Biological identity and diversity in photosynthesis and respiration: structure of the lumen-side domain of the chloroplast Rieske protein

Christopher J Carrell, Huamin Zhang, William A Cramer and Janet L Smith*

Background: The cytochrome *b*$_f$ complex functions in oxygenic photosynthesis as an integral membrane protein complex that mediates coupled electron transfer and proton translocation. The Rieske [2Fe–2S] protein subunit of the complex functions at the electropositive (p) membrane interface as the electron acceptor for plastoquinol and donor for the cytochrome *f* subunit, and may have a dynamic role in catalyzing electron and proton transfer at the membrane interface. There are significant structure/function similarities to the protein and for the pH of the intraorganelle compartment.

Results: The 1.83 Å crystal structure of a 139-residue C-terminal fragment of the Rieske [2Fe–2S] protein, derived from the cytochrome *b*$_f$ complex of spinach chloroplasts, has been solved by multiwavelength anomalous diffraction. The structure of the fragment comprises two domains; a small 'cluster-binding' subdomain and a large subdomain. The [2Fe–2S] cluster-binding subdomains of the chloroplast and mitochondrial Rieske proteins are virtually identical, whereas the large subdomains are strikingly different despite a common folding topology. A structure-based sequence alignment of the *b*$_f$ and *bc*$_1$ groups of Rieske soluble domains is presented.

Conclusions: The segregation of structural conservation and divergence in the cluster-binding and large subdomains of the Rieske protein correlates with the overall relatedness of the cytochrome *b*$_f$ and *bc*$_1$ complexes, in which redox domains in the aqueous *p* phase are dissimilar and those within the membrane are similar. Distinct sequences and surface charge distributions among Rieske large subdomains may provide a signature for interaction with the *p*-side oxidant protein and for the pH of the intraorganelle compartment.

Introduction

Energy transduction in oxygenic photosynthesis is carried out by four integral membrane protein complexes: the photosystem II reaction center, the cytochrome *b*$_f$ complex, the photosystem I reaction center and the ATP synthase. In addition to mediating electron transfer between the two reaction center complexes, the cytochrome *b*$_f$ complex accomplishes the translocation of protons across the membrane. The cytochrome *b*$_f$ complex comprises as many as seven subunits. Four 'large' subunits (molecular weight [MW] 18 000–32 000 Da) are products of the pet*A–D* genes, and as many as three small hydrophobic subunits (MW ~4000 Da) are products of the pet*G*, *L*, and *M* genes. The four large subunits are cytochrome *f* (with one c heme), cytochrome *b*$_6$ (with two *b* hemes), the Rieske protein (containing a high-potential [2Fe–2S] cluster), and subunit IV (Figure 1; summarized in [1]).

The photosynthetic cytochrome *b*$_f$ complex is related in function to the cytochrome *bc*$_1$ complex that has a central role in the mitochondrial respiratory chain and in bacterial photosynthesis. The two complexes are a study of structural identity and diversity. The mitochondrial cytochrome *b* is a fusion of chloroplast cytochrome *b*$_6$ and subunit IV protomers [2], but the *c*-type cytochromes *f* and *c*$_1$ are unrelated. The relation of the Rieske proteins is discussed in the present study. In both the *b*$_{h,f}$ and *bc*$_{1}$ cytochrome complexes, the redox-active domains of the Rieske protein and the *c*-type cytochrome reside in the aqueous phase on the electropositive (p)-side of the energy-transducing membrane while those of the *b*-type cytochrome are located within the bilayer. The electron donors to the cytochrome *b*$_{h,f}$ and *bc*$_{1}$ complexes are the related membrane-soluble small molecules plastoquinol and ubiquinol, respectively. The electron acceptors are the unrelated soluble proteins plastocyanin and cytochrome *c*. Atomic level structural information from these integral membrane protein complexes is limited although high-resolution crystal structures of soluble, redox-active domains of chloroplast cytochrome *f* [3,4] and of the mitochondrial Rieske protein [5] have been reported. Preliminary crystal structures of intact cytochrome *bc*$_1$ have been presented recently [6–9].

*Corresponding author.
E-mail: smithj@bragg.bio.purdue.edu

Key words: cytochrome *b*$_f$ complex, cytochrome *f*, energy transduction, plastoquinone, proton translocation

Received: 19 August 1997
Revisions requested: 29 September 1997
Revisions received: 15 October 1997
Accepted: 16 October 1997

Structure 15 December 1997, 5:1613–1625
http://biomednet.com/elecref/0969212600501613

© Current Biology Ltd ISSN 0969-2126
addition, an 8 Å projection structure of an algal cytochrome b_{6f} complex has been obtained [10].

The Rieske [2Fe–2S] protein of oxygenic photosynthesis is the primary oxidant of plastoquinol in the cytochrome b_{6f} complex on the p-side (lumen) of the membrane, generating plastoquinone. The protons released in the oxidation are ultimately transferred to the p-side bulk aqueous phase [11], possibly via the Rieske protein and its electron acceptor cytochrome f [4]. The mature intact Rieske protein in the cytochrome b_{6f} complex of spinach chloroplasts contains 179 residues [12], and appears to be anchored in the membrane by electrostatic forces [13] and by one relatively hydrophobic N-terminal transmembrane helix [14]. A soluble 139-residue C-terminal polypeptide fragment of the Rieske protein has been characterized and crystallized [15]. The electron paramagnetic resonance (EPR) spectrum of the reduced fragment (g = 2.03, g_y = 1.90, and a broad g_x = 1.74) and the room temperature midpoint oxidation/reduction potential (E_m = 320 mV at pH 7) are similar to the values of those parameters in the cytochrome b_{6f} complex. The E_m of the fragment is pH-dependent, with E_m = 359 mV at pH 6 and pK_m = 6.5.

A soluble 129-residue C-terminal polypeptide fragment of the related Rieske protein from the cytochrome b_{6f} complex of bovine heart mitochondria was characterized biochemically [16] and a crystal structure determined [5]. The EPR spectrum of the reduced fragment (g = 2.01–2.02, g_y = 1.90, g_x = 1.76–1.80) and the midpoint oxidation-reduction potential (E_m = 306 mV at pH 7) are similar to the values of those parameters in the intact cytochrome b_{6f} complex. The [2Fe–2S] center at one end of the Rieske soluble domain is coordinated by two cysteine ligands to one Fe and two histidines to the other Fe. The two histidine ligands are exposed to the solvent.

The three-dimensional structure of the chloroplast Rieske protein, together with the previously determined structure of its electron acceptor cytochrome f, may provide insight into the mechanism of electron transfer in the complex. Electron transfer from membrane-bound plastoquinol through the Rieske [2Fe–2S] cluster to the cytochrome f heme may occur over a long distance because the heme Fe is 45 Å from the site of membrane attachment in the elongate structure of cytochrome f. The distance will depend on the orientation of cytochrome f relative to the membrane surface and the packing arrangement of both the Rieske protein and cytochrome f in the complex. Interestingly, there is indirect evidence for a flexible orientation of both proteins. EPR spectra of cytochrome f in oriented membranes indicated a range of orientations of the heme relative to the membrane plane, suggestive of limited flexibility for the soluble domain [17,18]. The chloroplast Rieske N-terminal membrane anchor is linked to the soluble domain by a very glycine-rich peptide, which may be unstructured and quite flexible. The rate constant for electron transfer from the Rieske protein to cytochrome f is approximately 10^{14} s^{-1} [19], which corresponds to a donor–acceptor distance of 12–17 Å by current estimates of electron transfer rates in proteins [20,21]. A dynamic role for the Rieske protein is also emerging from structural studies of the cytochrome bc_{1} complex. Large-scale motion of the entire Rieske soluble domain about a flexible tether to the membrane anchor has been observed in cytochrome bc_{1} [9], apparently in association with the occupancy of the ubiquinol site [8].

We report here the 1.83 Å crystal structure of the soluble domain of the spinach chloroplast Rieske protein by multiwavelength anomalous diffraction (MAD) analysis utilizing the two intrinsic Fe atoms. The structures of the lumen-side domains of the Rieske protein (139 of the 179 residues of the mature protein) and of cytochrome f (252 of 285 residues) constitute virtually the entire extrinsic p-side domain of the cytochrome b_{6f} complex except for connecting loops of the transmembrane α helices of cytochrome b_{6} and subunit IV.

**Results and discussion**

**Structure determination**

The crystal structure of the Rieske soluble domain was determined by MAD from the iron atoms in the [2Fe–2S] cluster and was refined against diffraction data to 1.83 Å. The 2.1 Å electron-density map calculated with MAD phases was of high quality (Figure 2) and was easily interpreted in terms of a model for residues 53–179. The first twelve residues (41–52) of the soluble Rieske fragment are disordered. Residues 41–52 of the proteolytic fragment, including six glycines and two prolines, appear to form a

---

**Figure 1**

Schematic of the membrane topology of a monomer of the cytochrome b_{6f} complex. The 12 transmembrane helices are shown: four associated with cytochrome b_{6} (in pink), three with subunit IV (in yellow), and one each with cytochrome f (in green), the Rieske protein (in red) and the petG, M and L gene products [47–50] (in brown). The inference of one helix for the Rieske protein is based on the structure of the bovine mitochondrial cytochrome bc_{1} complex [6–9].

---
Figure 2

Electron density for the soluble domain of the spinach chloroplast Rieske protein. (a) Experimental map with MAD phases at 2.1 Å resolution. (b) 2F₀−Fᵡ map with phases of the refined model at 1.83 Å resolution. The Gly141-Pro142-Ala143-Pro144-Leu145 proline loop, including a cis-peptide at Pro142, is shown in stereo for both maps; atoms are shown in standard colors. Contours are drawn at the root mean square density level. (Maps were drawn in the program O [43].)
Flexible linker between the soluble domain of the Rieske protein and its N-terminal membrane anchor. The final model of the 127-residue Rieske fragment had an R factor of 17.0%, and no residues with disallowed $\phi/\psi$ values (Table 1). The Rieske soluble domain is the first structure solved by MAD in space group P1 with a centrosymmetric arrangement of anomalous scatterers. This crystallographic special situation presented no problems in MAD phasing.

**Protein structure**

The chloroplast Rieske soluble domain is bilobal with dimensions of approximately $40 \AA \times 30 \AA \times 20 \AA$. The overall fold is dominated by antiparallel $\beta$ secondary structure, with the [2Fe–2S] cluster bound near the top of the molecule, and the N-terminal connection to the flexible linker and missing membrane-anchor peptide near the bottom (Figure 3a). The only $\alpha$ helix, Ala65–Thr71, is at the bottom of the molecule, between $\beta_1'$ and $\beta_2$ (Figure 3b). The protein has two subdomains. The smaller ‘cluster-binding’ subdomain comprises the [2Fe–2S] cluster and residues 105–147, including $\beta_4$–$\beta_8$. The large subdomain includes residues 53–104 and 148–179 (strands $\beta_1$–$\beta_4$, $\beta_9$–$\beta_{10}$ and $\beta_1'$). Each subdomain has the topology of a simple antiparallel $\beta$ barrel such that adjacent $\beta$ strand ‘staves’ of the barrels are also contiguous in the primary sequence. In the large subdomain, the $\beta$ strands form an irregular six-stranded $\beta$ barrel in the order $\beta_1$–$\beta_2$–$\beta_3$–$\beta_4$–$\beta_9$–$\beta_{10}$ with $\beta_1$ and $\beta_{10}$ hydrogen bonded to complete the barrel; strand $\beta_{1}'$ and the $\alpha$ helix cap the bottom end of the barrel. The small cluster-binding subdomain is inserted between $\beta_4$ and $\beta_9$. The long axis of the molecule is spanned by $\beta_4$ with opposite ends of the strand contributing to the $\beta$ barrels of the two subdomains. The cluster-binding subdomain $\beta$ barrel, which is incomplete, is formed by $\beta$ strands in the order $\beta_4$–$\beta_5$–$\beta_6$–$\beta_7$–$\beta_8$. An extended

![Figure 3](image)

_Figure 3_  
Fold of the soluble domain of the spinach chloroplast Rieske protein. (a) Ribbon diagram showing the secondary structure elements and [2Fe–2S] cluster. The $\alpha$ helix is shown in red, $\beta$ strands are in blue and loops are in yellow; Fe and S atoms are shown as red and yellow spheres, respectively. (b) Stereo Co trace of the 127 residues that are ordered in the crystal structure; the [2Fe–2S] cluster is shown as black spheres. The view is the same in (a) and (b). (The figures were drawn in _SETOR_ [51] and _MOLSCRIPT_ [52].)
peptide completes the barrel between β4 and β8 without forming hydrogen bonds to either strand and reconnects the cluster-binding subdomain to the large subdomain.

**[2Fe–2S] cluster**

The [2Fe–2S] cluster is bound by the three, strictly conserved, outermost loops of the cluster-binding subdomain: the β4–β5, β6–β7 and β8–β9 loops (Figure 4). The structure of the cluster, including protein ligands and hydrogen bonds, is identical within experimental error to that of the bovine mitochondrial Rieske protein [5]. Fe ligands are carried on the β4–β5 and β6–β7 loops, which have very similar conformations to the cluster-binding peptides of rubredoxin [24]. The protein Fe ligands are Cys107 and His109 from the β4–β5 loop, and Cys125 and His128 from the β6–β7 loop. His109 and His128 ligate the outermost Fe atom and are otherwise solvent exposed. Cys107 and Cys125 ligate the innermost Fe atom and are buried in the protein. The ligand loops are buttressed by two supporting structures. The first of these is a disulfide bridge between Cys112 and Cys127, analogous to the Cys144–Cys150 disulfide bridge of the bovine mitochondrial Rieske protein. The conservation of the disulfide supports an earlier hypothesis that a disulfide within the Rieske protein contributes to the stability of the cytochrome b₆f₇ complex [25]. On the opposite side of the cluster from the disulfide bridge is the β8–β9 loop, also known as the ‘proline loop’ because it contains the invariant Gly-Pro-Ala-Pro peptide (residues 141–144 in the spinach chloroplast protein and 174–177 in the bovine mitochondrial protein). The pair of histidine ligands is flanked on one side by the sidechain of Leu110, which is adjacent to the disulfide bridge, and on the other side by the sidechain of Pro142 in the proline loop. The [2Fe–2S] cluster is enveloped by the tops of the β strands, the backbones of the ligand and proline loops, and the sidechains of His107, His128, Leu110 and Pro142.

There is a small difference in the redox behavior of the soluble Rieske fragment compared with the protein in the intact cytochrome b₆f complex. The midpoint potential of the soluble fragment at pH 6 is approximately 70 mV more positive than that in the complex [15]. This may be due to the greater solvent exposure of the [2Fe–2S] cluster-binding subdomain in the soluble fragment than in the intact complex.

Despite the sequence identity of the three cluster-binding loops, the peptide bond between Gly141 and Pro142 of the proline loop is in the cis conformation in the crystal structure of the chloroplast protein and in the trans conformation in the mitochondrial protein crystal structure. The observed peptide conformers in both structures are stabilized by intramolecular interactions. In the chloroplast Rieske soluble domain, the cis-peptide carbonyl oxygen is hydrogen bonded through a well ordered water molecule to the backbone carbonyl of His128 and to the sidechain of Arg140. In the mitochondrial Rieske structure, the trans-peptide carbonyl oxygen is directly hydrogen bonded to the sidechain of Arg118, which is part of a 23-residue mitochondrial insertion with no counterpart in the chloroplast Rieske protein. The observed conformations for this peptide are not obviously affected by crystal packing. In neither structure do crystal lattice contacts stabilize the observed conformer, or preclude formation of the other conformer. While there is no evidence for a mixture of states in the crystalline chloroplast protein (Figure 2), cis and trans conformers of both proteins may co-exist in solution.

**Comparison of chloroplast and mitochondrial Rieske proteins**

The three-dimensional structures of the Rieske soluble domains demonstrate that the proteins from the spinach chloroplast (this work) and the bovine mitochondrion [5]
are related. Structure conservation, however, is very unevenly distributed between the cluster-binding and large subdomains. The cluster-binding subdomains are virtually identical (Figure 5a), whereas the large subdomains are substantially different despite the identical folding topology (Figure 5b). Structural alignment and assignment of residue pairs for the chloroplast and mitochondrial large subdomains is difficult due to a different packing of the shorter β strands (β1-β10-β9) on the longer β strands (β2-β3-β4) in the two proteins. Less than half of the residues in the large subdomain are spatially equivalent in the chloroplast and mitochondrial proteins. The segregation of conservation and divergence is also evident in the primary sequences (Figure 6). The relationship of sequence identity to tertiary structural similarity in homologous proteins was quantitated by Chothia and Lesk [24]. The chloroplast and mitochondrial Rieske soluble domains fit this model well when the subdomains are considered separately, and in one molecule represent the extremes of similarity (Table 2).

The identical lengths of the chloroplast and mitochondrial soluble fragments (127 amino acids) also belie a lack of simple one-to-one equivalence of residues from the N to the C terminus. The chloroplast Rieske soluble fragment has an eight-residue insertion between β1 and β2 relative to the mitochondrial fragment, and is longer by 15 residues at the C terminus. The mitochondrial fragment has a 23-residue insertion between β3 and β4 relative to the chloroplast protein. These alignment gaps are modifications on the common architecture of the two proteins. They are responsible for the bilobal shape of the chloroplast domain and the single-lobed shape of the mitochondrial domain. These variations are also associated with the largest differences in the three-dimensional structures. For example, the lone α helix occurs in the internal insertion of each protein. The helices are not spatially equivalent because the internal insertions occur on opposite sides of the large domain.

**Conservation and divergence among Rieske proteins**

Alignment of the three-dimensional structures of spinach chloroplast and bovine mitochondrial Rieske fragments (Figure 5; Table 2) revealed the correct alignment of their primary sequences, which differs somewhat from alignments based on sequence data only. Sequences of other Rieske proteins were aligned to the spinach chloroplast and bovine mitochondrion pair, using the structure-based
Sequence alignment of the soluble domains of Rieske [2Fe–2S] proteins from cytochrome b$_{5}$f and cytochrome bc$_{1}$ complexes. Sequences of b$_{5}$f Rieske proteins are highlighted in green, with dark shading for invariant positions and light shading for sites of conservative substitution. Sequences of bc$_{1}$ Rieske proteins are highlighted in red, with dark and light shading as for the chloroplast proteins. Yellow highlights are for residues invariant in both groups of Rieske proteins. The two groups of sequences were aligned simultaneously using ClustalW [46], with alignment of the spinach chloroplast and bovine mitochondrial proteins constrained by the structure alignment. Spatially equivalent residues are enclosed in boxes. Residue numbering is with respect to the mature spinach chloroplast protein at the top and the bovine mitochondrial protein at the bottom. Secondary structure elements for both proteins are indicated with the residue numbering; α helices are shown as cylinders and β strands as arrows. The program ALSCRIPT [53] was used for display of the sequence alignment.
alignment as a fixed constraint (Figure 6). As has been established by extensive earlier work, most of the sequences segregate into two groups: the \( b_{6f} \) group, which includes the chloroplast photosynthetic proteins from cytochrome \( b_{6f} \) complexes; and the \( bc_1 \) group, which includes the mitochondrial respiratory proteins from the cytochrome \( bc_1 \) complexes. The most important conclusions from the sequence/structure comparison are that conservation between groups is limited to the cluster-binding subdomain and that the large subdomains of the \( b_{6f} \) and \( bc_1 \) groups are very dissimilar. Sequences of the cluster-binding subdomains of members of the \( b_{6f} \) group are 47–61% identical to those of the \( bc_1 \) group. The large subdomains, however, are only 5–18% identical between groups, exactly mirroring the structural comparison (Figure 5; Table 2). Indeed three-dimensional structures are required for accurate sequence alignment when the identity is so low. If the \( b_{6f} \) and \( bc_1 \) Rieske proteins have a common ancestor, as seems likely given their common function and fold, then the cluster-binding subdomains have been highly conserved while the Rieske large subdomains are considerably diverged in cytochromes \( b_{6f} \) and \( bc_1 \).

### The large subdomain and interaction with the electron acceptor

The structure-based sequence alignment reveals that the type of electron acceptor employed by the Rieske protein is strongly correlated with the type of electron acceptor employed by the Rieske protein in cytochrome complexes. Within the \( b_{6f} \) and \( bc_1 \) groups, the large subdomains are highly conserved despite the striking differences between groups (Figure 6). Cytochrome \( f \) is the electron acceptor for the \( b_{6f} \) Rieske proteins, and cytochrome \( c_1 \) for the \( bc_1 \) group. Rieske proteins that function in energy-transducing membrane protein complexes in *Chlorobium*, *Bacillus* and *Sulfobatus* species are dissimilar to both of the \( b_{6f} \) and \( bc_1 \) groups of Rieske proteins, and have neither cytochrome \( f \) nor cytochrome \( c_1 \) as electron acceptors. The cytochrome \( h \) complex of the green photosynthetic bacterium *Chlorobium limicola* does not include any membrane-bound \( c \)-type cytochrome [25]. Similarly, the archaebacterium *Sulfolobus acidocaldarius* produces no \( c \)-type cytochrome [26,27]. The cytochrome \( bc \) complexes of *Bacillus subtilis* [28] and *Bacillus stearothermophilus* [29] do include membrane-bound \( c \)-type cytochromes but these differ from both cytochromes \( f \) and \( c_1 \) and have an N-terminal domain resembling subunit IV of the cytochrome \( b_{6f} \) complexes. In the familiar pattern of segregated conservation and divergence, the cluster-binding subdomains of the *Chlorobium*, *Bacillus* and *Sulfobatus* Rieske proteins are highly similar to those of members of the \( b_{6f} \) and \( bc_1 \) groups (33–58% identity). The large subdomains, however, have such low sequence identity with the \( b_{6f} \) and \( bc_1 \) Rieske proteins (3–13% identity) or with each other (6–12% identity) that the alignments themselves are of dubious accuracy in the absence of three-dimensional structures.

### Distribution of surface charge

The chloroplast and mitochondrial Rieske soluble domains have very different charges at neutral pH. The chloroplast protein has an excess of four acidic amino acid sidechains (14 Asp + Glu residues and 10 Lys + Arg), whereas the mitochondrial protein has an excess of two basic sidechains (14 Asp + Glu residues and 16 Lys + Arg). The \( pI \) of the chloroplast Rieske soluble domain (4.9, as measured by isoelectric focusing; data not shown) is more acidic than the mitochondrial protein (\( pI = 6.1 \), as measured electrochemically; [16]). Anionic regions are prominent on the surface of the chloroplast protein (Figure 7a), whereas the mitochondrial protein has several basic surface regions and a more globular shape (Figure 7b). The substantial charge difference, however, is offset by the substantial difference in acidity of the compartments where the chloroplast and mitochondrial Rieske domains function in vivo. The chloroplast lumen is two to three pH units more acidic than the mitochondrial intermembrane space. Thus the average surface charge of the two Rieske proteins may be similar in situ.

The distribution of surface charge has relevance to the interactions of the Rieske soluble domain with its redox partners and with other proteins of the membrane protein complex. Thus the different distribution of surface charge on the chloroplast and mitochondrial Rieske soluble domain is consistent with the different functions in vivo.
domains may reflect their interaction with the very different electron acceptors cytochrome $f$ and cytochrome $c_1$. In both the chloroplast and the mitochondrial Rieske soluble domains, a relatively nonpolar surface surrounds the [2Fe–2S] cluster, and is derived from the strictly conserved ligand-binding loops and the proline loop. In recently described preliminary crystal structures of cytochrome $bc_1$ [6–9], the protein surface of the cluster-binding subdomain interacts with cytochrome $b$ and forms part of the binding niche of the ubiquinol/ubiquinone couple that is the electron donor to the Rieske protein. On the basis of the similar lack of polarity and the conservation of this region of the chloroplast Rieske protein, we conclude that the cluster-binding subdomain of the chloroplast Rieske protein interacts with cytochrome $b_6$ and the plastoquinol/plastoquinone-binding site of cytochrome $b_6f$ in a similar manner.

The Rieske protein contacts the membrane near the N terminus of the soluble domain. The location of this terminus is uncertain because of the disorder of the N-terminal 12 residues. A nonpolar surface might be expected for any part of the Rieske protein that interacts with the interfacial region of the membrane. The lack of such a surface, excepting the cluster-binding subdomain discussed above, argues against the interaction of the soluble domain with hydrophobic regions of the membrane. This is consistent with the participation of the soluble domain of the Rieske protein of cytochrome $bc_1$ in only protein–protein interactions with cytochrome $b$ and not protein–membrane interactions. Thus the Rieske protein may behave as a soluble protein, which is prevented from diffusing away from the membrane by its long, flexible tether.

**The N-terminal flexible tether**

Residues 41–52 of the spinach chloroplast Rieske soluble fragment tether residues 53–179, whose structure we describe here, to the membrane anchor. The 12-residue tether peptide is extremely flexible, as evidenced by its total disorder in the crystal structure, by the preponderance
of glycine residues in the peptide (six of twelve amino acids), and by the accessibility of residue 41 to the protease thermolysin when the soluble fragment is prepared from the intact cytochrome $b_{6}/f$ complex [15]. A long, flexible tether for the Rieske protein is consistent with the emerging picture of a large-scale motion of the Rieske protein associated with electron transfer and energy transduction in the cytochrome $b_{6}$ complex. The Rieske protein occupies different positions in three different crystal forms of cytochrome $b_{6}$ complexes, dependent on the binding of the drug stigmatellin to the $p$-side quinol site of the complex, whereas the distances between the cytochrome $c_{1}$ heme Fe and the cytochrome $b$ heme Fe atoms are constant [6–9]. Efficient electron transfer by the Rieske protein is a problem given the apparently fixed, greater than 30 Å distance from the quinol-binding site to the cytochrome $c_{1}$ redox center in cytochrome $b_{6}$. A large-scale motion by a flexibly tethered Rieske $[2Fe–2S]$ domain is one possible solution. A similar situation may occur in cytochrome $b_{ history. The heme Fe of cytochrome $f$ is distant (45 Å) from its membrane anchor with limited space in the sequence for a flexible tether between the soluble domain and the membrane tether [3].

Role of aromatic residues in intraprotein pathways of electron transfer

Current discussions of the pathways of electron transfer through proteins focus on whether the pathways are independent of protein structure [20], or sensitive to local structure, occurring with decreasing efficiency through covalent bonds, hydrogen bonds and space [21,30]. A role for the $\pi$-electron systems of aromatic amino acids in electron transfer through proteins has long been discussed [31–37], although there currently exists no example of electron transfer through a protein facilitated by a chain or cluster of aromatic sidechains.

The spinach chloroplast Rieske protein includes a large number of conserved aromatic residues (Figure 6). Nine of the conserved aromatic sidechains form two clusters in the protein, suggesting that they may function in conductive pathways of electron transfer. The first cluster includes Tyr132, Phe123 and Tyr89, which are conserved in all Rieske protein soluble domains, and Phe116. The second cluster, including Trp176, Trp177, Trp164, Phe101 and Phe169, is conserved in all $b_{6}/f$ Rieske proteins, but has no counterpart in the $b_{6}$ Rieske proteins. Within each cluster, aromatic sidechains are in van der Waals contact. The invariant Tyr132 of the first cluster is connected to the $[2Fe–2S]$ cluster via a hydrogen bond to the Fe ligand Cys107 $S_{y}$. Based on results obtained with model aromatic compounds [32,33], it can be hypothesized that either of these aromatic clusters could greatly facilitate electron transfer within the Rieske protein, because each cluster traverses a total distance of approximately 15 Å. Other factors, however, argue against this interpretation. The second cluster is approximately 12 Å from the $[2Fe–2S]$ cluster, and even further when through-bond pathways are considered. In addition, mutagenesis experiments with the yeast Rieske protein do not implicate the conserved residues of the first aromatic cluster in electron transfer function (BL Trumpower, personal communication). Thus, it is not possible at this time to infer a role for these conserved clusters of aromatic residues in electron transfer through the Rieske protein.

Pathways of proton transfer

Proton translocation by the cytochrome $b_{6}/f$ complex is tightly coupled to electron transfer. The source of protons is plastoquinol, which is deprotonated in conjunction with its oxidation. A simple mechanism for coupling electron transfer and proton translocation was proposed to involve proton transfer to a buried water chain in cytochrome $f$ concomitant with reduction of its heme Fe [4]. The most obvious proton source is the cytochrome $f$ reductant, the Rieske protein. The one set of residues in the Rieske protein that suggest themselves as candidate proton carriers are the histidine ligands His109 and His128. It is not at all obvious that the pK of either histidine sidechain is appropriate for this function. However, the redox titration of the Rieske soluble fragment has a pK of 6.5 [15], which is most readily associated with a sidechain near the $[2Fe–2S]$ cluster. The structure of the chloroplast Rieske protein soluble domain includes no bound internal water chain analogous to the buried water chain in cytochrome $f$. Nor is there an obvious hydrogen-bonded network of protein groups through the chloroplast Rieske protein. If the water chain in cytochrome $f$ is used in proton translocation, then it does not connect with a similar structure in the Rieske protein. If there is an intraprotein proton pathway involving the Rieske protein that would ultimately donate protons to cytochrome $f$, then it would seem to involve the interface between cytochrome $f$ and the Rieske protein during electron transfer.

Conservation and diversity in cytochrome $b_{6}/f$ and cytochrome $b_{6}$

The structure of the chloroplast Rieske soluble domain together with other results suggests a striking separation of structural conservation and diversity at the membrane surface in the cytochrome $b_{6}/f$ and $b_{6}$ complexes. The energy-transducing $b_{6}/f$ and $b_{6}$ complexes are very similar within the membane bilayer and very different in the $p$-side aqueous phase. The cytochrome $b_{6}$–subunit IV tandem pair of cytochrome $b_{6}/f$ has a similar sequence, and by inference a similar structure, throughout its length to cytochrome $b$ of the cytochrome $b_{6}$ complex, although these proteins differ in some important biochemical details. This is in marked contrast to the soluble domains on the $p$-side of the membrane. Conservation and divergence are sharply delineated in the Rieske soluble domain, possibly at the membrane border. The invariant cluster-binding
subdomain may be conserved to interact with the similar \(b\)-type cytochromes and plastoquinol/ubiquinol-binding sites within the membrane bilayer. In contrast, the Rieske large subdomains are very different and may have evolved to pair with the unrelated membrane-bound \(c\)-type cytochromes \((f\) and \(c_1)\) in the \(p\)-side aqueous phase of chloroplasts and mitochondria.

The difference in conservation of the intra- and extramembrane components complicates transfer of experimental results from one complex to the other. Little common ground exists for structural comparison of redox complexes between the unrelated \(c\)-type cytochromes and their divergent Rieske protein partners. The fundamentally different electrostatic potential surfaces of the chloroplast and mitochondrial Rieske proteins (Figure 7) reflect this fact as well as the large difference in acidity of their \textit{in vivo} compartments.

**Biological implications**

The integral membrane cytochrome \(b_{6f}\) complex transfers electrons between photosystems II and I and contributes to the proton gradient of the energy-transducing photosynthetic membrane. Together with the previously determined structure of the lumen-side domain of cytochrome \(f\), the 1.83 Å crystal structure of the soluble domain of the spinach chloroplast Rieske [2Fe–2S] protein, described in the present work, essentially completes an atomic description of the cytochrome \(b_{6f}\) complex on the electropositive (\(p\))-side of the membrane. Structures of the lumen-side domains of the chloroplast Rieske protein and cytochrome \(f\) suggest possibilities for intra- and interprotein pathways of electron and proton transfer in cytochrome \(b_{6f}\) that are distinct from those of the related respiratory cytochrome \(b_{1}\) complex. The interaction between the Rieske and cytochrome \(f\) redox partner domains in the chloroplast lumen remains to be determined.

In general, elucidation of the structure of an integral membrane protein with extended peripheral domains is greatly facilitated by the determination of accurate structures for the isolated soluble domains. Comparison of structures and sequences of the soluble domains of chloroplast and mitochondrial Rieske proteins, together with preliminary structural data for the mitochondrial complex, has already highlighted important differences in the protein subunits of the cytochrome \(b_{6f}\) and \(b_{1}\) complexes. The complexes are highly similar for domains within the membrane bilayer and dissimilar in domains that function on the \(p\)-side of the membrane. Diversity is most dramatic in the unrelated \(p\)-side cytochrome \(f\) and \(c_1\) subunits. Segregation of conservation and divergence is especially evident in the Rieske protein. The Rieske [2Fe–2S] cluster-binding subdomains, which interact with electron donors within the bilayer, are virtually identical in the chloroplast and mitochondrial complexes. The Rieske large subdomains, however, appear to have evolved to interact with unrelated redox partners in the \(p\)-side aqueous phase, and are dissimilar in the chloroplast and mitochondrial complexes.

**Materials and methods**

**Crystallization**

The Rieske soluble domain was prepared and crystallized as in Zhang et al. [15]. The Rieske soluble domain was crystallized by the hanging-drop vapor diffusion method with a reservoir solution of 100 mM sodium acetate (\(pH\) 4.6), 100–200 mM ammonium acetate and 30% polyethylene glycol (PEG)-4000. Hanging drops were a 1:1 mixture of protein solution (30 mg/ml Rieske soluble domain, 20 mM morpholino-propanesulfonic acid (MOPS) buffer, \(pH\) 7.2) and reservoir solution.

**Data collection**

Crystals were stabilized in crystallization reservoir solution and cryoprotected in 75 mM sodium acetate (\(pH\) 4.6), 150 mM ammonium acetate, 30% PEG-4000 and 25% ethylene glycol. After a short cryoprotection period (<5 min), crystals were removed in fiber loops and immediately flash-frozen in liquid ethane. MAD data were collected at three wavelengths on Beamline 19 (IIBM14) at the European Synchrotron Radiation Facility (ESRF), Grenoble, France. X-ray fluorescence from a Rieske protein crystal was recorded from 6.93 keV to 8.36 keV to select wavelengths for data collection. Wavelengths were selected at the peak (maximum \(f\)’, 7.130 keV, 1.7389 Å) and the inflection point (minimum \(f\)’, 7.118 keV, 1.7419 Å) of the Fe K edge and at one remote wavelength (8.266 keV, 1.5000 Å). The remote energy was unusually distant from the Fe K edge in order to obtain high angle, high quality data for eventual model refinement. Severe air absorption at the low energy Fe K edge limits the quality of data at the edge energies. Data were measured as successive 360° sweeps at the three wavelengths (72 images of 5° each) from one frozen crystal using an 18 cm MARsearch image-plate detector. Diffraction images were processed using the HKL program package [38,39] and scaled using SCALA and PHASES. MAD data were collected at three wavelengths were put on an approximately fully merged data at \(\lambda = 1.5000\) Å. Data quality is summarized in Table 1.

The crystals belong to the space group P1 with cell parameters \(a = 29.05\) Å, \(b = 31.87\) Å, \(c = 35.79\) Å, \(\alpha = 95.6^\circ\), \(\beta = 106.1^\circ\) and \(\gamma = 117.3^\circ\). The asymmetric unit contains one molecule of the Rieske soluble domain, with a calculated solvent content of 31%.

**MAD structure determination**

The data at the three wavelengths were put on an approximately absolute scale using SCALEIT from the CCP4 suite. The positions of the two Fe atoms were determined by inspection of an anomalous difference Patterson map from data at the wavelength of maximum \(f\)’ (1.7389 Å). Refinement of the Fe partial structure and MAD phasing were performed by the pseudo-isomorphous approach using MLPHARE [41]. The data at the edge-remote energy (\(\lambda = 1.5000\) Å) were designated as the ‘native’ data, and data at the other two wavelengths as ‘derivatives’. Phasing statistics are summarized in Table 3. Phases were refined by density modification in the form of solvent flattening and histogram matching using the program DM [42] with a 30% solvent mask. The real-space free residual decreased from 0.273 to 0.161, and the interpretability of the electron-density map improved somewhat.

**Model building and refinement**

The atomic model was constructed using the program O [44] and refined against the higher quality, more extensive data at \(\lambda = 1.5000\) Å using XPLOR [44,45]. The initial model of 127 amino acids had an R work of 0.349 and R free of 0.376 for data from 20.0 Å to 2.1 Å. After
Table 3

MAD data collection and phasing.

<table>
<thead>
<tr>
<th></th>
<th>(\lambda = 1.7389,\text{Å})</th>
<th>(\lambda = 1.7419,\text{Å})</th>
<th>(\lambda = 1.5000,\text{Å})</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. observations</td>
<td>22,187</td>
<td>18,692</td>
<td>28,882</td>
</tr>
<tr>
<td>No. unique reflections</td>
<td>5633</td>
<td>5603</td>
<td>8758</td>
</tr>
<tr>
<td>dsym (Å)</td>
<td>2.1</td>
<td>2.1</td>
<td>1.83</td>
</tr>
<tr>
<td>Redundancy</td>
<td>3.9</td>
<td>3.3</td>
<td>3.3</td>
</tr>
<tr>
<td>Completeness (%)</td>
<td>91.6</td>
<td>91.1</td>
<td>89.7</td>
</tr>
<tr>
<td>(R_{sym}) (%)</td>
<td>5.2</td>
<td>4.3</td>
<td>4.8</td>
</tr>
<tr>
<td>(R_{calc}) (%)</td>
<td>(8.5)</td>
<td>(7.2)</td>
<td>(10.8)</td>
</tr>
<tr>
<td>Phasing power‡</td>
<td>1.76</td>
<td>2.30</td>
<td></td>
</tr>
</tbody>
</table>

\(R_{sym} = \frac{\sum_{i} |I_{o,i} - I_{c,i}|}{\sum_{i} I_{c,i}}\). Phasing power \(= (\langle F_o \rangle/d)\) lack of closure rms isomorphous difference. 

References

16. Link, T.A., Hagen, W.R., Pleier, A.J., Assmann, C. & von Jagow, G. (1992). Determination of the redox properties of the Reseke \(2Fe-2S\) cluster of bovine heart cytochrome \(bc_1\) complex and 143 water molecules. Model quality is summarized in Table 1. Least-squares superpositions of the spinach chloroplast protein and its bovine mitochondrial counterpart [5] were performed in the program O. Sequence alignment was done with the program ClustalW [46]. Acceision numbers Coordinates for the model (1ufs) and the MAD data (1ufsasf) have been deposited in the Brookhaven Protein Data Bank. Acknowlegments The studies described in this manuscript were supported by USDA (US Department of Agriculture) grant 95-37306-2045 to JLS and WAC and by NIH grant GM-18457 to WAC. This work was supported by an NIH Biophysics Predctoral Training Grant. We thank AW Thompson for expert advice at ESRF beamline 19, JM Krahn, HB Gray and BL Trumpower for helpful discussions and JL Hollister for skillful and dedicated assembly of the text. References


