, J., and Münck, E. (1996) *J. Am. Chem. Soc. 118*, 8168-8169.

16. Bominaar, E. L., Achim, C., and Borshch, S. A. (1999) *J. Chem. Phys. 110*, 11411-11422.

17. Achim, C., Bominaar, E. L., Meyer, J., Peterson, J., and Münck, E. (1999) *J. Am. Chem. Soc. 121*, 3704-3714.

18. Berg, J. M., and Holm, R. H. (1982) *Structures and reactions of iron-sulfur protein clusters and their synthetic analogs*, John Wiley.

19. Strop, P., Takahara, P. M., Chiu, H. J., Angove, H. C., Burgess, B. K., and Rees, D. C. (2001) *Biochemistry 40*, 651-656.

20. Colbert, C. L., Couture, M. M. J., Eltis, L. D., and Bolin, J. T. (2000) *Structure 8*, 1267-1278.

21. Iwata, S., Saynovits, M., Link, T. A., and Michel, H. (1996) *Structure 4*, 567-579.

22. Bailey, S. (1994) *Acta Crystallogr. Sect. D Biol. Crystallogr. 50*, 760-763.

23. Derewenda, Z. S., Lee, L., and Derewenda, U. (1995) *J. Mol. Biol. 252*, 248-262.

24. Bes, M. T., Parisini, E., Inda, L. A., Saraiva, L. M., Peleato, M. L., and Sheldrick, G. M. (1999) *Struct. Fold. Des. 7*, 1201-1211.

25. Hurley, J. K., Weber Main, A. M., Hodges, A. E., Stankovich, M. T., Benning, M. M., Holden, H. M., Cheng, H., Xia, B., Markley, J. L., Genzor, C., GomezMoreno, C., Hafezi, R., and Tollin, G. (1997) *Biochemistry 36*, 15109-15117.

26. Xiao, Z. G., Lavery, M. J., Ayhan, M., Scrofani, S. D. B., Wilce, M. C. J., Guss, J. M., Tregloan, P. A., George, G. N., and Wedd, A. G. (1998) *J. Am. Chem. Soc. 120*, 4135-4150.

27. Meyer, J., and Moulis, J. M. (2001) in *Handbook of Metalloproteins* (Messerschmidt, A., Huber, R., T., P., and Wieghardt, K., Eds.) pp 505-517, John Wiley and Sons, Chichester, UK.

28. Mansy, S. S., Xiong, Y., Hemann, C., Hille, R., Sundaralingam, M., and Cowan, J. A. (2002) *Biochemistry 41*, 1195-1201.

29. Rebelo, J. M., Dias, J. M., Huber, R., Moura, J. J. G., and Romao, M. J. (2001) *J. Biol. Inorg. Chem. 6*, 791-800.

30. Mayer, S. M., Lawson, D. M., Gormal, C. A., Roe, S. M., and Smith, B. E. (1999) *J. Mol. Biol. 292*, 871-891.

31. Peters, J. W., Stowell, M. H. B., Soltis, S. M., Finnegan, M. G., Johnson, M. K., and Rees, D. C. (1997) *Biochemistry 36*, 1181-1187.

32. Salifoglou, A., Simopoulos, A., Kostikas, A., Dunham, R. W., Kanatzidis, M. G., and Coucouvanis, D. (1988) *Inorg. Chem. 27*, 3394-3406.

33. Otwinowski, Z., and Minor, W. (1997) *Methods Enzymol. 276*, 307-326.

34. Kissinger, C. R., Gehlhaar, D. K., and Fogel, D. B. (1999) *Acta Crystallogr. Sect. D Biol. Crystallogr. 55*, 484-491.

35. Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) *Acta Crystallogr. Sect. D Biol. Crystallogr. 54*, 905-921.

36. Jones, T. A., Zou, J. Y., Cowan, S. W., and Kjeldgaard, M. (1991) *Acta Crystallogr. Sect. A 47*, 110-119.

37. Read, R. J. (1986) *Acta Crystallogr. Sect. A 42*, 140-149.

38. Sheldrick, G. M., and Schneider, T. R. (1997) *Methods Enzymol. 277*, 319-343.

39. Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) *Acta Crystallogr. Sect. D Biol. Crystallogr. 53*, 240-255.

40. Cruickshank, D. W. J. (1999) *Acta Crystallogr. Sect. D Biol. Crystallogr. 55*, 583-601.

41. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) *J. Appl. Crystallogr. 26*, 283-291.

42. Mayerle, J. J., Denmark, S. E., DePamphilis, B. V., Ibers, J. A., and Holm, R. H. (1975) *J. Am. Chem. Soc. 97*, 1032-1045.

Table 4.1. *Summary of data collection statistics.*

	WT	**C55S**	**C59S**
Wavelength (Å)	0.9918	0.8860	0.9580
Space group	C2	C2	C2
Unit cell dimensions			
a (Å)	67.2	67.3	67.3
b (Å)	59.8	59.8	59.8
c (Å)	47.2	46.9	46.8
β(deg.)	110.3	109.8	109.3
Maximum resolution (Å)	1.50	1.25	1.05
Total reflections	98,554	183,676	358,550
Unique reflections	27,758	48,256	77,950
Completeness (%) [a]	98.9 (97.1)	99.7 (99.5)	95.6 (92.0)
I/σ(I)	7.8 (1.5)	28.5 (3.8)	38.2 (6.4)
R_{sym} (%) [b]	5.1 (40.4)	4.7 (26.1)	3.5 (21.6)

[a] Numbers in parentheses correspond to values in the highest resolution shell.

[b] $R_{sym} = (\Sigma_{hkl} \Sigma_i | I_i (hkl) - < I (hkl) > |) / (\Sigma_{hkl} \Sigma_i I (hkl))$.

Table 4.2. *Final* refinement statistics for WT, C55S, and C59S models of Fd4. Average temperature factors are given for subunits A and B individually.

	WT	**C55S**	**C59S**
Resolution limits (Å)	31.5-1.50	44.3 – 1.25	31.2 - 1.05
R-factor[a]	18.4	0.144	0.138
R-free	21.6	0.196	0.162
estimated coordinate error (Å)[b]	0.09	0.05	0.03
No. of nonhydrogen atoms in a.u.			
Protein	1590	1575	1576
Iron-sulfur	8	8	8
Water	187	215	198
Zinc	4	--	--
Sulfate	4	--	--
RMS deviations from ideal values			
Bond lengths (Å)	0.024	0.015	0.017
Bond angles (deg)	2.192	3.056	2.666
Dihedral angles (deg)	25.11	26.48	26.38
Improper torsion angles (deg)	1.60	1.64	1.83
Average temperature factor ($Å^2$)			
Protein	22.1, 20.6	21.0, 19.9	18.1, 16.1
Iron-sulfur	13.0, 12.8	12.9, 12.2	9.5, 9.7
Water	36.8	39.3	31.9
Zinc	37.1	--	--
Sulfate	44.3	--	--
Ramachandran plot[c]			
residues in most favored regions (%)	90.2	93.2	93.2
residues in additional allowed regions (%)	9.1	6.8	6.8
residues in generously allowed regions (%)	0.6	0.0	0.0
residues in disallowed regions (%)	0.0	0.0	0.0

[a] R-factor = $\Sigma(|F_{obs}|-|F_{calc}|) / \Sigma|F_{obs}|$

[b] coordinate errors were obtained from the diffraction-component precision index calculated from the values of R-free by the method of Cruickshank (*40*).

[c] As determined by PROCHECK (*41*).

Table 4.3. Cluster geometry in molecules A and B for WT, C55S, and C59S Fd4.[a]

	Residue	WT		C55S		C59S	
Bonds (Å)		**A**	**B**	**A**	**B**	**A**	**B**
Fe1 – Fe2		2.74	2.72	2.692	2.681	2.697	2.690
S1 – S2		3.52	3.52	3.590	3.592	3.548	3.546
Fe1-S1		2.24	2.22	2.262(7)	2.244(7)	2.216(4)	2.214(3)
Fe1-S2		2.21	2.19	2.215(9)	2.222(8)	2.215(4)	2.217(4)
Fe2-S1		2.23	2.24	2.20(1)	2.200(9)	2.232(4)	2.220(4)
Fe2-S2		2.25	2.25	2.305(8)	2.308(8)	2.217(4)	2.253(4)
Sγ-Fe1	**C9**	2.27	2.26	2.330(7)	2.328(6)	2.303(4)	2.302(4)
Sγ-Fe1	**C22**	2.31	2.28	2.302(8)	2.306(8)	2.302(4)	2.304(4)
Sγ-Fe2	**C55**	2.34	2.37	1.97(1)	2.01(2)	2.318(4)	2.315(5)
Sγ-Fe2	**C59**	2.29	2.33	2.300(9)	2.296(9)	1.940(9)	1.942(8)
Angles (°)							
S1-Fe1-S2		104.8	106.0	106.7(3)	107.1(3)	106.4(2)	106.3(2)
S1-Fe2-S2		103.7	103.3	105.6(3)	105.6(3)	104.6(2)	104.9(2)
Fe1-S1-Fe2		75.6	75.1	74.2	74.2	74.7	74.7
Fe1-S2-Fe2		75.7	75.3	73.1	72.5	74.3	74.0
Sγ-Fe1-Sγ	**C9/C22**	105.0	104.3	105.7(3)	104.6(2)	106.1(1)	106.1(1)
O/Sγ-Fe2-O/Sγ	**C/S55-C/S59**	90.5	90.4	96.6(5)	96.3(5)	100.0(3)	98.2(3)
Sγ-Fe1-S1	**C9**	105.4	106.4	103.9(3)	104.6(3)	104.2(1)	104.4(1)
Sγ-Fe1-S2	**C9**	116.7	116.2	115.7(3)	115.2(2)	116.5(1)	116.1(1)
Sγ-Fe1-S1	**C22**	115.5	113.7	115.7(3)	114.5(3)	115.4(1)	115.4(1)
Sγ-Fe1-S2	**C22**	109.9	110.6	109.5(3)	110.9(3)	108.5(1)	108.9(1)
O/Sγ-Fe2-S1	**C/S55**	115.4	115.7	109.2(6)	109.6(6)	114.7(2)	114.9(2)
O/Sγ-Fe2-S2	**C/S55**	112.7	113.5	112.7(5)	112.0(5)	112.8(1)	112.7(2)
O/Sγ-Fe2-S1	**C/S59**	113.2	112.2	110.8(4)	111.8(4)	107.6(3)	107.9(3)
O/Sγ-Fe2-S2	**C/S59**	121.8	122.3	121.6(3)	121.2(3)	117.5(3)	118.6(3)
Torsion Angle (deg.)							
Fe1-S1-S2-Fe2		-174.4	-172.0	-170.8	-170.6	-176.7	-176.7

[a] Numbers in parentheses correspond to standard uncertainties in the last digit. Due to the lower resolution of the WT structure, coordinate uncertainties of the individual atoms could not be obtained.

Table 4.4. *Average* stereochemical parameters for Fd4 structures discussed in text (from Table 4.3). Standard values are the average values for [2Fe-2S] clusters coordinated by four cysteine ligands in protein structures refined at resolutions $\leq$ 1.4 Å. These structures include PDB entries 1QT9, 1AWD, and the two distinct clusters in 1HLR (*5, 24, 29*). The model compound parameters are derived from the structure of $(Fe_2S_2(SC_6H_4CH_3)_4)^{2-}$ (*42*).

Parameter	WT	C55S	C59S	Protein Standards	Model Compound
Fe-S (Å)	2.23	2.24	2.23	2.23 ± 0.03	2.201
Fe—Fe (Å)	2.73	2.69	2.69	2.73 ± 0.02	2.691
S—S (Å)	3.52	3.59	3.55	3.51 ± 0.04	3.483
S-Fe-S (deg.)	104.5	106.2	105.6	104.3 ± 1.8	104.6
Fe-S-Fe (deg.)	75.4	73.5	74.4	75.5 ± 1.0	75.4
9Sγ -Fe- 22Sγ (deg.)	104.6	105.1	106.1	105.1 ± 1.7	111.2
55S/Oγ–Fe–59S/Oγ (deg.)	90.5	96.4	99.1	--	--
|Fe-S-S-Fe| torsion angle (deg.)	173	171	177	175	180

Table 4.5. Hydrogen bonding geometry in the [2Fe-2S] cluster environment. D (donor), A (acceptor), with separation distance measured in Å. Although not specifically indicated, the D-H–A angles for these interactions are all > 125°.

		WT		C55S		C59S	
		A	B	A	B	A	B
D	A	D-A	D-A	D-A	D-A	D-A	D-A
Cys22N	S1	3.39	3.44	3.43	3.43	3.40	3.40
Met56N	S2	3.26	3.29	3.35	3.34	3.34	3.35
Ala58N	S2	3.55	3.55	3.57	3.45	3.63	3.58
Arg13NH1	S1	3.54	3.42				
Wat A	S1			3.56	3.62	3.76	3.73
Gln11N	Cys9Sγ	3.58	3.52	3.57	3.48	3.55	3.54
Val64N	Cys22Sγ	3.51	3.6	3.48	3.53	3.54	3.57
Wat B	Cys22Sγ	3.11	3.19	3.19	3.25	3.30	3.28
Arg13NH2	55S/Oγ	3.17	3.06	3.68	3.74	3.42	3.48
Asn57N	55S/Oγ	3.25	3.29	3.51	3.54	3.30	3.35
Wat C	55S/Oγ					3.84	3.73
Wat D	55S/Oγ			2.76			
Wat E	55S/Oγ					3.63	4.15
Arg13NH1	59S/Oγ	3.46	3.39				
Wat A	59S/Oγ			3.36	3.35	2.93	2.95
Wat C	59S/Oγ					2.63	2.58
Wat D	59S/Oγ			3.77			
Wat F	59S/Oγ	3.45	3.42				
Wat G	59S/Oγ		3.67		3.79		

Table 4.6. Average positional shifts (in Ångstroms) of the Cα, Sγ, and/or Oγ atoms between WT, C55S, and C59S structures. Superpositions were based on the Cα positions of residues 3-103 of the corresponding subunits of the respective pairs.

Atom		C55S/WT	C59S/WT	C55S/C59S
Cys9	Cα	0.15	0.16	0.04
Cys9	Sγ	0.12	0.14	0.05
Cys22	Cα	0.12	0.11	0.05
Cys22	Sγ	0.14	0.11	0.09
Cys/Ser55	Cα	0.12	0.07	0.15
Cys/Ser55	Sγ/Oγ	0.64	0.19	0.56
Cys/Ser59	Cα	0.23	0.42	0.24
Cys/Ser59	Sγ/Oγ	0.25	0.70	0.74
FES	Fe1	0.13	0.13	0.13
FES	S1	0.10	0.15	0.19
FES	Fe2	0.18	0.17	0.21
FES	S2	0.21	0.12	0.19

Figure 4.1. Superposition of the A subunits of the former (orange) and current (cyan) WT, C55S (yellow), and C59S (purple) variant forms of Fd4. The green line denotes the twofold rotation axis relating subunits A (left) and B (right) in a dimer, while the magenta line corresponds to the axis about which the B subunit of the previous WT structure (*11*) is rotationally shifted relative to the B subunit of the current WT form. This rigid body rotational shift as well as regions of the old form that differ significantly from the new form (residues 13-20 and 39-46 highlighted in red) are likely the results of the different crystal packing between the two WT forms. A sulfate anion lying along the twofold axis may further stabilize the dimer interface in the new WT structure, although this species is not present in the C55S or C59S variants.

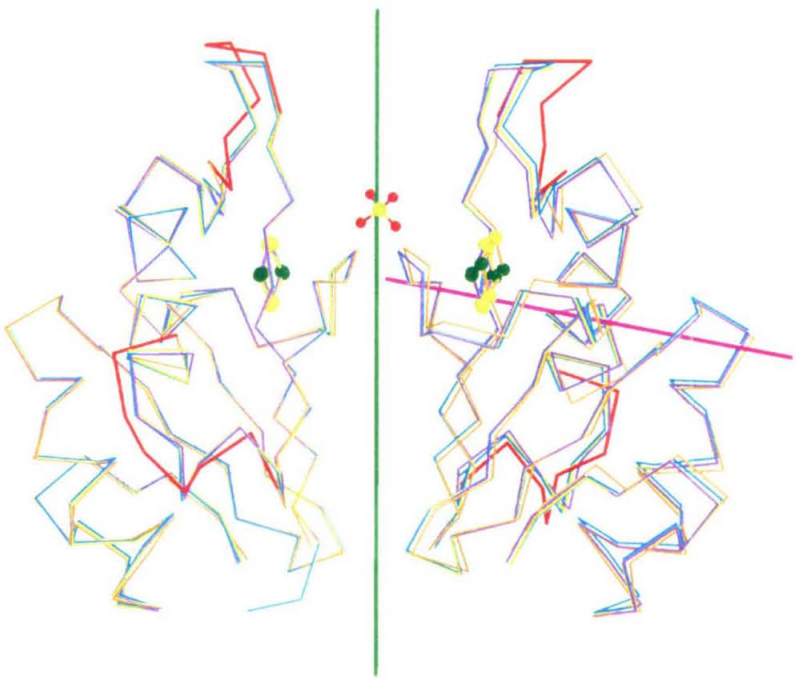


Figure 4.2. Stereoview of the [2Fe-2S] cluster and its immediate environment, showing the extensive hydrogen bonding network (listed in Table 4.5) involving the cluster, ligands, and surrounding residues. Iron and sulfur atoms are colored green and yellow, respectively. Fe1 is ligated by Cys9 and Cys22 while Fe2 is ligated by Cys55 and Cys59.

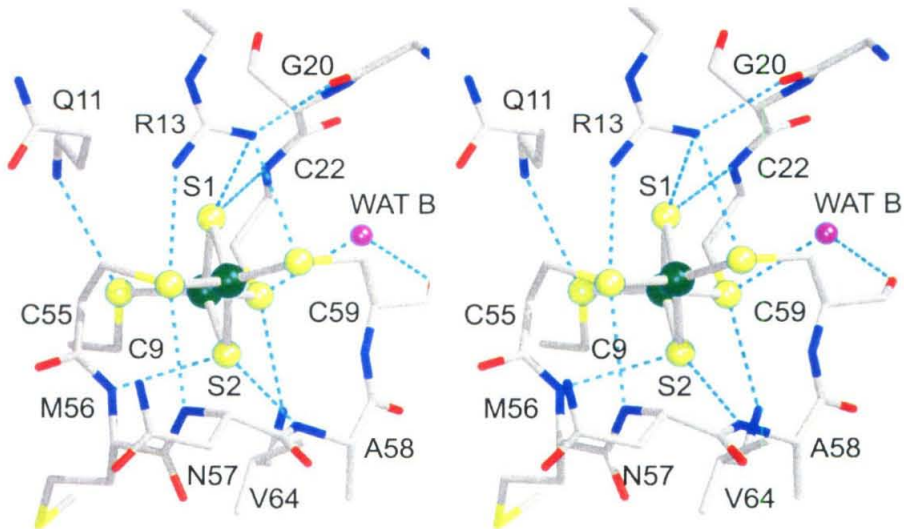


Figure 4.3. Stereoviews of the potential interactions between Cys ligands across the Fe_2S_2 face of iron-sulfur clusters. (a) The [2Fe-2S] cluster of Fd4, illustrating the conformation of Cys55 in which the Cα-H group is positioned over the cluster to hydrogen bond (dashed line) with the Cys9 Sγ. (b) a similar configuration is observed in the HiPIP from *Chromatium vinosum* ((*43*), PDB entry 1CKU); while Cys 43 and 46 have the same general relationship to the common face of the [4Fe-4S] cluster in this protein, the increased pucker of the Fe_2S_2 core precludes the formation of a Cα-H---Sγ hydrogen bond. Two of the cysteine ligands in HiPIP have been omitted from the figure for clarity. Iron, sulfur, nitrogen, oxygen and Cα atoms are colored green, yellow, blue, red, and black, respectively.

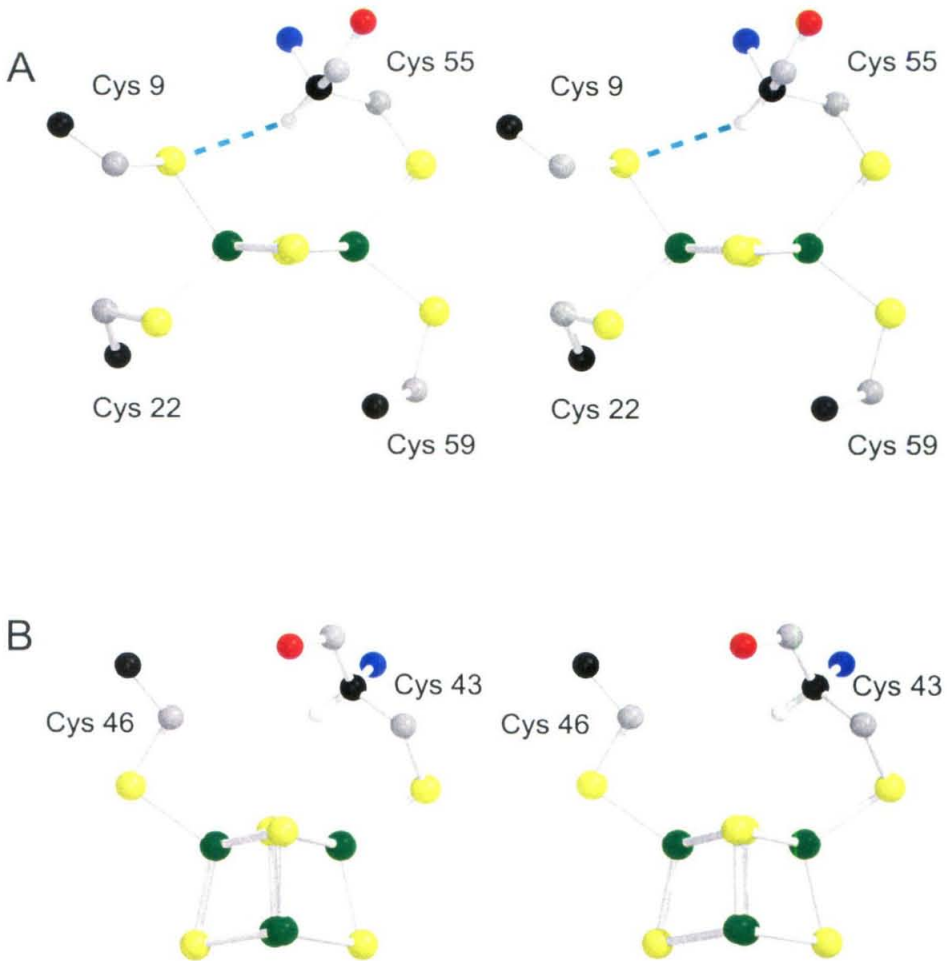


Figure 4.4. *The* [2Fe-2S] cluster, its ligands, and local secondary structure showing the shorter bond formed between (a) Fe2 and the Ser55 Oγ atom in the C55S structure and (b) Fe2 and the Ser59 Oγ atom in the C59S structure. In both figures, simulated-annealing $2|F_o|-|F_c|$ omit electron density contoured at 1.5 σ level is shown as a cyan mesh.

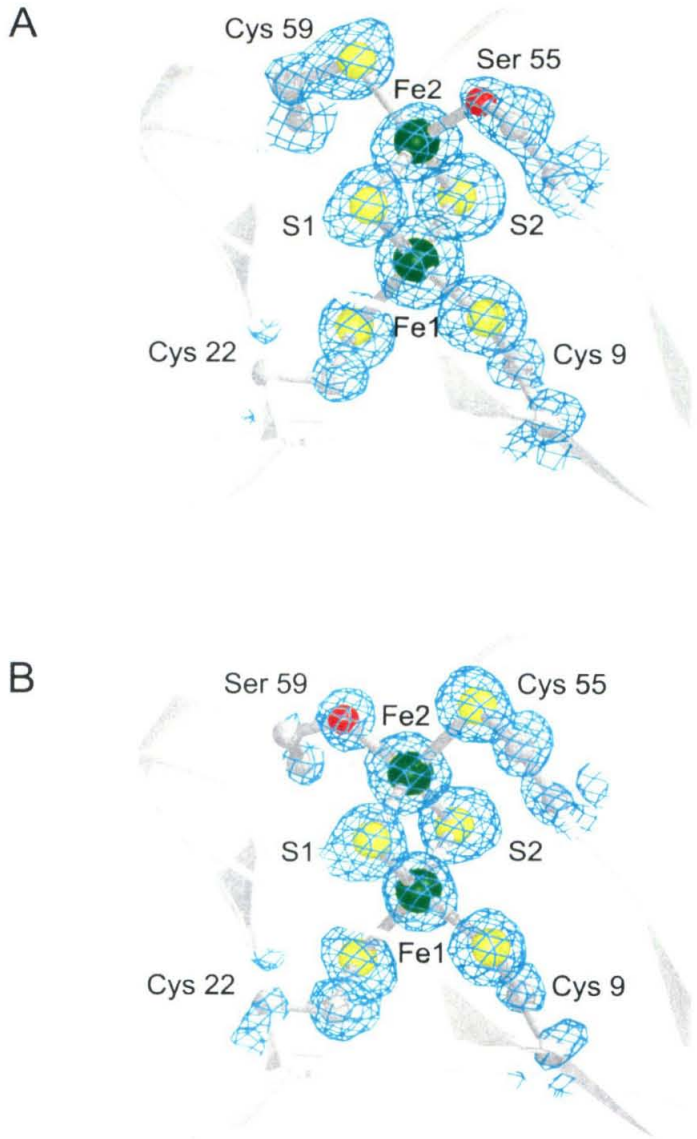


Figure 4.5. (a) The [2Fe-2S] cluster and its ligands from the WT (cyan), C55S (yellow), and C59S (purple) structures upon superposition of the corresponding 101 Cα atoms of the three structures, showing the varying degree of positional shifts that occur in the inorganic core as well as residues 55 and 59 due to the cysteine to serine substitutions. (b) perpendicular view showing the varying degree to which the inorganic core is distorted in each structure. The color scheme is the same as in panel (a).

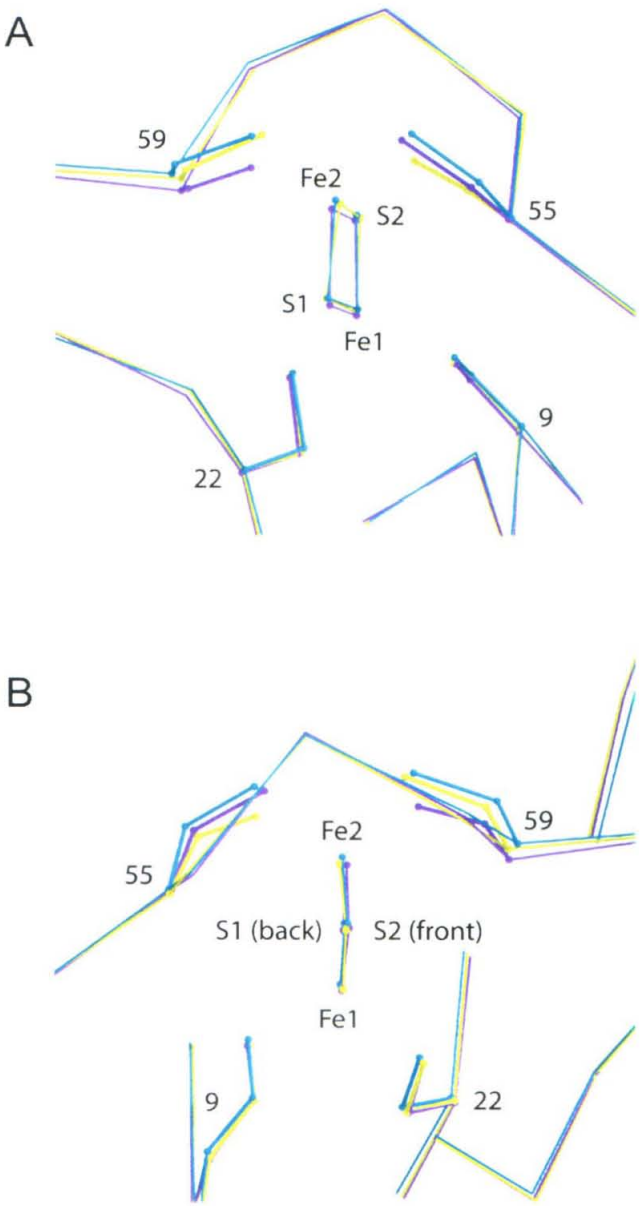
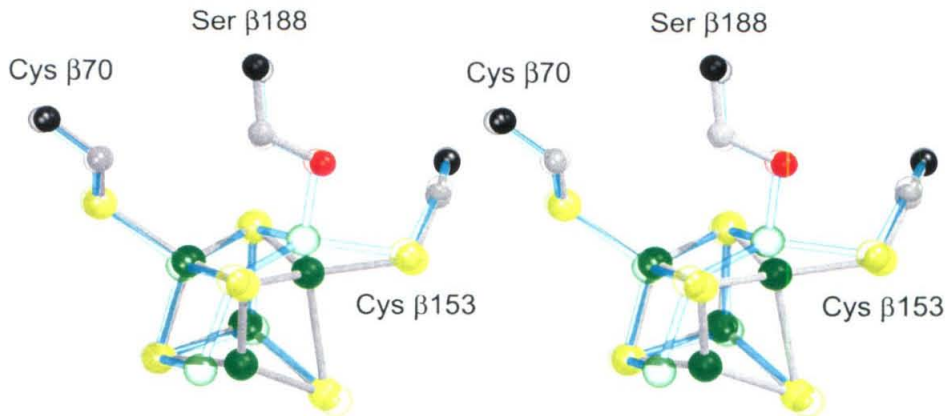


Figure 4.6. Stereoview comparing the P-cluster of nitrogenase in the oxidized (transparent ball-and-stick model in cyan) and reduced (solid ball-and-stick model in gray) states (*30, 31*). In the oxidized state, one of the irons is coordinated by the sidechain of Ser β188, while in the reduced state this iron is shifted and coordinates an inorganic sulfur in the cluster instead. The coloring scheme is as in Figure 4.3. PDB entries 2MIN (oxidized) and 3MIN (reduced) were used for this figure.



CHAPTER 5

Crystal Structures of the Stigmatellin-Bound *Rhodobacter sphaeroides* Photosynthetic Reaction Center in the Dark- and Light-Adapted States: Towards the Elucidation of Structural Changes Associated with the Light-Induced $D^+Q_A^-$ Charge-Separated State

Crystal Structures of the Stigmatellin-Bound *Rhodobacter sphaeroides* Photosynthetic Reaction Center in the Dark- and Light-Adapted States: Towards the Elucidation of Structural Changes Associated with the Light-Induced $D^+Q_A^-$ Charge-Separated State

Andrew P. Yeh[1], Edward C. Abresch[2], S. Michael Soltis[3], Herbert L. Axelrod[2], Melvin Y. Okamura[2], George Feher[2], and Douglas C. Rees[1,4†]

1 Division of Chemistry and Chemical Engineering, 147-75CH, California Institute of Technology, Pasadena, California 91125

2 Department of Physics, 0319, University of California, San Diego, 9500 Gilman Drive, La Jolla, California 92093

3 Stanford Synchrotron Radiation Laboratory, P.O. Box 4349, MS 69, Stanford, California 94309

4 Howard Hughes Medical Institute, California Institute of Technology, Pasadena, California 91125

† Corresponding author

E-mail address of the corresponding author: dcrees@caltech.edu

Abstract

In the *Rhodobacter sphaeroides* photosynthetic reaction center (RC), structural changes are postulated to occur following absorption of light to generate the charge-separated states (i.e., $D^+Q_A^-Q_B$ and $D^+Q_AQ_B^-$). Kleinfeld *et al.* provided compelling evidence for this when they showed that the kinetics of electron transfer in RCs cooled to cryogenic temperatures in the dark (i.e., neutral state) differ from those cooled to cryogenic temperatures under illumination (i.e., charge-separated state) (*1*). Structural changes were indeed observed for RC in the $D^+Q_AQ_B^-$ charge-separated state. Towards understanding the difference in electron-transfer kinetics observed between RCs in the neutral DQ_AQ_B and charge-separated $D^+Q_A^-Q_B$ states, the crystal structures of *Rb. sphaeroides* RC, complexed with the herbicide stigmatellin, in both the neutral (DQ_AS) and charge-separated ($D^+Q_A^-S$) states have been determined to 2.2 Å and 2.66 Å resolution, respectively. The structures reveal that there is little structural change in the protein and cofactors upon formation of the $D^+Q_A^-S$ charge-separated state.

Introduction

The primary process of bacterial photosynthesis, light-induced trans-membrane charge separation, occurs in the reaction center (RC), an integral membrane protein-pigment complex. In the photosynthetic bacterium *Rhodobacter sphaeroides*, the RC comprises three protein subunits (L, M, and H) and a host of cofactors (four bacteriochlorophylls, two bacteriopheophytins, two ubiquinones (UQ-10), and one non-heme high-spin iron (Fe^{+2})). Charge separation occurs through a series of electron transfers via one branch of these cofactors. Absorption of a photon excites the donor, D, a specialized bacteriochlorophyll dimer, whereupon an electron is transferred from D via cofactors along the A-branch to the primary acceptor ubiquinone (Q_A). The Q_A^- semiquinone in this $D^+Q_A^-Q_B$ charge-separated state subsequently reduces the secondary acceptor ubiquinone (Q_B) to form the $D^+Q_AQ_B^-$ charge-separated state. Following the reduction of D^+ by cytochrome c_2, a second photon is absorbed, resulting in another electron transfer from D to Q_B^-. The doubly reduced Q_B^{-2} uptakes two protons to form Q_BH_2, which subsequently leaves the RC and is replaced by an exogenous quinone. The subsequent oxidation of Q_BH_2 back to Q_B, reduction of cytochrome c_2, and the pumping of protons across the membrane are all carried out by the cytochrome bc_1 complex. The proton gradient that is generated in such fashion provides the proton-motive force necessary for subsequent ATP synthesis (*2*).

A number of investigations point to the possibility that protein structural changes accompany the formation of the different charge-separated states (*1, 3*). Kleinfeld *et al.* provided compelling evidence for this when they showed that the kinetics of electron

transfer in RCs cooled to cryogenic temperatures in the dark (i.e., neutral state) differ from those cooled to cryogenic temperatures under illumination (i.e., charge-separated state) (*1*). In particular, they observed that (i) the rate of electron transfer from $D^+Q_A^-$ back to DQ_A (i.e., the rate of the charge recombination reaction) was smaller in RCs frozen under illumination than those frozen in the dark, and (ii) the rate of electron transfer from Q_A^- to Q_B was larger by several orders of magnitude in RCs frozen under illumination than those frozen in the dark.

To understand the above effects, structural information about the charge-separated states is necessary. Towards this end, we have undertaken to determine the crystal structure of RC in its different charge-separated states. The previously determined crystal structure of RC in one of its charge-separated states, $D^+Q_AQ_B^-$, showed significant movement (~4.5 Å) of Q_B upon formation of $D^+Q_AQ_B^-$, and based on this and other structural changes observed at and around the Q_B site, a model to explain the effect observed for the dark and light Q_A^- to Q_B electron transfer kinetics as described above was proposed (*4*). To determine any structural changes associated with the light-induced $D^+Q_A^-$ state, we have prepared RCs in which Q_B was replaced with the herbicide stigmatellin (Figure 5.1) and have determined the structure of this complex in the dark- (DQ_AS) and light- ($D^+Q_A^-S$) adapted states.

Results

Structure of stigmatellin-bound RC in the dark-adapted (DQ_AS) state

The refined model at 2.2 Å resolution of dark-adapted *Rb. sphaeroides* RC complexed with stigmatellin (DQ_AS) consists of two noncrystallographic symmetry related RC molecules that are nearly identical to each other, as indicated by an rmsd of 0.26 Å between 828 Cα atoms of the corresponding L, M, and H subunits. The overall structure of DQ_AS is very similar to previously reported RC structures (Figure 5.2) (*4, 5*). The rmsd between the Cα's of the L, M, and H subunits of DQ_AS and the corresponding atoms of a dark-adapted state RC, for example, (DQ_AQ_B; PDB entry 1AIJ) (*4*) is 0.20 Å

Examination of the simulated annealing $2|F_o|-|F_c|$ omit electron density revealed stigmatellin to be bound in the Q_B pocket at the site proximal to Fe^{+2}. The overall mode of binding is very similar to that reported for the structure of *Blastochloris viridis* RC complexed with stigmatellin (*6*), with interactions formed between stigmatellin and the protein environment in the two complexes being nearly identical (Figure 5.3). Superposition of DQ_AS with the DQ_AQ_B and $D^+Q_AQ_B^-$ structures (PDB entries 1AIJ and 1AIG) (*4*) reveals stigmatellin to be bound in a very similar fashion to Q_B^- in terms of its position and orientation (Figure 5.4a), although the stigmatellin bicyclic ring extends somewhat further into the pocket than does the Q_B^- ring. The angle between the planes of the stigmatellin and Q_B^- rings is approximately 13°, and like Q_B^-, stigmatellin makes multiple (seven) interactions with its surrounding protein environment, many of which are analogous to those made by Q_B^- (Table 5.3). In comparison to Q_B^-, the hydroxyl O8 atom of stigmatellin replaces the 1-carbonyl oxygen atom of Q_B^- and is within hydrogen

bonding distance to the Oγ of Ser L223 and the backbone amide nitrogens of Ile L224 and Gly L225. Furthermore, Nδ1 of His L190, which forms a hydrogen bond with the 4-carbonyl oxygen of Q_B^- in the $D^+Q_AQ_B^-$ structure, appears capable of forming a bifurcated hydrogen bond with the 4-carbonyl oxygen and 5-methoxy oxygen of stigmatellin. Finally, the 7-methoxy oxygen of stigmatellin is within hydrogen bonding distance of Asp L213 and the amide nitrogen of Thr L226.

The binding of stigmatellin in the Q_B pocket is associated with slight structural perturbations of residues lining the pocket. Compared to corresponding residues in the DQ_AQ_B structure, Ser L223 in DQ_AS moves slightly away from stigmatellin to accommodate the stigmatellin bicyclic ring, while the side chain of Asp L213 rotates slightly such that it can now form hydrogen bonds with the methoxy O7 as well as a water. Moreover, the phenyl ring of Phe L216 is reoriented by ~30° from the ring of the corresponding residue in the DQ_AQ_B structure so that it is now approximately parallel with the stigmatellin head group. This reorientation of the Phe L216 phenyl ring both prevents its steric hindrance with the stigmatellin tail and also allows it to form more favorable van der Waals interactions with the stigmatellin ring (Figure 5.4b).

In the Q_A pocket of the DQ_AS structure, strong simulated annealing $2|F_o|-|F_c|$ omit electron density was observed upon removal of Q_A from the model, thereby allowing the unbiased fit to the density of the Q_A ring and five of ten isoprenoid units of the chain. Comparison with the DQ_AQ_B structure reveals Q_A to be positioned similarly in both structures. The O1 and O4 atoms from the two carbonyl groups of Q_A are within

hydrogen bonding distance to the Nδ1 of His L190 and the amide nitrogen of Ala M260, respectively. Other residues, which do not form hydrogen bonds with Q_A, but are in van der Waals contact and appear integral to Q_A binding include Leu M215, Trp M252, Met M256, Phe M258, Ile M265, and Trp M268.

Additional features of DQ_AS, compared to the DQ_AQ_B structure, include a greater number of total water molecules (539 vs. 468) as well as lauryl-*N*,*N*-dimethylamineoxide (LDAO) detergent molecules (30 vs. 4) per the two RCs in the asymmetric unit. The majority of water molecules that are proposed to be part of the proton transfer pathway to Q_B (*4, 7*) are conserved in the DQ_AS structure. Moreover, compared to the DQ_AQ_B structure, an additional six water molecules are present in the vicinity of the Q_B binding pocket in the DQ_AS structure. These waters are within hydrogen-bonding distance to Asp L210, Asp L213, Thr L226, Glu H173, and Arg H177. Of these residues, Asp L213 and Ser L223 have been postulated to be associated with the proton transfer pathway to Q_B.

LDAO detergent molecules that were observed in the DQ_AQ_B structure are conserved in the DQ_AS structure; moreover, the electron density revealed several additional detergent molecules that are bound near the protein-lipid interface of the transmembrane α-helices (Figure 5.5). Two of these detergents are bound near the cytoplasmic side in a cleft formed by residues from the H subunit (Ile H28 of the transmembrane α-helix, Pro H55, and Phe H56) and from the L and M subunits (Pro L28 and Phe M258). It is interesting to note that in the structures of *Bcl. viridis* and *T. tepidum* RC (PDB entries 3PRC and 1EYS), lipid or detergent molecules are also

observed in the corresponding site (*8*). At the site in *Bcl. viridis* RC, two LDAO detergent molecules and a sulfate anion are bound while in *T. tepidum* RC, one phospholipid molecule (dipalmitoyl-3-sn-phosphatidyl-ethanolamine) is bound. Two additional LDAO molecules are present on the other side of the DQ_AS structure, at a site related to the one described above by the pseudo twofold symmetry axis present in the RC. This site consists of a cleft, formed by residues from the N-termini of an M subunit α-helix (M55-M62) and a β-sheet comprising stretches of β-strands from both the L and M subunits (M30-M33, M47-M50, and L220-L222) (Figure 5.5).

Structure of stigmatellin-bound RC in the light-adapted ($D^+Q_A^-S$) state

The overall structure at 2.66 Å resolution of the RC in the $D^+Q_A^-S$ state is very similar to that in the DQ_AS state, with an rmsd between 828 Cα atoms of the corresponding L, M, and H subunits of 0.21 Å. One main difference between the structures is that, due to the lower resolution of the data, fewer water and LDAO detergent molecules are observed in the $D^+Q_A^-S$ structure compared to in the DQ_AS structure.

The simulated annealing $2|F_o|-|F_c|$ omit electron density contoured at one sigma level reveals the presence of stigmatellin in the Q_B binding pocket of the $D^+Q_A^-S$ structure; however, the electron density is weaker than for that observed in the DQ_AS structure, suggesting that either the stigmatellin in $D^+Q_A^-S$ is either more disordered and/or at a lower occupancy. Further evidence that this may be the case stems from the higher average B-factor for the stigmatellin (73 $Å^2$) compared to that for all the other cofactor

atoms (54 $Å^2$). Nevertheless, contouring the electron density at a lower sigma level (~ 0.5 σ) allowed placement of stigmatellin into the model and revealed that the mode of binding is nearly identical to that in the DQ_AS structure. Several residues that are in contact with stigmatellin, including Asn L213, Phe L216, Ser L223, and Ile L224, also appear to exhibit greater disorder in the $D^+Q_A^-S$ structure, as evidenced by their weaker electron density. Finally, slightly removed from the Q_B pocket is Glu H173, a residue which has been proposed to play an important role in the mechanism of the proton-coupled electron transfer to Q_B. This residue in $D^+Q_A^-S$ appears disordered, as indicated by the lack of electron density for the side chain. This is similar to what was observed in the light-adapted native structure ($D^+Q_AQ_B^-$) as well.

As was the case for the Q_B pocket, the simulated annealing $2|F_o|-|F_c|$ omit electron density at the Q_A site in $D^+Q_A^-S$ is weaker than in the DQ_AS structure (Figure 5.6). While the weaker electron density would appear to suggest that Q_A is either more disordered and/or at lower occupancy in the $D^+Q_A^-S$ structure, comparison of the average B-factor of Q_A^- (47 $Å^2$) with that of the other cofactors (55 $Å^2$) is not indicative of that. Displaying the electron density at a lower contour level allowed for the placement of the Q_A^- ring as well as six of the ten units of the isoprenyl tail. Superposition of the DQ_AS and $D^+Q_A^-S$ structures shows that the position and orientation of Q_A in both its neutral and reduced forms are virtually identical (Figure 5.7). Weaker electron density was observed not only for Q_A^- but also for several residues lining the Q_A^- pocket, including Trp M268 (located on the other side of the Q_A^- pocket), and Trp M252, Arg M253, Trp M254, and Thr M255 (all of which reside at the C-terminus of an α-helix that partly forms the other side of the

Q_A^- pocket). Again, the weaker electron density for these residues suggests that these residues may be more disordered than they are in RC in the charge-neutral state.

Discussion

Stigmatellin binding

The structure of *Rb. sphaeroides* RC complexed with stigmatellin (DQ_AS) reported here is in agreement with the previously reported structure (*6*) of the complex between *Bcl. viridis* RC and stigmatellin in terms of the mode of binding in the Q_B pocket. These structures provide a structural basis for the potent inhibitory effects of stigmatellin in bacterial photosynthetic reaction centers. In addition to the hydrogen bonds that are formed by Q_B^-, stigmatellin participates in three additional hydrogen bonds. These additional hydrogen bonds, which are formed between the methoxy O7 and the Asp L213 Oδ1 atom and the amide nitrogen of Thr L226, and between the methoxy O5 and His L190 Nδ1, may contribute to the higher affinity of stigmatellin to the Q_B pocket.

The DQ_AS structure also provides a structural rationale for herbicide resistance conferred by several RC mutants: Ser L223 → Ala, Ile L229 → Met, and Tyr L222 → Gly. With the exception of Tyr L222, these mutated residues all contact the stigmatellin bicyclic ring. Ser L223 plays a crucial role in stigmatellin binding due to the hydrogen bond that is formed with the stigmatellin O8 atom, while Ile L229 forms van der Waals interactions with the stigmatellin chromone ring. Tyr L222, while not in direct contact with stigmatellin, is in a stretch of residues that form part of the Q_B pocket and may

contribute to the structural rigidity of that region. The substitution of Tyr L222 with Gly may result in an increased flexibility that leads to instability of the binding pocket.

Greater disorder at the Q_A and Q_B pockets in the $D^+Q_A^-S$ state

Comparison of the electron density at the corresponding Q_A and Q_B pockets in the DQ_AS and $D^+Q_A^-S$ structures reveals these regions to be more disordered in the $D^+Q_A^-S$ than in the DQ_AS structure. One possible cause of the apparent greater disorder observed for stigmatellin and surrounding residues that form the Q_B pocket is that a small percentage of the stigmatellin dissociates from the Q_B pocket and is replaced by exogenous Q_B from the surrounding environment upon the one electron reduction of Q_A. Evidence suggesting such a possibility stems from a study which showed that the dissociation constant of various Q_B pocket binding inhibitors, including stigmatellin, increases in the presence of Q_A^- (*9*). Of the inhibitors tested, however, such an effect was the smallest with stigmatellin. Nevertheless, if Q_B were indeed to have replaced stigmatellin in a fraction of the RCs upon illumination, a mixture of both $D^+Q_A^-S$ as well as $D^+Q_AQ_B^-$ states would have been present in the crystal. This may in particular explain the disorder observed for Glu H173 in the $D^+Q_A^-S$ structure, as such disorder was observed for the same residue in the $D^+Q_AQ_B^-$ structure. Another possible cause of the greater disorder observed at the quinone binding pockets is the intense illumination to which the crystal was subjected. A recent study demonstrated that illumination of carotenoidless RC by intense light (1100 $\mu Em^{-2}s^{-1}$) causes protein damage, including modification of the Q_B binding site, via the formation of singlet oxygen radical species (*10*).

Structural changes in RC in the light-induced $D^+Q_A^-S$ state

Previous studies have suggested that structural changes within the RC may occur upon the reduction of Q_A to Q_A^-. Based on the kinetics observed for $D^+Q_A^- \rightarrow DQ_A$ charge recombination in RCs cooled to cryogenic temperatures either under illumination or in the dark, for example, Kleinfeld *et al.* suggested the possibility of an approximately 1 Å increase in the distance between the "specialized pair" bacteriochlorophyll dimer (D) and Q_A upon Q_A reduction (*1*). Based on observations from light-induced voltage changes in RCs, on the other hand, Brzezinski *et al.* proposed that either Q_A^- may move towards D^+ after charge separation, or proton transfer between protonatable residues of the RC and the solvent environment may occur (*3*). Significant movement of neither Q_A upon one electron reduction nor of D^+ occurs, as a comparison of the DQ_AS and $D^+Q_A^-S$ structures reveals no detectable changes in the structure and position of the cofactors, including D and Q_A, upon $D^+Q_A^-$ charge separation. This is in agreement with EPR (*11, 12*) and FTIR (*13, 14*) studies, which have suggested minimal repositioning of Q_A upon reduction. The extent of changes that are observed for Q_A^- by such spectroscopic methods include a slight reorientation ($\leq 8°$) of the Q_A^- ring with respect to that of neutral Q_A and slight increases and decreases (by ~0.05 Å) in the lengths of the C=O and C=C bonds, respectively, in the Q_A ring upon reduction. While such changes associated with Q_A^- may indeed occur, our structure of RC in the $D^+Q_A^-S$ at present cannot confirm these changes due to the current resolution (2.66 Å) of the model. Proton transfer between solvent and protonatable amino acids in the RC, as proposed by Brzezinski *et al.*, also can neither be eliminated nor confirmed by the current $D^+Q_A^-S$ structure as such

conformational changes would necessitate structures of significantly higher-resolution (~ ≤ 1 Å) to visualize the protonation states of individual amino acids.

Conclusions

The structures of stigmatellin-bound RCs in the dark- and light-adapted states at 2.2 Å and 2.66 Å, respectively, reveal no significant structural perturbations associated with formation of the $D^+Q_A^-$ charge-separated state. The absence of significant displacement of Q_A upon charge separation is consistent with spectroscopic evidence, which suggest at most a slight reorientation of the Q_A^- ring and slight lengthening and shortening of bonds in the Q_A^- ring. The origins of the kinetic effects observed by Kleinfeld *et al.* may be due to structural changes that are currently beyond the limits of the current resolutions in our investigation. It is hoped that higher-resolution structures of the RC in both DQ_AS and $D^+Q_A^-S$ states may be achieved in the future and provide a clearer understanding of structural changes associated with the $D^+Q_A^-$ state.

Materials and Methods

Generation of stigmatellin-bound RC crystals

RCs from *Rb. sphaeroides* were expressed, purified, and crystallized as previously described (*15*). Crystals belong to space group $P4_32_12$ and contain two RC molecules per asymmetric unit (Table 5.1). Stigmatellin (Fluka) was added to RC molecules in two different manners. For RC crystals that were used for the DQ_AS state structure, 0.5 µl of a 10 mM stigmatellin solution in ethanol was added to ~10 µl of mother liquor containing fully grown crystals several days before data collection. The

crystals used for the $D^+Q_A^-S$ state structure determination were grown in the presence of 100 μM stigmatellin.

The charge recombination kinetics (i.e., $D^+Q^- \rightarrow DQ$) upon photoexcitation of the RC crystals were measured to determine the fraction of RCs in the crystal containing Q_B. With Q_B bound to the RC, the charge recombination time ($t_{1/e}$) is ~1 s. In the absence of or displacement of Q_B by stigmatellin, charge recombination occurs from Q_A^- with a $t_{1/e}$ of 0.1 s. The charge-recombination kinetics were modeled with a two exponential fit. For crystals used in both the DQ_AS and $D^+Q_A^-S$ structure determinations, the fraction of slow charge-recombination (i.e., $t_{1/e} \cong 1$ s) was 0.2 ± 0.1, indicating ~20% Q_B occupancy.

Data collection, structure determination, and refinement

Diffraction data to 2.2 Å resolution for RC in the DQ_AS state were collected under cryogenic conditions on beamline 7-1 ($\lambda = 0.98$ Å) at the Stanford Synchrotron Radiation Laboratory (SSRL) with a 30 cm MAR Research imaging plate system. RC in the $D^+Q_A^-S$ state was prepared by illuminating an RC crystal of 0.15 mm thickness for ~30 ms with a filtered tungsten light source (bandpass 400 nm to 900 nm) of an intensity of 1 W/cm^2 and then immediately plunging the crystal in a liquid nitrogen bath to trap the charge-separated state. A control crystal of similar thickness was found under the same illumination conditions to be ~80% in the charge separated $D^+Q_A^-S$ state. Diffraction data to 2.66 Å resolution for $D^+Q_A^-S$ were collected at ~70 K by using a liquid helium cryostat on beamline 9-2 at SSRL ($\lambda = 1.033$ Å) with an Area Detector Systems Corp.

(ADSC) Quantum-4 CCD detector. Data sets were processed and scaled using DENZO and SCALEPACK (*16*) (Table 5.1).

Based on the scaling of the structure factors calculated from various *Rb. sphaeroides* RC models to the DQ_AS diffraction data, the structure of RC in the $D^+Q_AQ_B^-$ charge-separated state (*4*) (PDB entry 1AIG) was deemed to be the best starting model. The structure, with the waters and both quinones removed, was used as the starting model for the DQ_AS structure. Positional and individual B-factor refinement of the initial structure was carried out in CNS (*17*), whereupon $2|Fo|-|Fc|$ σ_a-weighted and $|Fo|-|Fc|$ σ_a-weighted electron density maps (*18*) were calculated. The maps revealed electron density corresponding to Q_A and stigmatellin in the Q_A and Q_B pockets, respectively. Based on the electron density at these pockets, Q_A and stigmatellin were built into the model at their respective sites. Multiple rounds of positional and individual isotropic B-factor refinement with the program CNS (*17*) were alternated with model rebuilding in the molecular graphics program O (*19*). Noncrystallographic symmetry (NCS) restraints were applied in the initial rounds of refinement and released in later rounds. Addition of solvent and detergent molecules and completion of refinement with CNS resulted in a final R-factor and R-free of 20.9% and 25.1%, respectively. The final DQ_AS model comprises two RC molecules, 538 water molecules, and 30 lauryl-*N*,*N*-dimethylamineoxide (LDAO) detergent molecules. Each RC molecule consists of three protein subunits, the bacteriochlorophyll dimer special pair (D_A and D_B), two accessory bacteriochlorophylls (BCh_A and BCh_B), two bacteriopheophytins (BPh_A and BPh_B), one ubiquinone (Q_A), one stigmatellin, and one non-heme iron. Residues 1-281, 3-301, and

11-258 of the L, M, and H subunits, respectively, of both RC molecules are present in the model.

The refined DQ_AS model, excluding water molecules, was used as the starting structure for the $D^+Q_A^-S$ model. The initial model was refined using positional and individual B-factor protocols in CNS, whereupon both $2|F_o|-|F_c|$ and $|F_o|-|F_c|$ simulated annealed (SA) omit maps of the region surrounding Q_A and stigmatellin were generated. Both SA omit maps revealed electron density corresponding to Q_A and stigmatellin in their respective binding pockets. Q_A and stigmatellin were subsequently incorporated into the model. Upon further iterative rounds of model refinement and rebuilding, and solvent and detergent addition, the R-factor and R-free converged to 20.8% and 27.3%, respectively. The final $D^+Q_A^-S$ model consists of two RC molecules, 185 water molecules, and six LDAO detergent molecules. The composition of the RC molecules is nearly identical to that of the DQ_AS state, with the exception that two additional residues at the N-terminus of the M and O subunits, which were not observed in the electron density maps of DQ_AS, were included in the model. Final refinement and model statistics for RC in the DQ_AS and $D^+Q_A^-S$ states are listed in Table 5.2.

References

1. Kleinfeld, D., Okamura, M. Y., and Feher, G. (1984) *Biochemistry 23*, 5780-5786.
2. Cramer, W. A., and Knaff, D. B. (1990) *Energy Transduction in Bioilogical Membranes*, Springer-Verlag, New York.
3. Brzezinski, P., Okamura, M. Y., and Feher, G. (1992) in *The Photosynthetic Bacterial Reaction Center II* (Breton, J., and Verméglio, A., Eds.) pp 321-330, Plenum Press, New York.
4. Stowell, M. H. B., McPhillips, T. M., Rees, D. C., Soltis, S. M., Abresch, E., and Feher, G. (1997) *Science 276*, 812-816.
5. Ermler, U., Fritzsch, G., Buchanan, S. K., and Michel, H. (1994) *Structure 2*, 925-936.
6. Lancaster, C. R. D., and Michel, H. (1997) *Structure 5*, 1339-1359.
7. Abresch, E. C., Paddock, M. L., Stowell, M. H. B., McPhillips, T. M., Axelrod, H. L., Soltis, S. M., Rees, D. C., Okamura, M. Y., and Feher, G. (1998) *Photosynth. Res. 55*, 119-125.
8. Fathir, I., Mori, T., Nogi, T., Kobayashi, M., Miki, K., and Nozawa, T. (2001) *Eur. J. Biochem. 268*, 2652-2657.
9. Ginet, N., and Lavergne, J. (2001) *Biochemistry 40*, 1812-1823.
10. Tandori, J., Hideg, E., Nagy, L., Maroti, P., and Vass, I. (2001) *Photosynth. Res. 70*, 175-184.
11. van der Est, A., Bittl, R., Abresch, E. C., Lubitz, W., and Stehlik, D. (1993) *Chem. Phys. Lett. 212*, 561-568.

12. Isaacson, R. A., Lendzian, F., Abresch, E. C., Lubitz, W., and Feher, G. (1995) *Biophys. J. 69*, 311-322.

13. Brudler, R., Degroot, H. J. M., Vanliemt, W. B. S., Steggerda, W. F., Esmeijer, R., Gast, P., Hoff, A. J., Lugtenburg, J., and Gerwert, K. (1994) *Embo J. 13*, 5523-5530.

14. Breton, J., Boullais, C., Burie, J. R., Nabedryk, E., and Mioskowski, C. (1994) *Biochemistry 33*, 14378-14386.

15. Allen, J. P. (1994) *Proteins 20*, 283-286.

16. Otwinowski, Z., and Minor, W. (1997) *Methods Enzymol. 276*, 307-326.

17. Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) *Acta Crystallogr. Sect. D Biol. Crystallogr. 54*, 905-921.

18. Read, R. J. (1986) *Acta Crystallogr. Sect. A 42*, 140-149.

19. Jones, T. A., Zou, J. Y., Cowan, S. W., and Kjeldgaard, M. (1991) *Acta Crystallogr. Sect. A 47*, 110-119.

20. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) *J. Appl. Crystallogr. 26*, 283-291.

Table 5.1. Summary of data collection and phasing statistics for stigmatellin-bound RCs in the dark (DQ_AS) and light ($D^+Q_A^-S$) adapted states.

	DQ_AS	$D^+Q_A^-S$
Wavelength (Å)	1.08	1.033
Space group	P4(3)2(1)2	P4(3)2(1)2
Unit cell dimensions (Å)		
a	140.1	141.0
b	140.1	141.0
c	273.8	276.3
Maximum resolution (Å)	2.2	2.65
Total reflections	402,872	456,667
Unique reflections	130,987	74,424
Completeness (%) [a]	95.6 (85.6)	91.8 (79.7)
$I/\sigma(I)$	9.1 (1.7)	16.7 (3.7)
R_{sym} (%) [b]	10.5 (30.0)	8.2 (24.9)

[a] Numbers in parentheses correspond to values in the highest resolution shell

[b] $R_{sym} = (\Sigma_{hkl} \Sigma_i | I_i (hkl) - < I (hkl) > |) / (\Sigma_{hkl} \Sigma_i I (hkl))$

Table 5.2. Final refinement statistics for models of stigmatellin-bound RC in the dark (DQ_AS) and light ($D^+Q_A^-S$) adapted states.

	DQ_AS	$D^+Q_A^-S$
Resolution limits (Å)	48.8-2.2	49.4-2.66
R-factor [a]	20.9	20.8
R-free	25.1	27.3
No. of nonhydrogen atoms in a.u.		
Protein	13,014	13,042
Cofactor	986	888
Water	538	185
Detergent	480	96
RMS deviations from ideal values		
Bond lengths (Å)	0.01	0.02
Bond angles (deg.)	1.62	2.11
Dihedral angles (deg.)	21.4	22.1
Improper angles (deg.)	0.98	1.32
Average temperature factor ($Å^2$)		
Protein	38.6	49.4
Cofactor	46.2	55.7
Water	41.5	50.3
Detergent	79.1	81.9
Ramachandran plot [b]		
residues in most favored regions (%)	91.2	89.3
residues in additional allowed regions	8.2	10.0
residues in generously allowed regions	0.6	0.7
residues in disallowed regions (%)	0.0	0.0

[a] R-factor = $\Sigma(|F_{obs}|-|F_{calc}|) / \Sigma|F_{obs}|$

[b] as determined by PROCHECK (*20*)

Table 5.3. Average distances of potential hydrogen bonds formed between the protein environment and Q_A, Q_A^-, Q_B, Q_B^-, and/or stigmatellin in native RC structures in the dark- and light-adapted states (DQ_AQ_B and $D^+Q_AQ_B^-$; PDB entries 1AIJ and 1AIG) and stigmatellin-bound RC structures in the dark- and light-adapted states (DQ_AS and $D^+Q_A^-S$). D (donor), A (acceptor), with separation distance measured in Å.

D	A	DQ_AS	$D^+Q_A^-S$	DQ_AQ_B	$D^+Q_AQ_B^-$
Ala M260 N	Q_A/Q_A^- O1	2.67	2.73	2.78	2.82
His M219 Nδ1	Q_A/Q_A^- O4	2.71	2.72	2.85	2.89
His L190 Nδ1	STG O4	2.69	2.86	--	--
His L190 Nδ1	STG O5	3.15	3.23	--	--
Asp L213 Oδ1	STG O7	3.10	3.07	--	--
Ser L223 Oγ	STG O7	3.09	3.50	--	--
Ser L223 Oγ	STG O8	2.59	2.41	--	--
Ile L224 N	STG O8	3.00	2.78	--	--
Gly L225 N	STG O8	3.37	2.86	--	--
Ile L224 N	Q_B O4	--	--	2.99	--
His L190 Nδ1	Q_B^- O4	--	--	--	2.73
Ser L223 Oγ	Q_B^- O1	--	--	--	3.13
Ile L224 N	Q_B^- O1	--	--	--	2.86
Gly L225 N	Q_B^- O1	--	--	--	3.24
Gly L225 N	Q_B^- O2	--	--	--	3.30

Figure 5.1. Chemical structures, with numbering of (a) ubiquinone-10 and (b) stigmatellin.

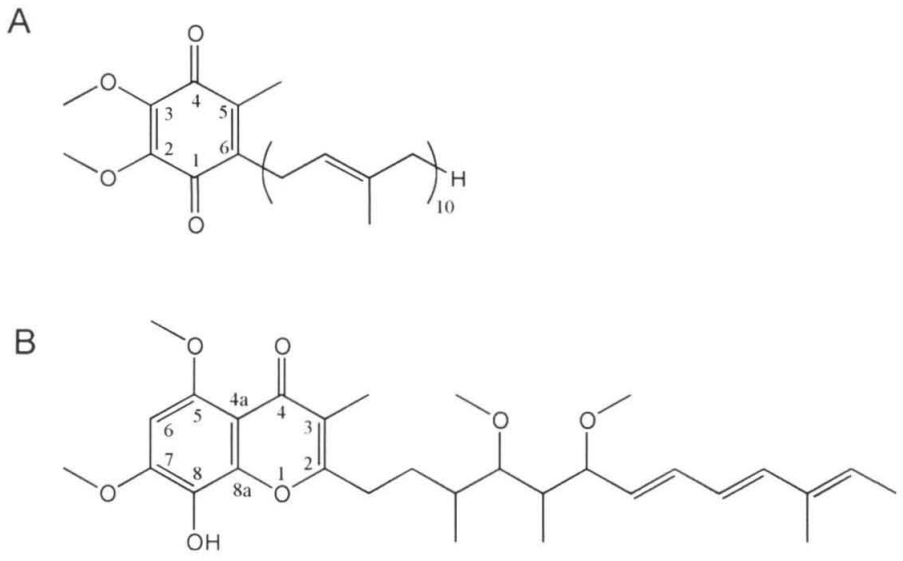


Figure 5.2. Overall structure of stigmatellin-bound RC in the dark-adapted state (DQ_AS), comprising the L (blue), M (yellow), and H (green) subunits, and the two branches of cofactors: bacteriochlorophylls (purple), bacteriopheophytins (cyan), and primary quinone Q_A (magenta). The tails of the bacteriochlorphyll and bacteriopheophytin cofactors have been omitted from this figure for clarity. The secondary quinone (Q_B) is replaced by the herbicide stigmatellin (red). The A branch, through which electron transfer occurs, is on the right side in this figure.

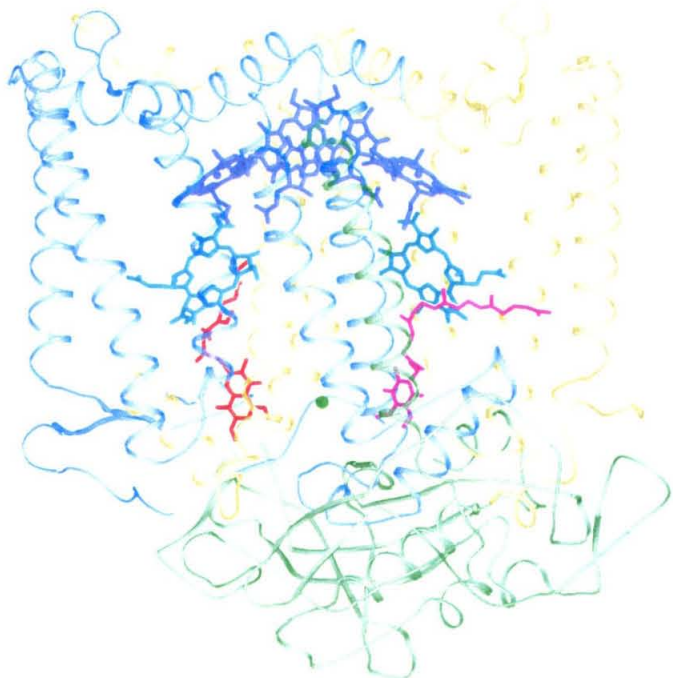


Figure 5.3. Superposition of the *Rb. sphaeroides* (yellow) and *Bcl. viridis* (cyan) stigmatellin-bound RC structures reveals stigmatellin to be bound in a similar mode at the proximal site of the Q_B pocket. Simulated annealing $2|F_o|-|F_c|$ electron density contoured at 1.0 σ level for the stigmatellin in the DQ_AS structure is depicted as a yellow mesh. Residue numbers correspond to those of *Rb. sphaeroides* RC.

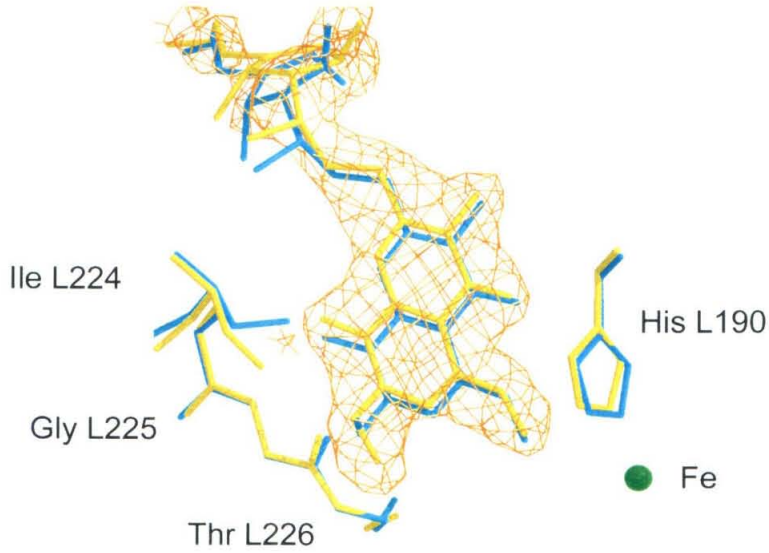


Figure 5.4. (a) Superposition of the DQ_AS (yellow), DQ_AQ_B (magenta; PDB entry 1AIJ), and $D^+Q_AQ_B^-$ (cyan; PDB entry 1AIG) structures shows the binding mode of stigmatellin is more similar to that of Q_B^- than of Q_B. In the DQ_AQ_B structure, four water molecules (magenta spheres) occupy the site at which stigmatellin and Q_B^- bind in their respective structures. Potential hydrogen bonds formed between stigmatellin and the protein environment are depicted by dashed lines. (b) another view showing the slight shift in several surrounding residues (e.g., Phe L216 and Ser L223) caused by stigmatellin binding. Color scheme is the same as in panel (a).

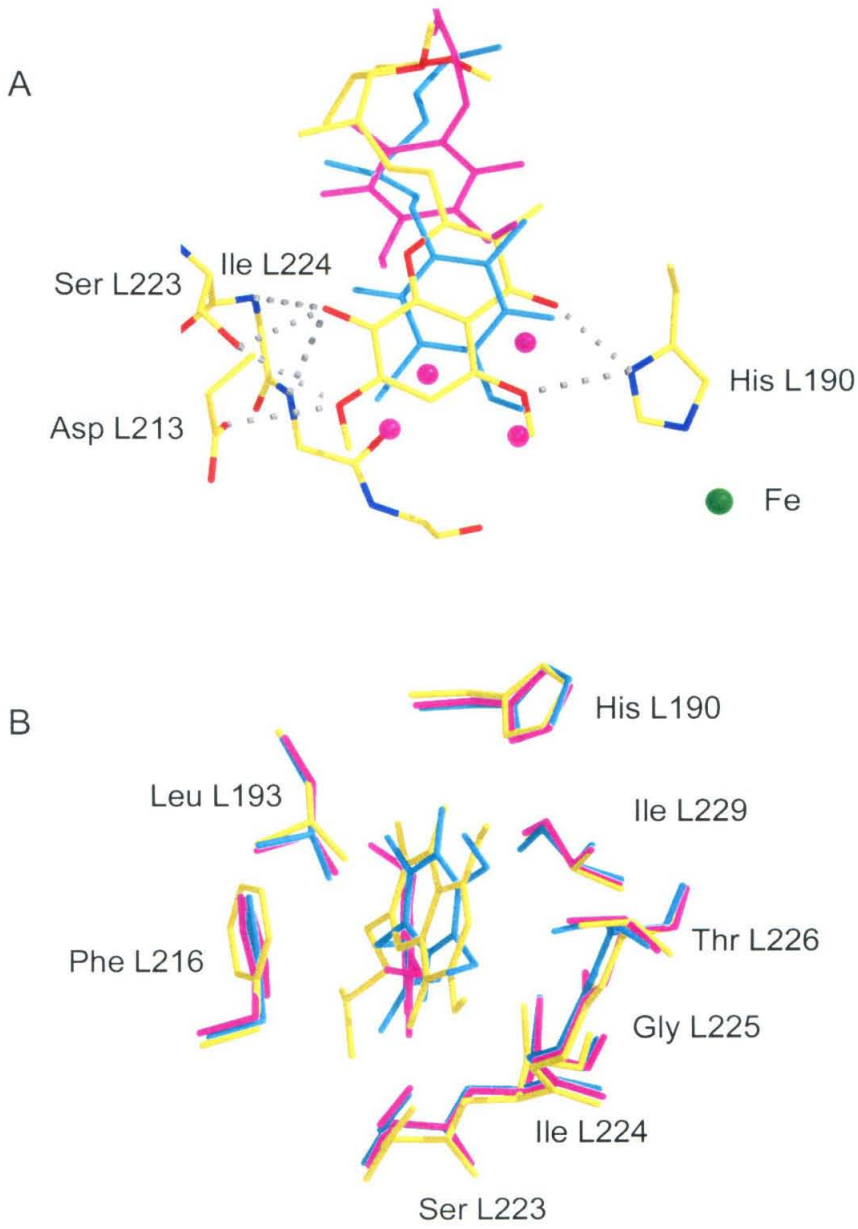


Figure 5.5. Overall structure of stigmatellin-bound RC in the dark-adapted state (DQ_AS) showing the L (blue), M (yellow), and H (green) subunits and bound LDAO detergent molecules (cyan and magenta). Detergents in magenta color are those that bind in clefts that are related by the internal pseudo-twofold symmetry axis, as described in the text. RC cofactors have been omitted from the figure for clarity.

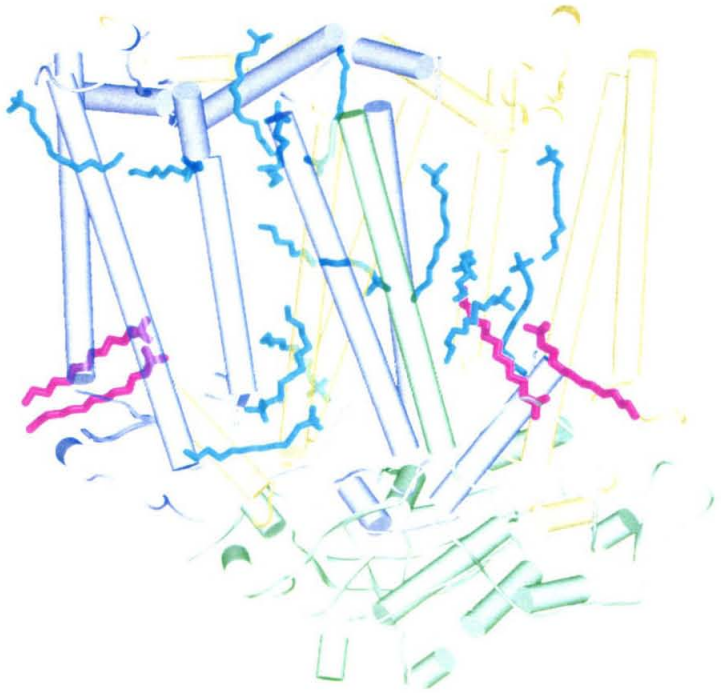


Figure 5.6. View of the Q_A binding pocket showing the structure of Q_A^- (gray) and several surrounding M subunit (blue) residues which appeared to be slightly disordered based on simulated annealing $2|F_o|-|F_c|$ electron density contoured at 1.0 σ level (green mesh). L and H subunits are colored yellow and green, respectively.

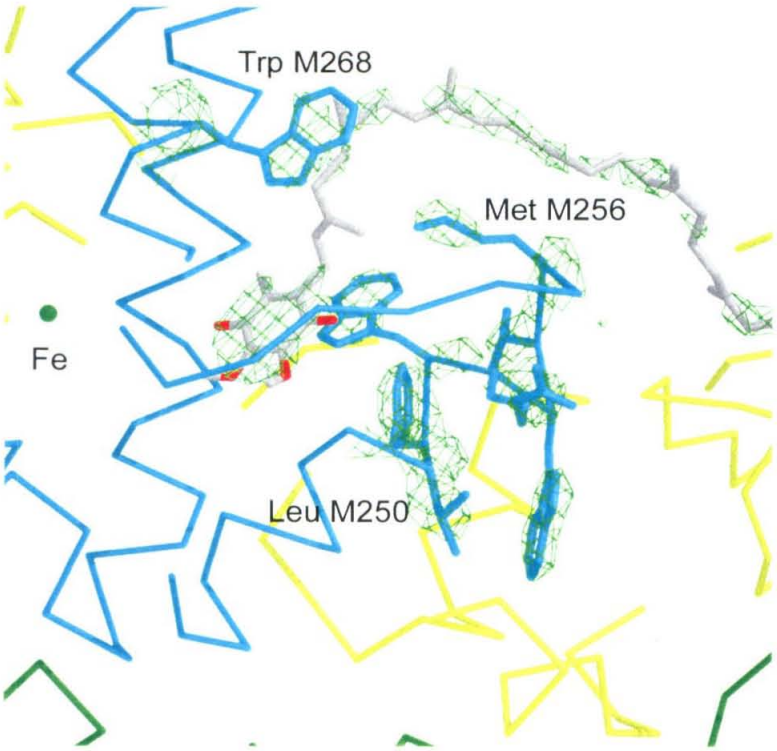
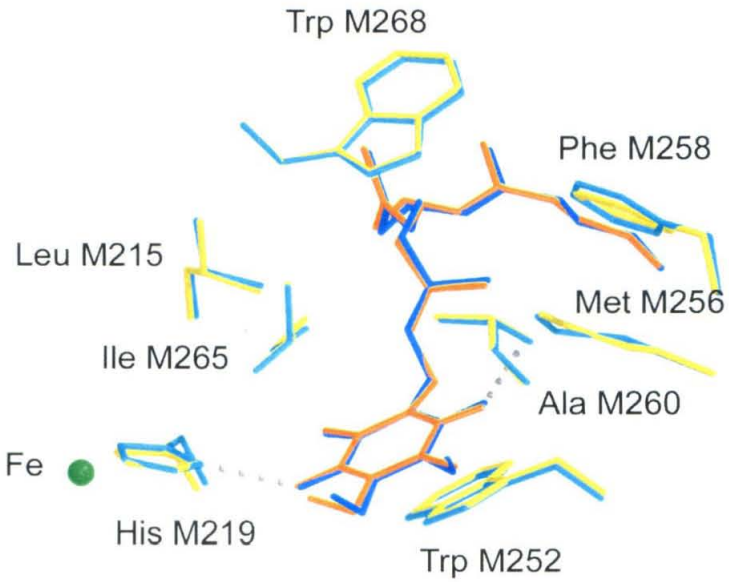


Figure 5.7. Superposition of the DQ_AS (yellow) and $D^+Q_A^-S$ (blue) structures showing the similarity of Q_A (darker shading) and surrounding residues. Hydrogen bonds formed between Q_A/Q_A^- and the protein environment are depicted by dashed lines.



CHAPTER 6

X-ray Structure Determination of the Cytochrome c_2:Reaction Center Electron Transfer Complex from *Rhodobacter sphaeroides* *

* Adapted from Axelrod, H. L., Abresch, E. C., Okamura, M. Y., Yeh, A. P., Rees D. C. and Feher, G. (2002) *J. Mol. Biol. 319*, 501-515.

X-ray Structure Determination of the Cytochrome c_2:Reaction Center Electron Transfer Complex from *Rhodobacter sphaeroides*

Herbert L. Axelrod[§], Edward C. Abresch[§], Melvin Y. Okamura[§], Andrew P. Yeh[⊥], Douglas C. Rees[⊥,‡], and George Feher[§*]

[§] Department of Physics, 9500 Gilman Drive, University of California, San Diego, La Jolla, CA 92093-0319

[⊥] Division of Chemistry and Chemical Engineering, 147-75CH, California Institute of Technology, Pasadena, California 91125

[‡] Howard Hughes Medical Institute, California Institute of Technology, Pasadena, California 91125

[*] To whom correspondence should be addressed

Phone: 858-534-4389, Fax 858-822-0007, E-mail: gfeher@ucsd.edu

Abstract

In the photosynthetic bacterium *Rhodobacter sphaeroides*, a water soluble cytochrome c_2 (cyt c_2) is the electron donor to the reaction center (RC), the membrane-bound pigment-protein complex that is the site of the primary light-induced electron transfer. To determine the interactions important for docking and electron transfer within the transiently bound complex of the two proteins, RC and cyt c_2 were co-crystallized in two monoclinic crystal forms. Cyt c_2 reduces the photo-oxidized RC donor (D^+), a bacteriochlorophyll dimer, in the co-crystals in ~0.9 μsec, which is the same time as measured in solution. This provides strong evidence that the structure of the complex in the region of electron transfer is the same in the crystal and in solution. X-ray diffraction data were collected from co-crystals to a maximum resolution of 2.40 Å and refined to an R-factor of 22% (R_{free}=26%). The structure shows the cyt c_2 to be positioned at the center of the periplasmic surface of the RC, with the heme edge located above the bacteriochlorophyll dimer. The distance between the closest atoms of the two cofactors is 8.4 Å. The side chain of Tyr L162 makes van der Waals contacts with both cofactors along the shortest intermolecular electron transfer pathway. The binding interface can be divided into two domains: (i) A short-range interaction domain that includes Tyr L162, and groups exhibiting nonpolar interactions, hydrogen bonding, and a cation-π interaction. This domain contributes to the strength and specificity of cyt c_2 binding. (ii) A long-range, electrostatic interaction domain which contains solvated complementary charges on the RC and cyt c_2. This domain, in addition to contributing to the binding, may help steer the unbound proteins toward the right conformation.

Introduction

Intermolecular electron transfer is a fundamental process required for energy conversion in biological systems (*1-3*). In photosynthetic bacteria, a membrane-bound 101 kDa pigment-protein complex called the reaction center (RC) interacts with a 14 kDa water-soluble cytochrome c_2 (cyt c_2) in an electron transfer cycle that converts light energy into chemical energy (*4-7*). The RC contains three subunits (L, M and H) as well as four bacteriochlorophyll molecules, two bacteriopheophytins, two ubiquinones and a non-heme Fe atom. Cyt c_2 contains a prosthetic heme group that is covalently bonded to two cysteines through thioether linkages. The X-ray crystal structures of the RC (*8-11*) and cyt c_2 (*12*) from *Rb. sphaeroides* have been determined.

Light excitation of the RC photo-oxidizes the primary donor, D, a specialized bacteriochlorophyll dimer, that initiates sequential electron transfer through a series of electron acceptors: bacteriochlorophyll, bacteriopheophytin, and two ubiquinone molecules Q_A and Q_B. Upon double reduction, Q_B binds two protons, dissociates from the RC and transfers electrons to the cyt bc_1 complex (reviewed in ref. (*13*)). The resultant proton gradient generated across the membrane drives ATP synthesis. Reduced cyt c_2 carries electrons from the cyt bc_1 complex back to the RC to reduce the oxidized donor, D^+. Thus, cyt c_2 acts as an electron shuttle between the RC and cyt bc_1, completing the electron transfer cycle. For cyt c_2 to function effectively in this cycle, its association and dissociation from the RC must be rapid, and in the bound cyt c_2:RC complex it must promote rapid electron transfer to D^+.

The interactions between cyt c_2 and RC have been studied extensively (reviewed in ref. (*14*)). The cyt c_2 and RC form a transient complex with a dissociation constant $K_D \cong 10^{-6}$ M at low ionic strength (10 mM) (*15-18*). The observed electron transfer rate consists of two main phases: a first-order ($k_{ET} \cong 10^6$ s^{-1}) phase attributed to cyt c_2 bound to the RC, and a slower, second-order phase due to RCs reacting with unbound cyt c_2 (*17-20*). The first-order rate depends on the driving force (i.e., the redox potential of D^+/D) (*21, 22*) and represents the intrinsic rate of electron tunneling from cyt c_2 to D^+. The second-order rate constant is independent of the driving force and is associated with the dynamics of cyt c_2 docking. Its value at low ionic strength is 10^9 M^{-1} sec^{-1}.

The second-order rate constant depends on ionic strength, indicating an involvement of electrostatic interactions in the association of the two proteins (*16-18*). This interaction takes place between a cluster of positively charged residues around the solvent-exposed edge of the cytochrome heme (*12, 18*) and a complementary cluster of negatively charged residues on the periplasmic surface of the RC (*8-11, 18, 23*). The involvement of specific Lys residues of the cyt c_2 and acidic residues on the periplasmic surface of the RC have been shown by chemical modification (*24, 25*), site directed mutagenesis experiments (*20, 26-28*), and chemical cross-linking experiments (*19, 29, 30*).

Different docking models have been proposed for the cyt c_2:RC complex (*23, 27, 31*), based on the formation of complementary ion pairs between the two proteins and on general features of the homologous RC from *Blastochloris viridis* (formerly called

Rhodopseudomonas viridis), which contains a non-dissociating tetraheme cytochrome that is attached to the periplasmic surface of the RC (*32*). However, the structure of the cyt c_2:RC complex from *Rb. sphaeroides* had not been established definitively up to now.

In this work, we report the X-ray crystal structure of the cyt c_2:RC complex from *Rb. sphaeroides*. This study is an extension of earlier work by Adir *et al.* (*27*) in which co-crystallization of the cyt c_2:RC complex was first reported. However, in that study, the occupancy of the cyt binding site in the co-crystal was low (0.25 cyt per RC) and consequently it was not possible to determine a reliable structure of the complex. In this work, new conditions for crystallization were established to improve the occupancy of the cyt c_2 binding in the co-crystal. The enhanced occupancy and improved X-ray diffraction resolution from these new co-crystals have enabled us to determine the structure of the cyt c_2:RC complex, and to establish the protein-protein interactions that govern the binding and the electron transfer processes. The structures of several co-crystals of water-soluble electron transfer complexes have been previously reported (*33-37*). However, the structure reported here is of a complex between an integral membrane protein and its exogenous, water-soluble, electron transfer partner. A preliminary account of this work has been presented elsewhere (*38, 39*).

Results

Electron transfer kinetics in co-crystals

The electron transfer kinetics between cyt c_2 and the photo-oxidized donor were measured at 550 nm. The results on a type I co-crystal are shown in Figure 6.1. The upper

trace shows the absorbance changes following a laser flash in the presence of oxidized cyt c_2, corresponding to the reaction:

$$\text{cyt } c_2^{3+}\text{:RC}(DQ_A) \xrightarrow{h\nu} \text{cyt } c_2^{3+}\text{:RC}(D^+Q_A^-) \qquad (1)$$

In this case, cyt c_2^{3+} cannot reduce D^+, and the absorption change is due to the formation of $D^+Q_A^-$. * Note that the charge recombination $D^+Q_A^- \rightarrow DQ_A$ is slow (~100 ms) and does not occur on the time scale of Figure 6.1. When cyt c_2 is reduced prior to the laser flash, it can transfer an electron to D^+ according to the reaction:

$$\text{cyt } c_2^{2+}\text{:RC}(DQ_A) \xrightarrow{h\nu} \text{cyt } c_2^{2+}\text{:RC}(D^+Q_A^-) \xrightarrow{k_{ET}} \text{cyt } c_2^{3+}\text{:RC}(DQ_A^-) \qquad (2)$$

The observed absorbance changes (lower trace in Figure 6.1) are due to the oxidation of cyt c_2 and the concomitant reduction of D^+. Similar results were obtained on type II co-crystals. From the exponential decay of the trace obtained from several co-crystals, the electron transfer time was determined to be

$$\tau_{ET} = 1/k_{ET} = 0.9 \pm 0.1 \mu s \qquad (3)$$

This value of τ_{ET} is the same as that observed in solution, showing that the complex is functionally equivalent in both the co-crystal and in solution. The question of the fraction of RCs in which electron transfer occurs (i.e., that have an active cyt c_2 molecule attached) is discussed in the next section.

---

* Electron density for Q_B is not observed in the co-crystals indicating a low occupancy of the secondary quinone acceptor, Q_B.

The stoichiometry of cyt c_2 and RC in the co-crystal

The molar ratio cyt c_2:RC in the co-crystal was determined by three independent methods.

In the first method, several batches of co-crystals were dissolved, and the concentrations of cyt c_2 and RC were determined spectrophotometrically using the known and measured extinction coefficients (see Materials and Methods). The value of the molar ratio obtained was cyt c_2/RC = 0.9 ($\pm$ 0.1).

The second method utilized the amplitudes of the absorbance changes in the co-crystal following a laser flash. In principle, the concentrations of cyt c_2 and RC in the co-crystal can be obtained from the absorption changes associated with their oxidation and reduction as seen in Figure 6.1. Since we do not know the required extinction coefficients in the crystal, this method is not feasible. However, if we monitor a wavelength where only RC absorbance changes are observed, we can determine the cytochrome occupancy without knowledge of the extinction coefficients. At a monitoring wavelength of 600 nm, the predominant absorption is due to the creation (and decay) of $D^+Q_A^-$. Only those RC's that have a bound cyt c_2^{2+} will be reduced following a laser flash (equation (2)). Thus, from the ratio of the amplitude of the absorption changes immediately following a laser flash (corresponding to $[D^+Q_A^-]$) to the amplitude after reduction by cyt c_2^{2+} (corresponding to $[DQ_A^-]$ = [cyt c_2]), the molar ratio of cyt c_2 to RC can be determined

without having to know the extinction coefficients*. Using this method (data not shown), we estimated an active molar fraction of cyt c_2:RC $\cong$ 0.8 ($\pm$ 0.1).

In the third method, we used the program MAPMAN (*40*) to integrate the difference electron density peaks in type II co-crystals surrounding the heme iron atom on cyt c_2 and the iron atom on the RC, with the coordinates of both irons omitted from the refined model. The ratio of the integrated $|F_o|$-$|F_c|$ electron densities surrounding the two metal atoms resulted in a c_2 to RC molar ratio of 1.0 ($\pm$ 0.1). This ratio is, within experimental error, in agreement with the values obtained from the two spectrophotometic determinations described above.

General structural features

The crystal data and refinement statistics obtained on the two co-crystal forms are summarized in Table 6.1. Least-squares overlap of the coordinates of the two crystal forms show the same relative position of the cyt c_2 with respect to the RC. This indicates a common docking structure in both crystal forms, despite some differences in the crystal lattice contacts.

---

* In practice, the contribution of cyt c_2 to the absorption changes cannot be neglected. It contributes ~10% to the absorption changes at 600 nm (using the extinction coefficients in solution), which needs to be corrected for. However, since this is a small contribution, our error in the extinction coefficient does not significantly affect the results. (An error in the extinction coefficients of a factor of two would change the molar ratio by only ~10%.)

The general structural features of the cyt c_2-RC complex are shown in Figure 6.2. The heme iron on cyt c_2 is located ~10 Å above the periplasmic surface of the RC displaced by ~2 Å from the pseudo-twofold symmetry axis relating the L and M subunits, and the two RC cofactor branches (*41-43*). The plane of the cyt c_2 heme is tilted toward the M side of the RC at an angle of ~10° from the membrane normal, and is rotated ~45° about the membrane normal with respect to the plane of the primary donor. The exposed heme edge is positioned directly above the primary donor and is in van der Waals contact with Tyr L162 on the RC. An expanded version of this region is shown, together with the electron density map, in Figure 6.3.

The contact region of the docking interface is a closely packed, solvent-inaccessible region. It is located near the intersection of four non-membrane spanning helices (*23*), *cd* and *e* on the L subunit as well as the *cd* and *e* helices on the M subunit. (Figure 6.2). These helices are sloped away from the interface, resulting in an increasing separation of the surfaces of the RC and cyt c_2 as one proceeds away from the contact point. These loosely packed regions contain many charged residues that are solvent accessible. We now proceed to discuss the interactions of the docking surfaces in more detail.

Electrostatic interactions

Views of the electrostatic interactions between oppositely charged side chains on the two proteins are shown in Figure 6.4. Most of the electrostatic interactions between cyt c_2 and RC are on the M side of the periplasmic surface even though the heme on cyt

c_2 docks near the center of the RC. Most of the charged groups at this interface are separated by more than 5 Å (Table 6.2) and are consequently solvent-accessible (*44*). Indeed, several water molecules were identified in the interface region (discussed below).

Several Lys side chains on cyt c_2 near the methionine ligand to the heme iron are pointing toward periplasmic acidic residues on the M subunit. These lysines, which are conserved in many species, create the region of highest positive electrostatic surface potential on cyt c_2 (*18, 20*). The region of strongest negative electrostatic potential on the RC is centered on the periplasmic surface of the M subunit near Asp M184, which is surrounded by Glu M95 and Glu M173 (*18, 20*). Two regions with the highest (complementary) potentials face each other in the complex; specifically Lys C99 is in the vicinity of Asp L257 and Asp L261, and Lys C103 and Lys C105 are in the vicinity of Asp M184, Glu M95, and Glu M173 (see Figure 6.4. and Table 6.2). The distances between these lysines and acidic side chains are > 4.5 Å, which is significantly larger than expected for an electrostatic salt bridge.

Van der Waals interactions

The interface in the vicinity of the cyt c_2 heme and RC donor is predominantly apolar in nature. Probably, the most important contact for electron transfer (detailed below) is between the CBC methyl group of the cytochrome and the Tyr L162 ring on the RC (see Figure 6.5 and Table 6.2). Other nonpolar contacts, in the vicinity of Tyr L162, are formed between Phe C102 on cyt c_2 and the side chains of Val M192 and Leu M191 on the RC. There are a number of additional close-range interprotein van der Waals

interactions (45) between atoms separated by ≤ 4 Å (Table 6.2). Most of these are in the closely packed contact region, and involve the side chains of Thr C17, Thr C36, and Thr C101, which surround the edge of the heme and Tyr L162.

Interprotein hydrogen bonds

Three potential hydrogen bonds between cyt c_2 and RC were identified (see dotted lines in Figure 6.6, and Table 6.2). Two of the hydrogen bonds involve the amide side chains of Asn M187 and Asn M188 on the M subunit of the RC and one involves an amide side chain Gln L258 on the C-terminal helix of the L subunit. All three amide side chains on the RC interact with polypeptide backbone atoms on cyt c_2 between Lys C99 and Lys C103 near the methionine ligand to the heme iron. These hydrogen bonds are located near the boundary separating a tightly packed region of the docking interface from the more loosely packed, surrounding region.

Cation-π interaction

An additional short-range interaction is provided by the close contact between Arg C32 and Tyr M295 (see Figure 6.4(a), Table 6.2). The distance between the charged guanidinium side chain and the aromatic ring is less than 4 Å, which is characteristic of a cation-π interaction (*46, 47*). The aromatic plane of the tyrosine side chain is parallel to the arginine side chain. The preference for this parallel stacking has been described (*47*). Tyrosine M295 is conserved in several species of photosynthetic reaction centers including *Blc. viridis* and *Rhodobacter capsulatus*; however, Arg C32 is not conserved. In *Rb. capsulatus* and *Paracoccus denitrificans,* this residue is replaced by lysine and in

Blc. viridis and *Rhodospirillum rubrum* the polypeptide segment containing this residue is missing.

Interactions involving water molecules

Two types of water molecules were identified at the docking interface in the Type II crystals: (a) bridging waters that form hydrogen bonds with both the RC and cyt c_2, and (b) water molecules that are hydrogen-bonded to only one of the two proteins. The seven bridging water molecules are located near the boundary separating the closely packed contact area from a more loosely packed region containing charged residues (see above). The bridging water molecules, together with their putative hydrogen bond partners and B-factors, are listed in Table 6.3. In the loosely packed, solvent-accessible region of the interface, there are many resolved water molecules of the second type. Several of these are located within hydrogen bonding distance of the side chains of acidic residues on the surface of the RC.

Conformational changes of the RC in the crystal structure of the complex

There are two types of structural changes that occur in the co-crystal complex. One is brought about by the crystallization process, the other by complex formation.

The first type of structural change of the RC observed in the co-crystal (see the arrow in Figure 6.2), is a kink in the backbone that involves the isomerization of the peptide bond of Pro M49 from the *cis*-configuration in the free RC to the *trans*-configuration in the co-crystal. Pro M49 is located on the cytoplasmic side of the RC near

the end of transmembrane helix A (*23*). This change alters the conformation of all preceding residues towards the N-terminus of the M subunit. Of particular note is the refolding of Pro M34-Gly M43 into an α-helix near the cytoplasmic end of the transmembrane A helix, the amino acid side chains of which form crystal lattice contacts with a cyt c_2 bound to a neighboring RC. The residues involved in this rearrangement are located on the side of the RC opposite from the cyt c_2 docking site.

More localized conformational changes due to the docking of the cytochrome are observed in both the RC and cyt c_2. Least-squares overlap of the docked and undocked RC structures (excluding the first 50 residues at the N-terminal region of M) shows a root mean square displacement of 0.75 Å. A similar overlap for cyt c_2 shows an average displacement of 0.50 Å. Closer examination reveals that the most significant structural change at the docking interface is the position of M295 that participates in the cation-π interaction. The aromatic ring moves ~2 Å toward Arg C32. On cyt c_2, the torsion angle ψ of Thr C101 shifts ~45° to permit formation of a hydrogen bond with the side chain of Asn M187.

Discussion

We have determined the structures of the cyt c_2:RC complex in *Rb. sphaeroides* by X-ray diffraction from two different monoclinic crystal forms. This work is an extension of the earlier work by Adir *et al.* (*27*) who, because of the low cyt c_2 occupancy (~25%) in the co-crystal, were unable to determine the structure accurately. We increased the occupancy to ~90% by lowering the ionic strength. This resulted in a

significantly higher-resolution (2.4 Å) than that reported previously (4.5 Å) (*27*). The higher-resolution enabled us to establish with confidence the position of the bound cytochrome, as well as most of the side chains at the interface, and to determine the locations of several water molecules.

The structure of the cyt c_2:RC complex gives insight into the design features of the molecular machinery that allows it to effectively perform its dual physiological functions of rapid electron transfer from the cyt c_2 to the bacteriochlorophyll dimer with rapid association and dissociation of cyt c_2 from RC that is required for an effective turnover rate of the photochemical cycle (*4, 5, 7, 8*). Before describing the structural features responsible for these properties, we need to address the question whether the crystallization process has affected the co-crystal structure.

Is the structure of the cyt c_2:RC complex in the co-crystal the same as in solution?

This kind of question is often asked of crystallographers and is inherent to any technique that requires crystals. The question is particularly relevant when one deals with a relatively loosely bound complex (i.e., cyt c_2:RC) whose structure can be affected by the requirement to form specific intermolecular contacts in the crystal.

The intermolecular electron transfer kinetics between cyt c_2 and D^+, k_{ET}, is expected to depend sensitively on the structure, i.e., the distance between the co-factors through the relation (*48-50*):

$$k_{ET} \propto e^{-\beta r} \qquad (4)$$

where r is the distance between the cofactors, and β is a constant estimated theoretically (*48*) and empirically by measuring the intramolecular electron transfer rate, k_{ET}, in a series of proteins (*50, 51*). Moser *et al.* (*50*) arrived at a value of β=1.4 $Å^{-1}$, and Tezcan *et al.* (*51*) at a value of 1.1 $Å^{-1}$. We have measured the interprotein electron transfer rate k_{ET} in the cyt c_2:RC co-crystal and found it to be, within experimental error (~10%), the same as in solution. Using an average value of β=1.25 $Å^{-1}$, equation (4) predicts that a 10% change in k_{ET} (our experimental uncertainty) is brought about by a ~0.1 Å change in r, which is below the accuracy of the position of the X-ray coordinates. This result provides strong evidence that crystallization does not alter the structure of the complex in the vicinity of the docking surface where electron transfer occurs. Additional, corroborative, evidence is provided by the fact that in the two different crystal forms (Table 6.1), each with a different set of crystal contacts, k_{ET} was found to be the same. Although some structural changes were observed due to the crystallization process (see Results), these were far removed from the docking surface.

The electron transfer pathway

The structure of the cyt c_2:RC complex in the region between the heme and the primary donor provides information on the likely route of the electron transfer path. The distance between the closest atoms on the two cofactors in the complex, the CBC methyl group on the heme and the ethyl group on either of the bacteriochlorophyll rings II of the dimer. (Figure 6.5) is 8.4 Å. This distance is bridged by the aromatic side chain of Tyr L162, which makes van der Waals contacts to both cofactors. We therefore postulate that

the most likely electron transfer pathway proceeds from the heme edge via Tyr L162 to the bacteriochlorophyll dimer.

The RC from the related photosynthetic bacterium *Blc. viridis* contains a non-dissociating tetraheme cytochrome molecule. The distance between closest atoms on the heme and donor is 8.7 Å (*32*), similar to that in *Rb. sphaeroides*. In both species, a tyrosine side chain is in van der Waals contact with the cofactors (*32, 52, 53*). Superimposing the coordinates of the primary donors of the two species results in a displacement of 1.5 Å between the heme irons. The electron transfer between the cytochrome and primary donor is about three times faster in *Blc. viridis* than in *Rb. sphaeroides* (*54, 55*). This may be either due to a difference in the reorganization energy λ (which has not been determined for *Blc. viridis*) or a different path length (the distance between conjugated atoms on the heme and primary donor is 12.3 Å in *Blc. viridis* compared to 14.2 Å in *Rb. sphaeroides*).

The role of Tyr L162 has been investigated by site directed mutagenesis. In *Rb. sphaeroides*, mutations of Tyr L162 to the non-aromatic residues Met, Leu, Gly, and Ser resulted in a ~100 fold reduction in the rate of electron transfer (*52, 53*). In contrast, the analogous mutations of Tyr L162 in *Blc. viridis* had only a small effect on the electron transfer rate (typically a factor of 2) (*56*). How can we reconcile these disparate results? In *Blc. viridis*, the cytochrome is permanently bound to the RC, keeping the distance between the heme and bacteriochlorophyll dimer the same in the native and mutant structures. A different bridging residue can apparently take over the role of Tyr L162 in

the electron transfer. A possible explanation for the different result in the *Rb. sphaeroides* complex is a change in the position of the cyt c_2 on the mutant RC surface. Using β=1.25 $Å^{-1}$, equation (4) predicts for an increase in distance of 3.7 Å a ~100 fold reduction in k_{ET}. We conclude, therefore, that Tyr L162 in *Rb. sphaeroides* is likely to play an important role in attaining the optimum docking position of the cyt c_2.

Let us now compare the structural results and the measured value of k_{ET} with theoretical predictions. Aquino *et al.* (*57*) have made pathway calculations for different positions of the cytochrome on the RC surface. They concluded that the best pathway for electron transfer is through Tyr L162, in accord with the structure presented in this work. These authors, however, calculated only relative electron transfer rates for different pathways. To obtain absolute values, we use the empirical relation of Moser *et al.* (*50*) (β=1.4 $Å^{-1}$) and correct for the Frank-Condon factor (*49*) using a reorganization energy λ=0.96 eV and ΔG= -0.160eV (*22*). For the 14.2 Å distance between conjugated ring atoms in the cyt c_2:RC complex, we expect the measured rate to be $k_{ET}=8 \times 10^3$ s^{-1} and 2×10^5 s^{-1} for β=1.4 $Å^{-1}$ and β=1.1 $Å^{-1}$, respectively (*50, 51*)*. Taking these results at face value, one concludes that the intermolecular electron transfer rate in the cyt c_2:RC complex ($k_{ET} \cong 10^6$ s^{-1}) is larger than expected for an average protein. Perhaps a more realistic conclusion is that using an average value of β is too simplistic an assumption and

* The value of β=1.4 $Å^{-1}$ is based on distances between conjugated bonds on the cofactors (*50, 58*), whereas β=1.1 $Å^{-1}$ is based on distances between redox-active metal atoms in metalloproteins (*51*). An analogous metal-to-metal distance for the *Rb. sphaeroides* cyt c_2:RC complex cannot be defined, since the Mg^{2+} atoms on the donor do not change redox state during the electron transfer.

a more sophisticated theory that takes into account the nature of the specific bonds involved is required (*59, 60*). The structural results of this work should provide a useful model against which theories can be tested.

The two-domain docking model

The interface between cyt c_2 and the RC can be divided into two distinct, spatially separated domains. There is a central region, which encompasses the contact point between the heme edge and Tyr L162. In addition, this region contains the Arg C32-Tyr M295 cation-π interaction as well as several residues on the cyt c_2 and RC that make van der Waals contact with each other (see Table 6.2). This region contributes to the strength and specificity of the binding of the cyt c_2 through short-range interactions (Figure 6.7).

Surrounding the short-range interaction domain is the electrostatic domain, which contains oppositely charged residues on cyt c_2 and on RC, respectively. The electrostatic domain provides the long-range interactions (Figure 6.7) that play a role in the dynamics of the docking process (*20, 28*).

The interactions in the electrostatic domain arise from the cluster of acidic residues surrounding Asp M184 on the M side of the RC surface facing oppositely charged residues on the cytochrome molecule. This is in agreement with mutational results, which show that the most important interactions are between Asp M184 and Lys C103 and between Asp L261 and Lys C99 (*28*). The structure shows that there is ample

space to solvate the charges. The solvation has important functional implications: it can enhance the rates of association of reduced cyt c_2 and dissociation of oxidized cyt c_2, ensuring that cyt c_2 turnover is not the bottleneck in the photosynthetic electron-transfer cycle.

The two-domain docking model for the cyt c_2:RC complex incorporates both the functional advantages of the short-range interaction region in which cyt c_2 is specifically oriented for rapid electron transfer and an electrostatic long-range interaction domain, which steers the unbound proteins towards the right configuration in the bound complex.

Previously proposed structural models

Three models for the docked complex between RC and cyt c_2 from *Rhodobacter sphaeroides* have been proposed (*23, 27, 31*). Allen *et al.* (*23*) positioned the heme over Tyr L162 through formation of a series of complementary salt bridges between the two molecules. Tiede *et al.* (*31*) positioned the cytochrome molecule toward the M side of the RC, while Adir *et al.* (*27*) located the cytochrome heme over Asp M184 on the M side. All three models show the major electrostatic interactions to be on the M side of the RC. The co-crystal structure described in this work is most similar to the docking model proposed by Allen *et al.* (*23*) These authors used the structure of the homologous cyt c_2 from *R. rubrum* since the X-ray crystal structure of cyt c_2 from *Rb. sphaeroides* had not been determined at that time. In their model Asp M184 was positioned close to C94 of the *R. rubrum* cyt c_2, which is equivalent to Lys C103 in cyt c_2 from *Rb. sphaeroides.* In the modeling of Allen *et al.*, short-range interactions were not considered explicitly.

Similarily, in the Tiede and Adir structures, the predominant weight was given to electrostatic interactions. Differences between the co-crystal structure presented here and the other models based predominantly on electrostatics suggest that in addition to electrostatic interactions, short-range interactions play an important role in the positioning of the cyt c_2 molecule.

Structures of other electron transfer complexes involving soluble cytochromes

The cyt c_2:RC co-crystal structure bears some similarities to the X-ray structures of other complexes that contain a transiently docked cytochrome (*33, 34*). For example, the cyt c:cyt c peroxidase complex, whose X-ray structure has been solved by Pelletier and Kraut (*33*), has a small hydrophobic contact region consisting of the exposed heme edge on the cytochrome and two Ala residues on the peroxidase that are connected by covalent bonds to the electron acceptor, an oxidized tryptophan radical. The structure contains relatively few (two) salt bridges. The cyt c_2:RC complex is also similar to the structure proposed for the cytochrome c: cytochrome oxidase complex, based on mutagenesis results (*61-63*) and electrostatic modeling (*64*). In the proposed structure of this complex, the heme edge contacts a Trp residue located in contact with the Cu_A center, the electron acceptor, at the center of a highly charged region on the surface of the oxidase, facing an oppositely charged surface on the cytochrome molecule.

The complexes described above share with the cyt c_2:RC complex the common motif of a small contact region that brings the cofactors into close contact for rapid electron transfer and of oppositely charged surfaces on the cofactors that electrostatically

steer the mobile cytochrome molecule towards the correct docking position on the surface of its partner. Further analyses may reveal that such a docking mode is prevalent among other systems that involve electron transfer between donor and acceptor proteins.

Materials and Methods

Protein purification and characterization

RCs from the carotenoidless mutant of *Rhodobacter sphaeroides* strain R26 were purified in a buffer containing 15 mM Tris-HCl at pH 8.0 in the presence of the detergent lauryldimethylamine-N-oxide (LDAO, Fluka) and 0.1 mM EDTA as described previously (*65*). The purity of the RC was monitored by the optical absorption ratio, A_{280}/A_{802}, which for the purified RC was 1.20. Cytochrome c_2 from the R26 strain of *Rb. sphaeroides* was purified as described previously (*12*). Following the final purification step, cyt c_2 was $\geq$ 90% in the reduced form. Purity was monitored by the optical absorption ratio, A_{280}/A_{417}, which in the fully reduced state was 0.25 (*66*). The purity of the two proteins was additionally assessed by SDS-PAGE as described in ref. (*27*). Prior to crystallization, the purified RC and cyt c_2 were dialyzed against 10 mM Tricine (pH 8.5), 0.025% LDAO, 0.1 mM EDTA., and concentrated; the RCs to ~12 mg/ml with Centricon 100 filters (Amicon) and cyt c_2 to ~20 mg/ml with Centricon-30 filters.

Crystallization

Co-crystals were obtained by vapor diffusion at 19° C in 20 μl sitting drops in 1 *ml* Cryschem type plates (Charles Supper Co. Natick, MA). The co-crystallization conditions were as described by Adir *et al.* (*27*) except that no NaCl was added either to

the crystallization solution or to the outer reservoir. The crystallization solutions contained 70 μM RC (7.7 mg/ml), 140 μM (2 mg/ml) reduced cyt c_2, 10% (w/v) PEG 4000, 0.06% (w/v) LDAO, 3.9%(w/v) heptanetriol, and 15 mM Tricine (pH 8.5). The reservoirs contained 22% (w/v) PEG 4000, 50 mM Tricine (pH 8.5). Co-crystals were observed within two to four weeks.

Determination of the cyt c_2-RC stoichiometry in the co-crystal

The cyt c_2/RC ratio in the co-crystal was determined spectrophotometrically on small co-crystals (typical dimensions ~50 x 50 x 50 μm) that were washed several times in an artificial mother liquor (AML) consisting of 22% PEG 4000, 20 mM Tricine (pH 8.5), 0.1% LDAO, 3.9% w/v heptanetriol. The washed crystals were centrifuged in a 1 ml Eppendorf tube, and dissolved in 50 μl of buffer containing 20 mM Tricine (pH 8.5), 0.1% LDAO, 0.1 mM EDTA. Absorption spectra were obtained in 100 μl quartz microcuvettes using a Cary 50 dual beam spectrophotometer (Varian). The RC concentration was determined from the absorbance at 802 nm using an extinction coefficient ε_{802}=2.88 x 10^5 $M^{-1}cm^{-1}$ (*67*). The cyt c_2 concentration was obtained from the reduced minus oxidized difference spectra at 550 nm, which we determined to be $\Delta\varepsilon^{550}$=21.5 (± 0.2) $mM^{-1}cm^{-1}$. To obtain this value, we used the reported extinction coefficient, ε_{550}=30.8 $mM^{-1}cm^{-1}$, for the reduced state (*66*). The oxidized spectrum was obtained on dissolved co-crystals by adding potassium ferricyanide to a final concentration of 60 μM, while the reduced spectrum was recorded after subsequent addition of sodium ascorbate to a final concentration of 1 mM. Two additional methods that were used to determine the stoichiometry are discussed in Results.

Measurement of electron transfer kinetics in co-crystals

The electron transfer kinetics were measured in a microspectrophotometer of local design as described (*27, 68*), but with the following modifications: a laser flash, producing the charge-separated $D^+Q_A^-$ was delivered by a YAG laser (λ=700 nm, 5 mJ, 20 ns pulse width) (Opotek, Carlsbad, CA) via a light guide to the crystal oriented perpendicularly to the monitoring light. Kinetic data were obtained at two wavelengths of the monitoring beam, 550 nm and 600 nm. To reduce laser flash artifacts, two SWP 560 filters and one SWP 620 filter (Linos Photonics) were placed in front of the phototube for measurements at 550 nm, and a 600 nm interference filter (Corion) and a SWP620 filter for measurements at 600nm. The response time of the system was 3×10^{-8} seconds.

Co-crystals were incubated in a buffer solution containing 20% (w/v) PEG 4000, 50 mM tricine (pH 8.5), 7.5% (w/v) heptanetriol, 0.1% (w/v) LDAO, 5 mM 2,3-dimethoxy-5-methylbenzoquinone (UQ_0) and 5 mM 2,3-dimethoxy-5-methylhydrobenzoquinol (UQ_0H_2); and mounted in 0.7 mm quartz capillaries (Charles Supper, Natwick MA). The mixture of UQ_0 and UQ_0H_2 served to reduce the cyt c_2 in the co-crystal, which allowed signal averaging of multiple flashes. Crystals were also soaked in buffer with 50 mM potassium ferricyanide to observe the kinetics when cyt c_2 is oxidized.

X-ray data collection

X-ray diffraction data were collected at either beamline 9-1 of the Stanford Synchrotron Radiation Laboratory (SSRL) or beamline 5.0.2 of the Advanced Light

Source (ALS, Berkeley, CA) on two different crystal forms. Crystals were mounted on nylon loops and soaked sequentially for ~30 sec in solutions containing 22% (w/v) PEG 4000, 20 mM Tricine (pH 8.5), 0.06%(w/v) LDAO, 3.9%(w/v) heptanetriol and increasing concentrations (5%, 15%, and 20% v/v) of glycerol cryoprotectant (*69*). Following the soaks, crystals were plunged into liquid nitrogen and mounted onto a goniostat. Data were collected on crystals cooled to ~100 K using a stream of cold nitrogen gas. The diffraction data were processed with the MOSFLM (*70*) or HKL (*71*) software packages, and scaled with either the CCP4 SCALA (*72*) program or the HKL SCALPACK (*71*) program. Data processing statistics are shown in Table 6.1.

X-ray structure determination

Since the RC comprises most of the scattering material in the co-crystal, molecular replacement methods were used to position the RC in the unit cell. RC coordinates refined at 2.2 Å resolution (*11*) were used in this calculation, with the molecular replacement implemented with the XPLOR (*73*), CNS (*74*), or the AMoRe (*75*) programs. Two co-crystal forms, designated Type I and Type II, were investigated. Both forms belong to space group $P2_1$, with two (type I) or one (type II) complex present per asymmetric unit.

Simulated annealing omit electron density maps of the form $2F_o$-F_c or F_o-F_c were calculated with CNS, contoured with the program MAPMAN (*40*), and displayed on computer workstations using the programs TOM/FRODO (*76*) or O (*77*). To determine the location of cyt c_2, difference electron density maps were computed with the RC

molecular replacement phases. These maps revealed a large peak, 5-6 σ above the background level, at the same relative location in both co-crystal forms. The magnitude and location of this peak suggested its assignment to the heme iron atom of cyt c_2. Solvent-flattened and non-crystallographic symmetry averaged difference maps revealed additional electron density corresponding to the heme and its axial ligands, His C19 and Met C100, and the α-helices of cyt c_2. The previously determined coordinates of cyt c_2 from *Rb. sphaeroides* (PDB ID 1CXC (*12*)) were subsequently modeled into this electron density at the periplasmic surface of the RC.

Rigid body, positional, simulated annealing, and isotropic temperature factor refinement were carried out with the CNS package incorporating maximum-likelihood refinement target functions (*78*), non-crystallographic symmetry restraints, and anisotropic bulk-solvent correction. Refinement statistics are shown in Table 6.1. The 2.2 Å resolution coordinates of the RC (*11*) and the 1.6 Å resolution coordinates of cyt c_2 (*12*) positioned by molecular replacement as described above were used as starting models for refinement in the two crystal forms. During the final stages of refinement, water molecules were incorporated into the 2.40 Å model for the type II form using automated CNS command scripts that locate difference electron density peaks that are 3 σ above the average background level of the map and within hydrogen bonding distance of potential donors or acceptors.

RCSB Protein Data Bank accession codes

The coordinates, isotropic temperature factors, and observed structure factors have been submitted to the RCSB Protein Data Bank (form I, ID code 1L9J; form II, ID code 1L9B).

Acknowledgments

We thank R. Isaacson for his expert help in designing and running the microspectrophotometer, Keith Henderson at the Advanced Light Source (ALS), and the staff at the Stanford Synchrotron Radiation Laboratory (SSRL) for their advice and assistance with the X-ray data collections; Mark Paddock, Raphael Calvo, Les Dutton, Marilyn Gunner, and Jose Onuchic for helpful discussions. The synchrotron X-ray data collection facilities are supported by the U.S. Department of Energy (ALS,SSRL) and NIH (SSRL). This work was supported by the National Institutes of Health (grant number GM-13191 to G.F. and GM-45162 to D.C.R) and National Science Foundation (grant number 9974568 to M.Y.O.).

References

1. Bendall, D. S. (Ed.) (1996) *Protein Electron Transfer*, Bios Scientific Publishers, Oxford, UK.
2. Graber, P., and Milazzo, G. (Eds.) (1997) *Bioenergetics*, Vol. 4, Birkhauser, Boston, MA.
3. Mathews, F. S., Mauk, G. A., and Moore, G. R. (2000) in *Protein-Protein Recognition* (Kleasnthous, C., Ed.) pp 60-101, Oxford University Press, Oxford, UK.
4. Feher, G., Allen, J. P., Okamura, M. Y., and Rees, D. C. (1989) *Nature 339*, 111-116.
5. Blankenship, R. E., Madigan, M. T., and Bauer, C. E. (Eds.) (1995) *Anoxygenic Photosynthetic Bacteria*, Kluwer Academic Publishers, Dordrecht, Netherlands.
6. Michel-Beyerle, M.-E. (Ed.) (1996) *Reaction Center of Photosynthetic Bacteria: Structure and Dynamics*, Springer, Berlin, Germany.
7. Breton, J., Nabedryk, E., and Vermeglio, A. (1998) *Photosynth. Res. 55*, 117-378.
8. Allen, J. P., Feher, G., Yeates, T. O., Rees, D. C., Deisenhofer, J., Michel, H., and Huber, R. (1986) *Proc. Natl. Acad. Sci. U. S. A. 83*, 8589-8593.
9. Chang, C. H., Tiede, D., Tang, J., Smith, U., Norris, J., and Schiffer, M. (1986) *FEBS Lett. 205*, 82-86.
10. Ermler, U., Fritzsch, G., Buchanan, S. K., and Michel, H. (1994) *Structure 2*, 925-936.
11. Stowell, M. H. B., McPhillips, T. M., Rees, D. C., Soltis, S. M., Abresch, E., and Feher, G. (1997) *Science 276*, 812-816.

12. Axelrod, H. L., Feher, G., Allen, J. P., Chirino, A. J., Day, M. W., Hsu, B. T., and Rees, D. C. (1994) *Acta Crystallogr. Sect. D-Biol. Crystallogr. 50*, 596-602.

13. Crofts, A. R., and Berry, E. A. (1998) *Curr. Opin. Struct. Biol. 8*, 501-509.

14. Tiede, D. M., and Dutton, P. L. (1993) in *The Photosynthetic Reaction Center* (Deisenhofer, J., and Norris, J. R., Eds.) pp 257-298, Academic Press, San Diego, CA.

15. Overfield, R. E., Wraight, C. A., and Devault, D. (1979) *FEBS Lett. 105*, 137-142.

16. Rosen, D., Okamura, M. Y., and Feher, G. (1980) *Biochemistry 19*, 5687-5692.

17. Moser, C. C., and Dutton, P. L. (1988) *Biochemistry 27*, 2450-2461.

18. Tiede, D. M., Vashishta, A. C., and Gunner, M. R. (1993) *Biochemistry 32*, 4515-4531.

19. Rosen, D., Okamura, M. Y., Abresch, E. C., Valkirs, G. E., and Feher, G. (1983) *Biochemistry 22*, 335-341.

20. Tetreault, M., Rongey, S. H., Feher, G., and Okamura, M. Y. (2001) *Biochemistry 40*, 8452-8462.

21. Lin, X., Williams, J. C., Allen, J. P., and Mathis, P. (1994) *Biochemistry 33*, 13517-13523.

22. Venturoli, G., Drepper, F., Williams, J. C., Allen, J. P., Lin, X., and Mathis, P. (1998) *Biophys. J. 74*, 3226-3240.

23. Allen, J. P., Feher, G., Yeates, T. O., Komiya, H., and Rees, D. C. (1987) *Proc. Natl. Acad. Sci. U. S. A. 84*, 6162-6166.

24. Okamura, M. Y., and Feher, G. (1983) *Biophys. J. 41*, A122-A122.

25. Long, J. E., Durham, B., Okamura, M., and Millett, F. (1989) *Biochemistry 28*, 6970-6974.

26. Caffrey, M. S., Bartsch, R. G., and Cusanovich, M. A. (1992) *J. Biol. Chem. 267*, 6317-6321.

27. Adir, N., Axelrod, H. L., Beroza, P., Isaacson, R. A., Rongey, S. H., Okamura, M. Y., and Feher, G. (1996) *Biochemistry 35*, 2535-2547.

28. Tetreault, M., Cusanovich, M., Meyer, T., Axelrod, H. L., and Okamura, M. Y. (2001) *Biochemistry (in press).*

29. Drepper, F., Dorlet, P., and Mathis, P. (1997) *Biochemistry 36*, 1418-1427.

30. Drepper, F., and Mathis, P. (1997) *Biochemistry 36*, 1428-1440.

31. Tiede, D. M., and Chang, C. H. (1988) *Isr. J. Chem. 28*, 183-191.

32. Deisenhofer, J., Epp, O., Sinning, I., and Michel, H. (1995) *J. Mol. Biol. 246*, 429-457.

33. Pelletier, H., and Kraut, J. (1992) *Science 258*, 1748-1755.

34. Chen, L. Y., Durley, R. C. E., Mathews, F. S., and Davidson, V. L. (1994) *Science 264*, 86-90.

35. Schindelin, N., Kisker, C., Sehlessman, J. L., Howard, J. B., and Rees, D. C. (1997) *Nature 387*, 370-376.

36. Morales, R., Charon, M. H., Kachalova, G., Serre, L., Medina, M., Gomez-Moreno, C., and Frey, M. (2000) *EMBO Rep. 1*, 271-276.

37. Kurisu, G., Kusunoki, M., Katoh, E., Yamazaki, T., Teshima, K., Onda, Y., Kimata-Ariga, Y., and Hase, T. (2001) *Nat. Struct. Biol. 8*, 117-121.

38. Axelrod, H. L., Abresch, E., Isaacson, R., Tetreault, M., Okamura, M. Y., and Feher, G. (1998) *Biophys. J. 74*, A146-A146.

39. Axelrod, H. L., Abresch, E. C., Okamura, M. Y., Feher, G., Yeh, A. P., and Rees, D. C. (1999) *Biophys. J. 76*, A20-A20.

40. Kleywegt, G. J., and Jones, T. A. (1996) *Acta Crystallogr. Sect. D-Biol. Crystallogr. 52*, 826-828.

41. Yeates, T. O., Komiya, H., Rees, D. C., Allen, J. P., and Feher, G. (1987) *Proc. Natl. Acad. Sci. U. S. A. 84*, 6438-6442.

42. Komiya, H., Yeates, T. O., Rees, D. C., Allen, J. P., and Feher, G. (1988) *Proc. Natl. Acad. Sci. U. S. A. 85*, 9012-9016.

43. Allen, J. P., Feher, G., Yeates, T. O., Komiya, H., and Rees, D. C. (1987) *Proc. Natl. Acad. Sci. U. S. A. 84*, 5730-5734.

44. Lee, B. K., and Richards, F. M. (1971) *J. Mol. Biol. 55*, 379-400.

45. Creighton, T. E. (1993) *Proteins - Structures and Molecular Properties*, W. H. Freeman and Co., New York.

46. Ma, J. C., and Dougherty, D. A. (1997) *Chem. Rev. 97*, 1303-1324.

47. Gallivan, J. P., and Dougherty, D. A. (1999) *Proc. Natl. Acad. Sci. U. S. A. 96*, 9459-9464.

48. Hopfield, J. J. (1974) *Proc. Natl. Acad. Sci. U. S. A. 71*, 3640-3644.

49. Marcus, R. A., and Sutin, N. (1985) *Biochim. Biophys. Acta 811*, 265-322.

50. Moser, C. C., Keske, J. M., Warncke, K., Farid, R. S., and Dutton, P. L. (1992) *Nature 355*, 796-802.

51. Tezcan, F. A., Crane, B. R., Winkler, J. R., and Gray, H. B. (2001) *Proc. Natl. Acad. Sci. U. S. A. 98*, 5002-5006.

52. Farchaus, J. W., Wachtveitl, J., Mathis, P., and Oesterhelt, D. (1993) *Biochemistry 32*, 10885-10893.

53. Wachtveitl, J., Farchaus, J. W., Mathis, P., and Oesterhelt, D. (1993) *Biochemistry 32*, 10894-10904.

54. Ortega, J. M., and Mathis, P. (1993) *Biochemistry 32*, 1141-1151.

55. Chen, I. P., Mathis, P., Koepke, J., and Michel, H. (2000) *Biochemistry 39*, 3592-3602.

56. Dohse, B., Mathis, P., Wachtveitl, J., Laussermair, E., Iwata, S., Michel, H., and Oesterhelt, D. (1995) *Biochemistry 34*, 11335-11343.

57. Aquino, A. J. A., Beroza, P., Beratan, D. N., and Onuchic, J. N. (1995) *Chem. Phys. 197*, 277-288.

58. Farid, R. S., Moser, C. C., and Dutton, P. L. (1993) *Curr. Opin. Struct. Biol. 3*, 225-233.

59. Aquino, A. J. A., Beroza, P., Reagan, J., and Onuchic, J. N. (1997) *Chem. Phys. Lett. 275*, 181-187.

60. Beratan, D. N., Betts, J. N., and Onuchic, J. N. (1991) *Science 252*, 1285-1288.

61. Witt, H., Malatesta, F., Nicoletti, F., Brunori, M., and Ludwig, B. (1998) *J. Biol. Chem. 273*, 5132-5136.

62. Zhen, Y. J., Hoganson, C. W., Babcock, G. T., and Ferguson-Miller, S. (1999) *J. Biol. Chem. 274*, 38032-38041.

63. Wang, K. F., Zhen, Y. J., Sadoski, R., Grinnell, S., Geren, L., Ferguson-Miller, S., Durham, B., and Millett, F. (1999) *J. Biol. Chem. 274*, 38042-38050.

64. Roberts, V. A., and Pique, M. E. (1999) *J. Biol. Chem. 274*, 38051-38060.

65. Isaacson, R. A., Lendzian, F., Abresch, E. C., Lubitz, W., and Feher, G. (1995) *Biophys. J. 69*, 311-322.

66. Bartsch, R. G. (1978) in *The Photosynthetic Bacteria* (Clayton, R. K., and Sistrom, W. R., Eds.) pp 349-386, Plenum Press, New York.

67. Straley, R. A., Parson, W. W., Mauzerall, D. C., and Clayton, R. K. (1973) *Biochim. Biophys. Acta 305*, 597-609.

68. Allen, J. P., and Feher, G. (1984) *Proc. Natl. Acad. Sci. U. S. A. 81*, 4795-4799.

69. Rodgers, D. W. (1997) *Methods Enzymol. 276*, 183-203.

70. Leslie, A. G. W. (1992) *Joint CCP4 and ESF-EACMB Newsletter on Protein Crystallography, No. 26*, Daresbury Laboratory, Warrington, U. K.

71. Otwinowski, Z., and Minor, W. (1997) *Methods Enzymol. 276*, 307-326.

72. Bailey, S. (1994) *Acta Crystallogr. Sect. D-Biol. Crystallogr. 50*, 760-763.

73. Brünger, A. T. (1993) *X-PLOR, v. 3.1: A System for X-ray Crystallography and NMR*, Yale University Press, New Haven, CT.

74. Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) *Acta Crystallogr. Sect. D Biol. Crystallogr. 54*, 905-921.

75. Navaza, J. (2001) *Acta Crystallogr. Sect. D-Biol. Crystallogr. 57*, 1367-1372.

76. Jones, T. A. (1985) *Methods Enzymol. 115*, 157-171.

77. Jones, T. A., and Kjeldgaard, M. (1997) *Methods Enzymol. 277*, 173-208.

78. Brünger, A. T., Adams, P. D., and Rice, L. M. (1998) *Curr. Opin. Struct. Biol. 8*, 606-611.

79. Luzzati, P. V. (1952) *Acta Crystallogr. 5*, 802-810.

80. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) *J. Appl. Crystallogr. 26*, 283-291.

81. Carson, M. (1997) *Methods Enzymol. 277*, 493-505.

82. Esnouf, R. M. (1999) *Acta Crystallogr. Sect. D-Biol. Crystallogr. 55*, 938-940.

83. Merritt, E. A., and Bacon, D. J. (1997) *Methods Enzymol. 277*, 505-524.

84. Bogan, A. A., and Thorn, K. S. (1998) *J. Mol. Biol. 280*, 1-9.

Table 6.1. Data collection and refinement statistics.

Crystal Form	Type I	Type II
Crystal Data		
Space Group	$P2_1$	$P2_1$
a (Å)	77.9	78.2
b (Å)	80.3	115.6
c (Å)	246.6	79.7
β (°)	92.4	110.3
RC-cyt c_2 complexes/asymmetric unit	2	1
Data Collection		
Maximum Resolution (Å)	3.25	2.40
Total Observations (unique)	348,873(47,668)	198,053(50,331)
Mean $I/\sigma(I)$[a] (highest resolution shell)	4.8(1.7)	4.6(2.0)
R_{symm}[b] (highest resolution shell)(%)	11.9(45.0)	10.0(33.6)
Completeness[c] (last shell) (%)	98.9(98.8)	96.8(83.7)
Refinement		
Resolution Range (Å)	50-3.25	50-2.40
R-factor[d] (%)	26.9	23.4
Average B-factor RC ($Å^2$)	60	63
Average B-factor cyt c_2 ($Å^2$)	67	62
R_{free}[e] (%)	31.5	27.8
Deviation from ideal bond lengths (Å)	0.015	0.012
Deviation from ideal bond angles (°)	1.85	1.78
rms coordinate error[f] (Å)	0.4	0.3
Ramachandran Plot[g]		
Residues in most favored regions (%)	80.3	86.7
Residues in additional allowed regions (%)	17.3	11.5
Residues in generously allowed regions (%)	1.3	1.5
Residues in disallowed regions (%)	1.1	0.3

Table 6.1. *continued*

[a] $I/\sigma(I)$ is the ratio of the average of the diffraction intensities to the average background intensity.

[b] $R_{symm}=\Sigma_{hkl}\Sigma_j|\ I_{hkl} - \langle I_{hkl}\rangle|/\Sigma_{hkl}\Sigma_j|I_{hkl}|$ where $\langle I_{hkl}\rangle$ is the average intensity for a set of j symmetry-related reflections and $I_{hkl,}$ is the value of the intensity for a single reflection within a set of symmetry related reflections.

[c] Completeness is the ratio of the number of reflections measured to the total number of reflections possible.

[d] R-factor $=\Sigma_{hkl}\ |F_o|-|F_c|/\Sigma_{hkl}|F_o|$ where $|F_o|$ is the observed structure factor amplitude and $|F_c|$ is the calculated structure factor amplitude.

[e] $R_{free}= \Sigma_{hkl,T}|F_o|-|F_c|/\Sigma_{hkl,T}|F_o|$ where a test set , T, represented by 5% of the data, is omitted from the refinement.

[f] rms error in coordinates based on the method of Luzzatti (*79*).

[g] Ramachandran plots were determined with the program PROCHECK (*80*).

Table 6.2. Intermolecular contacts and distances between cyt c_2 and RC for different types of interactions.

Type of Intermolecular Interaction	Reaction Center	Cytochrome c_2	Interatomic Distance (Å)[a]
Oppositely Charged Residues[b]	Asp L155 OD1	Arg C32 NH1	6.7
	Asp L257 OD1	Lys C99 NZ	6.8
	Asp L261 OD1	Lys C97 NZ	5.7
	Asp L261 OD2	Lys C99 NZ	5.5
	Glu M95 OE1	Lys C103 NZ	8.3
	Glu M95 OE2	Lys C103 NZ	7.8
	Glu M173 OE2	Lys C105 NZ	8.9
	Asp M184 OD2	Lys C103 NZ	6.7
Van der Waals Contacts[c]	Tyr L162 CE1	Heme CBC	3.9
	Leu M191 CG	Heme CMC	3.9
	Leu M191 CB	Phe C102 CD1	3.4
	Tyr L162 CB	Thr C36 OG1	3.7
	Gly L165 CA	Thr C101 CG2	3.8
	Asn L166 N	Thr C101 CG2	4.0
	Gln L258 OE1	Thr C101 CG2	3.7
	Asn M188 OD1	Phe C102 CD1	3.5
	Tyr L162 CB	Thr C36 OG1	3.7
	Gly L165 CA	Thr C101 CG2	3.8
	Asn L166 N	Thr C101 CG2	4.0
	Gln L258 OE1	Thr C101 CG2	3.7
	Asn M188 OD1	Phe C102 CD1	3.5
	Leu M191 O	Gln C14 CB	3.4
	Leu M191 CD1	Thr C17 CG2	3.9
	Val M192 CA	Gln C14 OE1	3.5
	Gly M194 O	Arg C32 NH1	4.0
	Asn M293 ND2	Asn C13 CB	3.4
Interprotein Hydrogen Bonding	Gln L258 NE2	Lys 99 O	3.5
	Asn M187 ND2	Thr C101 O	3.1
	Asn M188 OD1	Lys 103 N	3.1
Cation-π	Tyr M295 CZ	Arg C32 NE	3.7

[a] The mean error in the bond lengths is ± 0.3 Å

[b] Oppositely charged pairs at distances ≤ 10 Å are listed.

[c] Interactions at distances ≤ 4.0 Å are listed

Table 6.3. Hydrogen bond contacts and B-factors for bridging water molecules.

	B-Factor (Å^2)	**Reaction Center**	**Cytochrome c_2**
Wat1	41	Leu M191 O	Gln C14 O, thr C17 Oγ1
Wat2	45	Asn L159 Oδ1	Thr C17 Oγ, Thr C17 O
Wat5	66	Gln L258 Oε1	Gly C98 O
Wat10	60	Ser L158 Oγ, Asn L159 Oδ1	Thr C17 O
Wat12	42	Asn L159 O, Thr L163 Oγ1	Thr C36 Oγ1
Wat13	71	Asp L257 Oδ2, Gln L258 N, Asp L261 Oδ2	Lys C99 Nζ
Wat20	64	Asn L159 Oδ1	Lys C35 N, Thr C36 N

Figure 6.1. Light-induced absorption changes at 550 nm following a single laser flash in a type I cyt c_2:RC co-crystal. The absorbance change of the top trace (a) with oxidized cytochrome is due to the formation of $D^+Q_A^-$ (equation (1)). The absorbance change of the bottom trace (b) is due to electron transfer from reduced cytochrome to the oxidized donor, D^+ (equation (2)). The characteristic electron transfer time in the co-crystal was determined to be 0.9 (± 0.1) μs, the same as in solution. This indicates that the structure of the complex in the region of electron transfer is the same in the co-crystal and in solution.

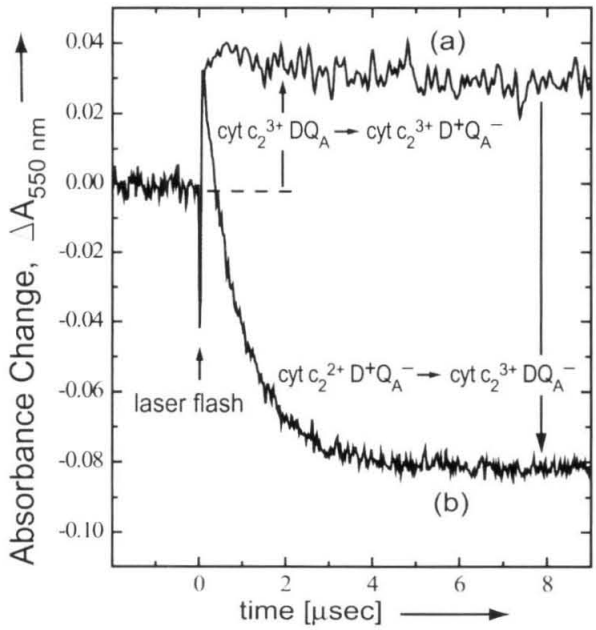


Figure 6.2. A stereoview of the cyt c_2:RC complex from *Rb. sphaeroides* showing the location of the bound cyt c_2 (lavender), the heme prosthetic group (turquoise), the RC L subunit (yellow), the RC M subunit (blue), the RC H-subunit (green), the RC primary donor (red), and non-heme Fe atom (red). The location of the conformational change in the co-crystal at the N-terminal end of the M subunit is indicated by an arrow. The twofold symmetry axis of the RC is parallel to the plane of the drawing approximately connecting the irons on the RC and the heme. The illustration was made with the program RIBBONS (*81*).

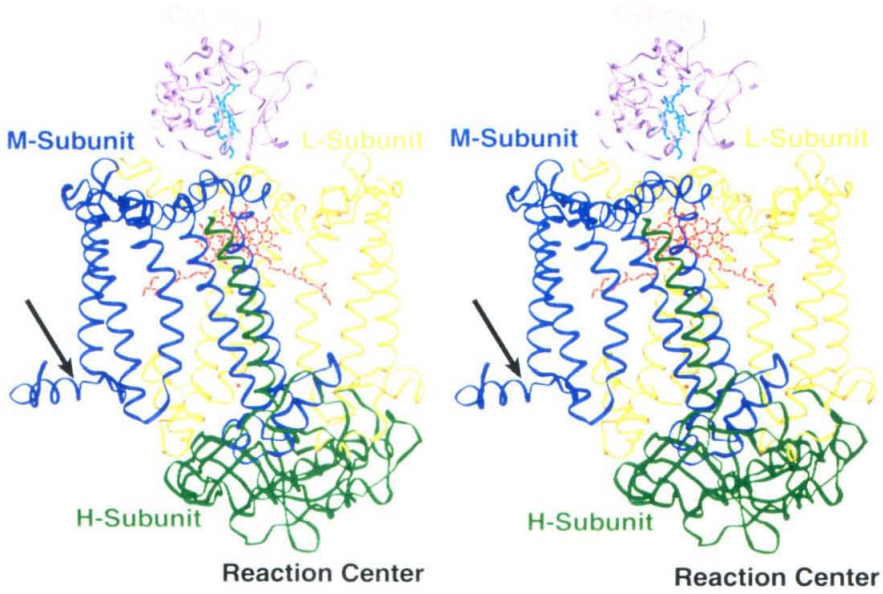


Figure 6.3. Stereoview of the cyt c_2:RC co-crystal structure in the vicinity of the bound cyt c_2 heme (turquoise) with cyt c_2 in lavender, the RC L subunit in yellow, the RC M subunit in blue, and the RC donor in red. A $2|F_o|-|F_c|$ simulated annealing omit electron density map (*74*) calculated at a resolution of 2.4 Å and contoured at 1.0 σ is shown in pink for cyt c_2 and green for the RC. To reduce phase bias, the cyt c_2 heme and its axial ligands (His C19 and Met C100) were excluded from the calculation of the map. The illustration was made with the programs BOBSCRIPT (*82*) and Raster3D (*83*).

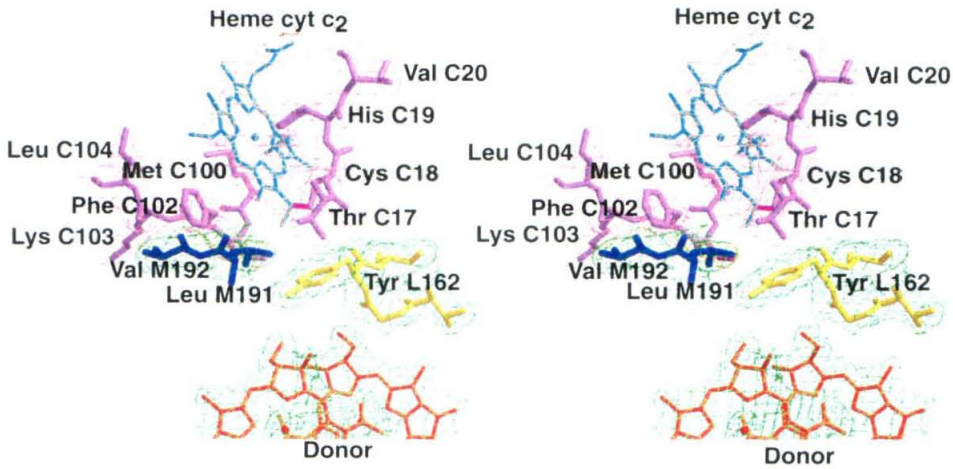
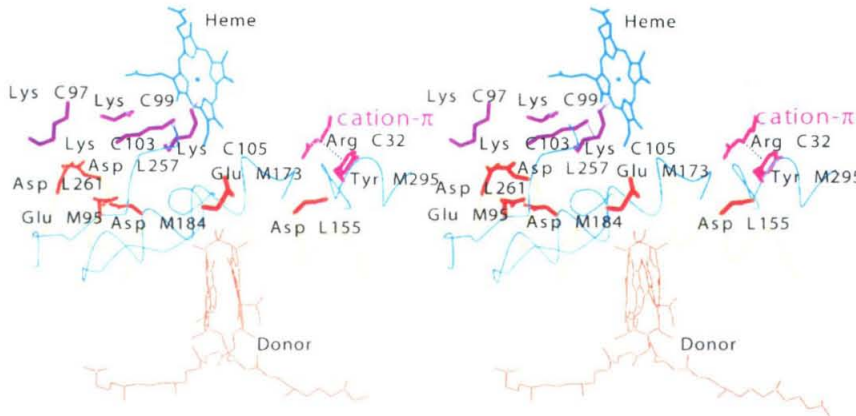
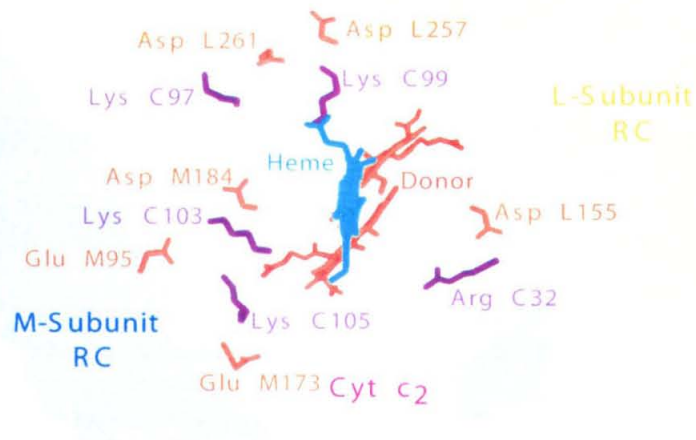


Figure 6.4. *Two views of electrostatically interacting residues in the cyt c_2:RC complex from Rb. sphaeroides.* (a) Stereoview showing positively charged lysine side chains (violet) on cyt c_2 facing negatively charged aspartic and glutamic acid side chains (red) on the persiplasmic surface of the L (yellow) and M (cyan) subunits of the RC. The heme on cyt c_2 is in turquoise and the primary donor, on the RC is in red. Most of the complementary electrostatic interactions are on the M side (left) of the RC near Asp M184 and Glu M95. The positively charged guanididium side chain of Arg C32 (magenta) is in close proximity to the side chain of Tyr M295 (magenta) forming a cation-π interaction. (b) A representation of the cyt c_2:RC complex in a view normal to the periplasmic surface of the RC. The molecular surfaces of the interacting subunits are outlined in yellow for L, blue for M, and lavender for cyt c_2. Positive side chains on cyt c_2 are in violet and negative side chains on the RC in red.



(a)



(b)

Figure 6.5. Stereo presentation of the interface region containing the closest contacts between the RC and bound cyt c_2. Side chains on cyt c_2 (lavender) interacting with RC side chains are shown. Interaction regions on the L subunit of the RC are in yellow and M subunit in blue. The primary donor on the RC is shown in red. The CBC methyl group on the cyt c_2 heme (turquoise) is in van der Waals contact with Tyr L162 on the RC. This region contains many short-range nonpolar interactions, including those between Phe C102 on cyt c_2 and Leu M191 and Val M192 on the RC (see Table 6.2). This region is believed to be important for intermolecular electron transfer between reduced cyt c_2 and the photo-oxidized donor.

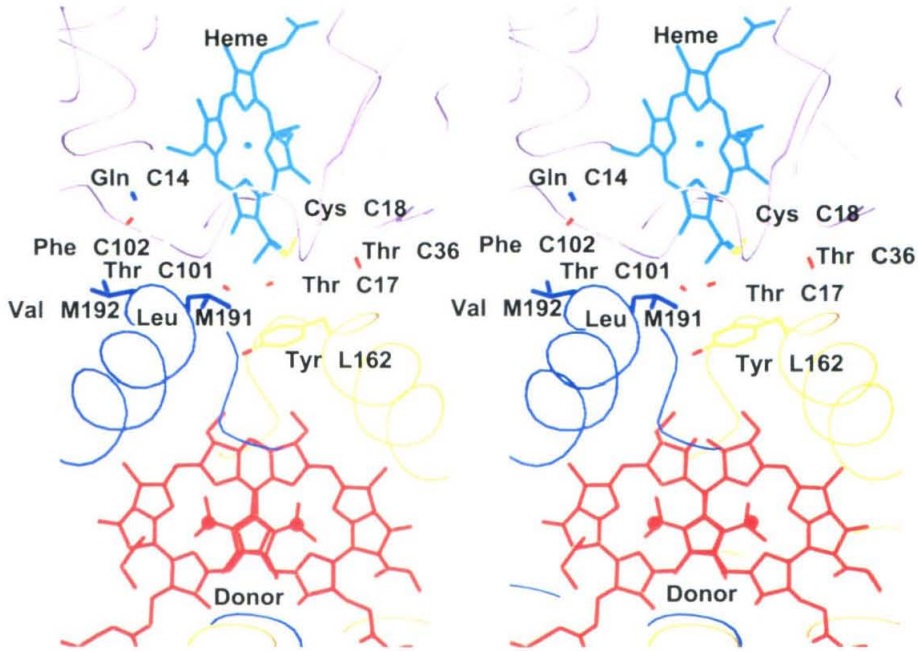


Figure 6.6. *Stereoview* of intermolecular hydrogen bonding in the cyt c_2:RC complex. Interacting nitrogen atoms are in blue and side chain oxygens and water oxygens are in red. Dashed lines represent hydrogen bonds. Three amide side chains on the RC (Gln L258, Asn M187, and Asn M188) are positioned to hydrogen bond with backbone atoms on cyt c_2 near the exposed region on the heme (turquoise). L subunit (yellow), M subunit (cyan), cyt c_2 (lavender).

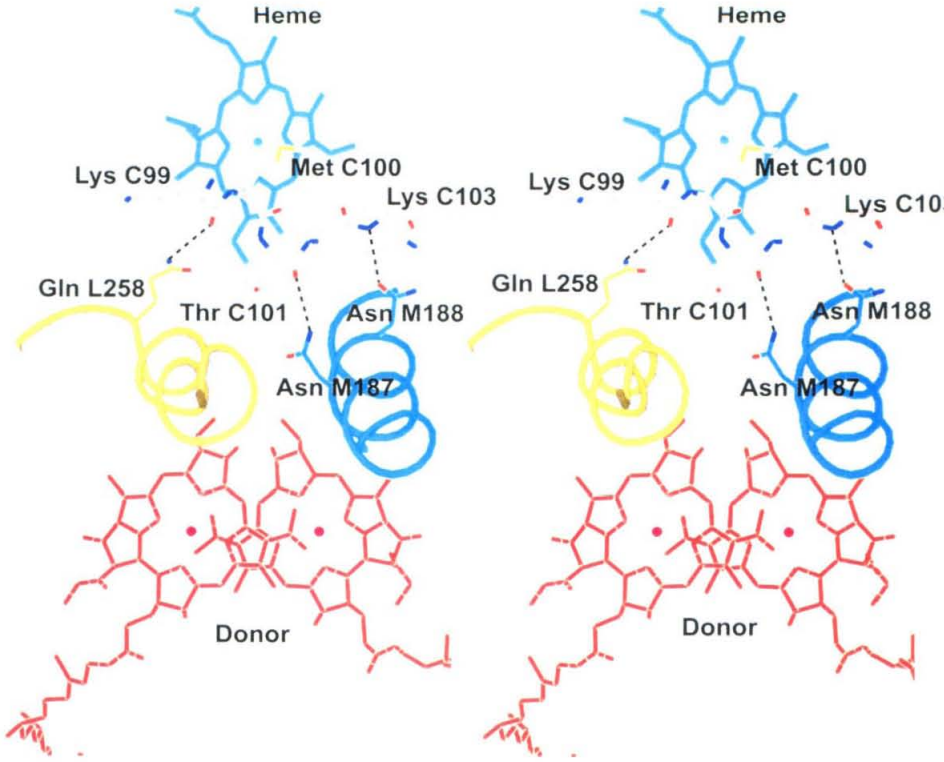


Figure 6.7. Interacting atoms in the cyt c_2:RC complex from *Rb. sphareoides* mapped onto the van der Waals surfaces of the individual proteins. The two proteins have been peeled apart to expose the docking interface on both partners by a 180° rotation of the cyt c_2 about a vertical axis (dashed) in the plane of the paper (*84*). The region of closest contact in the complex is between Tyr L162 on the RC (green) and the solvent-exposed edge of the heme on cyt c_2 (green). Surrounding atoms are color coded according to interaction type with short-range van der Waals interactions in yellow. Hydrophobic amino acid side chains on the RC (Leu M191, Val M192) that form van der Waals contacts with Phe C102 on cyt c_2 are in gray. Atoms on the RC in hydrogen bonding contact with backbone atoms on cyt c_2 are in cyan. The two residues that interact to form a cation-π interaction (Tyr M295 and Arg C32) are in pink. Charged side chains, lys on cyt c_2 (violet) and asp and glu on the RC (red) that are separated by 10 Å or less are shown. The short-range interactions (yellow, gray, cyan, and pink) are clustered around Tyr L162 and the heme edge, whereas the residues in the electrostatic domain are located further from the region of closest contact. The side chains in the electrostatic domain interact in solvent-accessible regions, whereas most of the contact regions shown in yellow, gray, cyan or pink are excluded from solvent. The figure illustrates the presence of two domains: A region containing short-range interactions (green, yellow, gray, cyan, pink) and one with long-range, electrostatic interactions (red and violet).