Structure of the Soluble Domain of Cytochrome f from the Cyanobacterium Phormidium laminosum † , ‡

ARTICLE *in* BIOCHEMISTRY · AUGUST 1999

Impact Factor: 3.02 · DOI: 10.1021/bi9903190 · Source: PubMed

Structure of the Soluble Domain of Cytochrome *f* from the Cyanobacterium *Phormidium laminosum*[†,‡]

Christopher J. Carrell,[§] Beatrix G. Schlarb,[||] Derek S. Bendall,[||] Christopher J. Howe,[||] William A. Cramer,[§] and Janet L. Smith*[,§]

Department of Biological Sciences, Purdue University, West Lafayette, Indiana 47907, and Department of Biochemistry and Centre for Molecular Recognition, University of Cambridge, Tennis Court Road, The Downing Site, Cambridge CB2 1QW, U.K.

Received February 9, 1999; Revised Manuscript Received May 17, 1999

ABSTRACT: Cytochrome *f* from the photosynthetic cytochrome b_6f complex is unique among *c*-type cytochromes in its fold and heme ligation. The 1.9-Å crystal structure of the functional, extrinsic portion of cytochrome *f* from the thermophilic cyanobacterium *Phormidium laminosum* demonstrates that an unusual buried chain of five water molecules is remarkably conserved throughout the biological range of cytochrome *f* from cyanobacteria to plants [Martinez et al. (1994) *Structure 2*, 95−105]. Structure and sequence conservation of the cytochrome *f* extrinsic portion is concentrated at the heme, in the buried water chain, and in the vicinity of the transmembrane helix anchor. The electrostatic surface potential is variable, so that the surface of *P. laminosum* cytochrome *f* is much more acidic than that from turnip. Cytochrome *f* is unrelated to cytochrome c_1, its functional analogue in the mitochondrial respiratory cytochrome bc_1 complex, although other components of the b_6f and bc_1 complexes are homologous. Identical function of the two complexes is inferred for events taking place at sites of strong sequence conservation. Conserved sites throughout the entire cytochrome b_6f/bc_1 family include the cluster-binding domain of the Rieske protein and the heme *b* and quinone-binding sites on the electrochemically positive side of the membrane within the *b* cytochrome, but not the putative quinone-binding site on the electrochemically negative side.

The cytochrome b_6f membrane protein complex of oxygenic photosynthesis couples the transfer of electrons between the photosynthetic reaction centers to proton translocation and generation of the proton electrochemical potential utilized for ATP synthesis (*1*, recent review). These functions are analogous to those carried out by the cytochrome bc_1 complex of the respiratory electron transport chain and of purple photosynthetic bacteria (*2*). The *b*-type cytochromes in these complexes are polytopic integral membrane proteins that contain two noncovalently bound hemes, one located toward the electrochemically negative side (*n*)[1] and one toward the electrochemically positive (*p*) side of the membrane (*3*−*8*). The *c*-type cytochromes in these complexes, cytochromes *f* and c_1, have a single covalently bound heme located in a *p*-side extrinsic domain, which is anchored by a single C-terminal transmembrane helix (*9*). The Rieske Fe_2S_2 proteins in these complexes also have a *p*-side extrinsic functional domain with a single N-terminal transmembrane helix.

The analogous functions of the b_6f and bc_1 complexes are thought to be a consequence of a close evolutionary relation between all or most of their redox-active subunits (*10*−*12*). From comparisons of primary sequence, distribution of hydrophobic domains, absolute hydrophobic character, and the motif of interhelix bis-histidine heme coordination, there is little question that the cytochrome *b* polypeptide of the bc_1 complexes, with eight transmembrane helices, shares a common ancestor with the smaller cytochrome b_6 and subunit IV of the b_6f complexes. Cytochrome b_6, with four transmembrane helices, corresponds to the N-terminal half of the bc_1 cytochrome *b* subunit, and subunit IV to its C-terminal half (*3*). The Rieske proteins of the b_6f and bc_1 complexes are more diverged than are the *b*-type cytochromes, although the Rieske subdomains that bind the Fe_2S_2 cluster are virtually identical (*13*, *14*). The *c*-type cytochromes *f* and c_1 are unrelated (*6*, *15*). Biochemical analogies, routinely made between the b_6f and bc_1 complexes, must take into account this patchwork of conservation and variability.

It appears that the hydrophobic subunits of these integral membrane cytochrome complexes are highly conserved and much more closely related than are the subunits with extensive peripheral and exposed domains. The presence of *a*- and *b*-type and absence of *c*-type cytochromes in the thermoacidophilic archaeon *Sulfolobus acidocaldarius* (*16*)

[†] Supported by USDA Grant 98-35306-6405 to J.L.S. and W.A.C. B.G.S. was supported by the Biotechnology and Biological Scences Research Council and the Deutscher Akademischer Auslandsdienst (Program HSPIII).

[‡] Coordinates have been deposited in the Protein Data Bank under access code 1ci3.

* To whom correspondence should be addressed. Phone: (765) 494-9246. Fax: (765) 496-1189. E-mail: smithj@bragg.bio.purdue.edu.

[§] Purdue University.

[||] University of Cambridge.

[1] Abbreviations: MES, 2-(*N*-morpholino)ethanesulfonic acid; *n*, *p*, electrochemically negative and positive sides of the membrane; PDB, Protein Data Bank; PEG, poly(ethylene glycol); Q_n, Q_p, quinone-binding sites on the *n* and *p* sides of the membrane; RMSD, root-mean-square deviation.

10.1021/bi9903190 CCC: $18.00
Published on Web 07/02/1999

suggests that the *c*-type cytochromes may have arisen after the divergence of archaebacteria and other organisms. Integral membrane proteins that are protected from the environment may have been more stable in the primitive extreme environments. Subunits with extensive protein mass and redox centers in the bulk aqueous phase may have arisen later when the aqueous environment was less harsh. On the other hand, Rieske-type proteins, which have an extensive peripheral *p*-side domain in mitochondria (*13*) and chloroplasts (*14, 17*), are present in *S. acidocaldarius (18, 19*).

A crystal structure of cytochrome *f* from turnip chloroplasts revealed several features unique among the *c*-type cytochromes (*15, 20*). The extrinsic portion of cytochrome *f* has two β-sheet domains arranged to form an elongated structure. The heme is bound in the larger domain with the N-terminal α-amino group as an Fe ligand. A chain of five water molecules is buried within the large domain, in contact with the His Fe ligand. The structure of the soluble portion of cytochrome *f* from the b_6f complex of the green alga *Chlamydomonas reinhardtii* has also been determined (*21*), and a preliminary 3.5-Å structure has been presented for this domain from the cyanobacterium *Spirulina maxima* (*22*). Properties of cytochrome *f* from the moderately thermophilic cyanobacterium *Phormidium laminosum* have been described (*23*). Here we report the crystal structure of the 250-residue soluble portion of cyanobacterial cytochrome *f* from the b_6f complex of *P. laminosum*. Cyanobacteria and plants represent the extent of biological divergence of cytochromes *f*. The high degree of similarity of cytochromes *f* from the cyanobacterium and higher plant chloroplasts implies that structural details have been preserved over approximately 3 billion years, from the early origins of energy-transducing membranes.

MATERIALS AND METHODS

Expression and Purification of P. laminosum Cytochrome f. P. laminosum cytochrome *f*, truncated after Arg250, was expressed under semi-anaerobic condidtions in *E. coli* W3110 transformed with the plasmid pUC19CF, and purified from periplasmic extracts (*24*). However, as the yield was low, two additional steps of purification were inserted. Pooled cytochrome *f* fractions from the first column (Whatman DE52) were fractionated with ammonium sulfate. The supernatant was brought to 60% saturation and applied to a Sepharose CL4B hydrophobic interaction column (2.5 cm diameter $\times$ 20 cm; Sigma) preequilibrated with 65% ammonium sulfate, pH 7, at 4 °C. The protein was eluted with a gradient (400 mL) of 60−30% ammonium sulfate, pH 7, and subjected to further chromatography (*24*). The final yield from 40 L of culture was about 7 mg of cytochrome *f* with an absorbance ratio (A_{280}/A_{556}) of 1.0 for the reduced cytochrome.

Crystallization and Data Collection. The purified protein was crystallized by the sitting drop vapor diffusion method using a 1:1 ratio of cytochrome *f* to reservoir solution. The protein solution was 10 mg/mL cytochrome *f* in 50 mM Tris buffer, pH 7.5, 50 mM NaCl, and trace ascorbic acid. The reservoir solution was 100 mM 2-(*N*-morpholino)ethanesulfonic acid (MES, pH 6.5), 200 mM zinc acetate, and 8−12% poly(ethylene glycol) (PEG) 8000 (Fluka). Crystals were pink triangular prisms approximately 0.45 mm on a side and 0.1 mm thick. A crystal was harvested into a solution of 100 mM MES (pH 6.5), 200 mM zinc acetate, and 10% PEG-8000. Cryoprotection was achieved by transferring a crystal through harvesting solutions with successively higher concentrations of glycerol in increments of 4%. The final glycerol concentration of 20% was reached in a total cryoprotection time of about 20 min. The crystal was removed immediately with a loop and flash-frozen in an Oxford Cryosystems cryostream.

Table 1: Summary of Diffraction Data

wavelength (Å)	1.5418
no. of measured reflections	126348
no. of unique reflections	26071
minimum *d*-spacing (Å)	1.9
average redundancy	4.8
% completeness	97.9 (83.8)[a]
R_{sym}[b] (%)	3.2 (12.8)

[a] Values in parentheses pertain to the outermost shell of data (1.97−1.90 Å). [b] $R_{sym} = \Sigma_{h,i}|I_{h,i} - \langle I_h\rangle|/\Sigma_{h,i}I_{h,i}$.

Diffraction data were measured using a rotating anode source (Rigaku RU-200) and a Molecular Structures Corp. R-Axis IV image plate system. Diffraction extended to a minimum *d* spacing of 1.9 Å in the space group *R32* with unit cell parameters $a = b = 109.6$ Å, $c = 145.1$ Å, $\alpha = \beta = 90°$, $\gamma = 120°$ in the hexagonal setting. The asymmetric unit includes one cytochrome *f* molecule and a solvent content of approximately 60% (v/v). The diffraction images were processed and scaled using the HKL program suite (*25*). The data quality is summarized in Table 1.

Structure Determination and Refinement. The structure of *P. laminosum* cytochrome *f* was solved by the molecular replacement method with AMORE (*26*), using the complete model of turnip cytochrome *f* (PDB code 1hcz; *20*) as a probe structure. The orientation and position of the molecule in the *R32* unit cell were found readily, despite a 4° difference in interdomain hinge between the unknown and probe structures. The starting *R*-factor was 43.0% for data in the range 20.0−4.5 Å. Rigid-body refinement of the complete polypeptide, then of the two domains, and finally of individual secondary structure elements was carried out using X-PLOR (*27*). The model was modified to reflect the sequence of the *P. laminosum* cytochrome *f*, yielding an *R*-factor of 44% for data to 3.0 Å.

The model was refined with one round of simulated annealing followed by alternating cycles of atomic refinement in X-PLOR and model-building in the program O (*28*). The CCP4 program suite (*29*) was used for the calculation of the $|2|F_o| - |F_c||$ and $||F_o| - |F_c||$ maps used in model building. A large electron density peak at a crystal lattice contact near the His55 side chain was interpreted as a zinc ion from the crystallization solution. The zinc atom was coordinated by His55, Glu125, Asp54 from a symmetry-related molecule, and one water molecule. Coordination stereochemistry could not be restrained across crystal lattice contacts in X-PLOR, so two cytochrome *f* molecules were refined in space group *R3*. Diffraction data for space group *R3* were obtained by rescaling the original unmerged data. The refinement test set of reflections was obtained by expansion of the *R32* test set. The model, including water molecules, was restrained to obey *R32* symmetry during

Table 2: Refinement Statistics for *P. Laminosum* Cytochrome *f*

residue range	1−249
no. of non-H protein atoms	2182
no. of waters	250
data range (Å)	30−1.9
R-factor[a] (%)	20.3
no. of selected reflections	23596
free-*R* factor (%)	23.5
no. of selected reflections	1993
data cutoff	$\|F\| > 0.0$
average *B* (Å^2)	
main chain	26.9
side chain	28.9
water	38.4
heme and Zn	13.9
all atoms	28.7
RMS deviations from target values	
bond lengths (Å)	0.010
bond angles (deg)	1.703
bonded *B* (Å^2)	1.53
Ramachandran outliers	none
cis peptides	Gly117−Pro118

[a] $R = \Sigma_h|F_{obs,h} - F_{calc,h}|/\Sigma_h F_{obs,h}$.

refinement in space group *R3*. After completion of refinement in space group *R3*, the model was converted back to *R32* symmetry for final refinement, during which the zinc ligands did not move appreciably. A second zinc ion was coordinated to the His150 side chain at a site of half-occupancy, which was not at a crystal lattice contact. The final model refinement was done with the program CNS (*30*) using all data between 30.0 and 1.9 Å. Statistics for the final model are summarized in Table 2.

To determine appropriate restraints for metal−ligand coordination distances for the zinc sites and for the axial iron ligands, tetrahedral complexes of zinc with imidazole or carboxylate functional groups, and octahedral complexes of iron with primary amine or imidazole ligands, were extracted from the Cambridge Structural Database. The metal−ligand coordination distances were restrained to the median bond distance for each chemical type. These distances are as follows: Zn−N, 2.05 Å; Zn−O, 2.10 Å; Fe−imidazole, 2.11 Å; Fe−NH_2, 2.15 Å.

Structure and Sequence Analysis. Structure comparisons were done using the superposition routines in the program O. Most of the structure comparisons were based on a superposition of 188 Cα atoms in the large domain (residues 1−169 and 231−249). The difference in interdomain hinge angle was determined by superposition of small domains (residues 170−230) in molecules with large domains previously superimposed. Common water sites were also defined based on the large-domain Cα superposition, using a cutoff distance of 1.5 Å, approximately half the interatomic separation of hydrogen-bonded waters. Similarity searches of the structure database were done with the program DALI (*31*), and of the sequence database with BLAST (*32*).

For sequence comparisons of cytochrome b_6 and subunit IV with cytochrome *b*, all 21 deposited cytochrome b_6 sequences and 22 subunit IV sequences were used. Among the several hundred deposited cytochrome *b* sequences, 24 were selected that represent the biological range of cytochrome *b*, including bacteria, fungi, algae, protists, insects, vertebrates, and plants. The program CLUSTALW (*33*) was used to align the cytochrome b_6 sequences with the N-terminal portion of cytochrome *b*, and the subunit IV sequences with the C-terminal portion. A total of 30 invariant residues were found in the 45-sequence, 215-residue cytochrome b_6−cytochrome *b* alignment (Pro25, Gly35, Gly49, Tyr56, Ala63, Ser66, Gly77, His84, Ala88, Ser89, His98, Arg101, Trp114, Gly117, Gly131, Tyr132, Leu134, Pro135, Gln138, Trp142, Ala153, Pro155, Gly158, Thr175, Leu176, Arg178, His183, Pro187, His197, Pro209), and 14 in the 46-sequence, 160-residue subunit IV−cytochrome *b* alignment (Pro248, Pro262, Thr265, Pro266, Ile269, Pro271, Glu272, Trp273, Tyr274, Leu282, Arg283, Lys288, Gly291, Pro306). Residue numbering is according to the chicken cytochrome *b* sequence.

RESULTS AND DISCUSSION

Structure. The crystal structure of *P. laminosum* cytochrome *f* in the reduced form was solved by molecular replacement using the 62% identical turnip cytochrome *f* as a probe (*20*). The 1.9-Å model, with $R_{work} = 20.3\%$ and $R_{free} = 23.5\%$, includes 249 amino acids, 1 heme, 250 water molecules, and 2 Zn ions. The polypeptide of the expressed extrinsic domain of cytochrome *f* is well ordered throughout its length (Figure 1) with the exception of residue 250, for which there is no electron density. In the intact cytochrome, residue 250 connects the extrinsic domain to the membrane-anchor helix. All residues but one are in the most favorable regions of the Ramachandran plot. Asp199, which is in a loop and has poor electron density, has ϕ/ψ angles in the generously allowed region. The model fits stereochemical quality criteria well. The mean error in atomic positions is 0.11 Å, estimated by the SIGMAA method (*34*), and 0.21 Å by the Luzatti method (*35*), using data from 5 to 1.9 Å.

The extrinsic portion of cyanobacterial cytochrome *f* is an elongate structure dominated by β-secondary structure (Figure 2), as for the turnip cytochrome (*15*, *20*). Both the heme and the C-terminal membrane anchor are linked to the larger of two structural domains (residues 1−169 and 231−249). The small domain (residues 170−230) and the membrane anchor are at opposite ends of the large domain. The heme is bound by the N-terminal 25 residues, which include the signature heme-binding peptide, Cys-X-Y-Cys-His, (residues 21−25) of *c*-type cytochromes. The heme iron is ligated by His25 and by the α-amino group of Tyr1.

Comparison of Structures of P. laminosum and Turnip Cytochrome f. The two structural domains of the plant cytochrome *f* were superimposed separately on the corresponding domains of cyanobacterial cytochrome *f* (Figure 3a). Root-mean-square deviations (RMSDs) are equal in the two domains, 0.82 Å for 188 Cα atoms in the large domain and 0.82 Å for 61 Cα atoms in the small domain. Sequence identity is significantly higher in the large domain (69%) compared to the small domain (40%). The correlation between overall 62% sequence identity and RMSD of Cα positions is within the expected range (*36*). The angle between the two domains is 4.5° smaller in *P. laminosum* cytochrome *f* than in the turnip cytochrome. This difference may reflect a flexible hinge between the domains, or it may be a small difference between the *P. laminosum* and turnip proteins. These possibilities cannot be distinguished because crystals of the *P. laminosum* and turnip cytochromes each provide only one view of the protein molecule. However, a

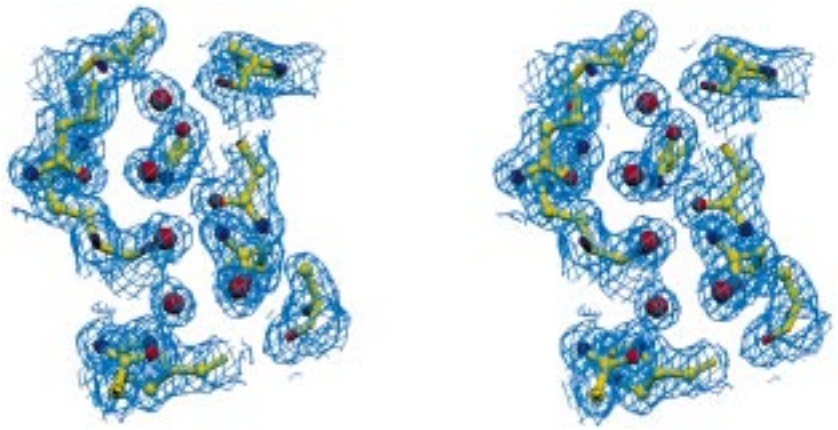

FIGURE 1: Electron density at 1.9-Å resolution and the refined model for the soluble fragment of *P. laminosum* cytochrome *f*. A portion of the buried water chain and some of the surrounding amino acids are shown in the stereo diagram. Carbon atoms are yellow, nitrogen blue, and oxygen red. Contours in the $|2|F_o| - |F_c||$, α_{calc} map are drawn in blue at the root-mean-square density level. The map was drawn in the program SETOR (*56*).

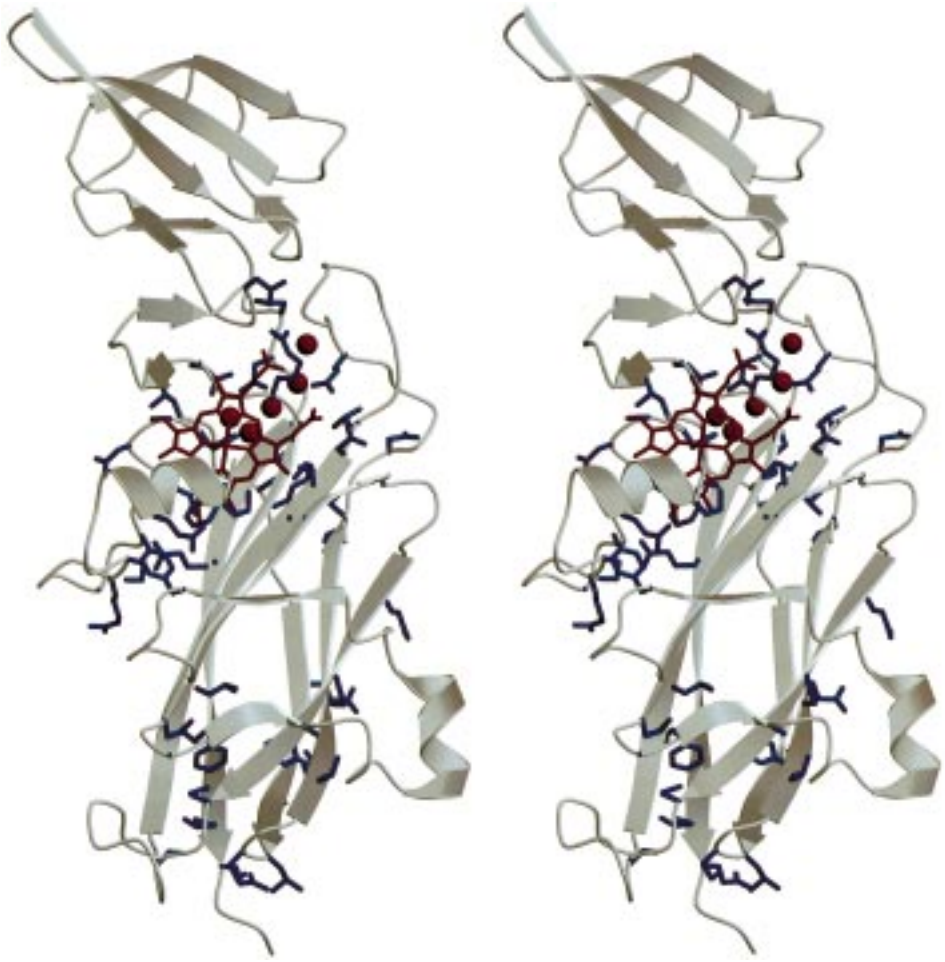

FIGURE 2: Stereo ribbon diagram of *P. laminosum* cytochrome *f*. The heme cofactor and buried water chain are drawn in red. Conserved side chains among 22 cytochrome *f* sequences are shown in blue. Invariant residues cluster around the water chain, the heme, and near the C-terminal connection to the transmembrane helix at the bottom of this figure. The figure was prepared with MOLSCRIPT (*57*) and RASTER3D (*58*).

small amount of flexibility is clear in the algal cytochrome *f* from *C. reinhardtii*, demonstrated by three different hinge angles for the three copies of the cytochrome in the crystal structure (*21*). Based on this result, we expect that the connection between large and small domains has this small degree of flexibility in most or all cytochromes *f*.

Buried Water Chain. A rare hydrogen-bonded chain of five water molecules buried within turnip cytochrome *f* (*20*) is conserved to a remarkable degree in the cyanobacterial cytochrome. The ordered water structure was determined independently for the cyanobacterial cytochrome *f* crystal structure. Peaks for the buried waters were among the

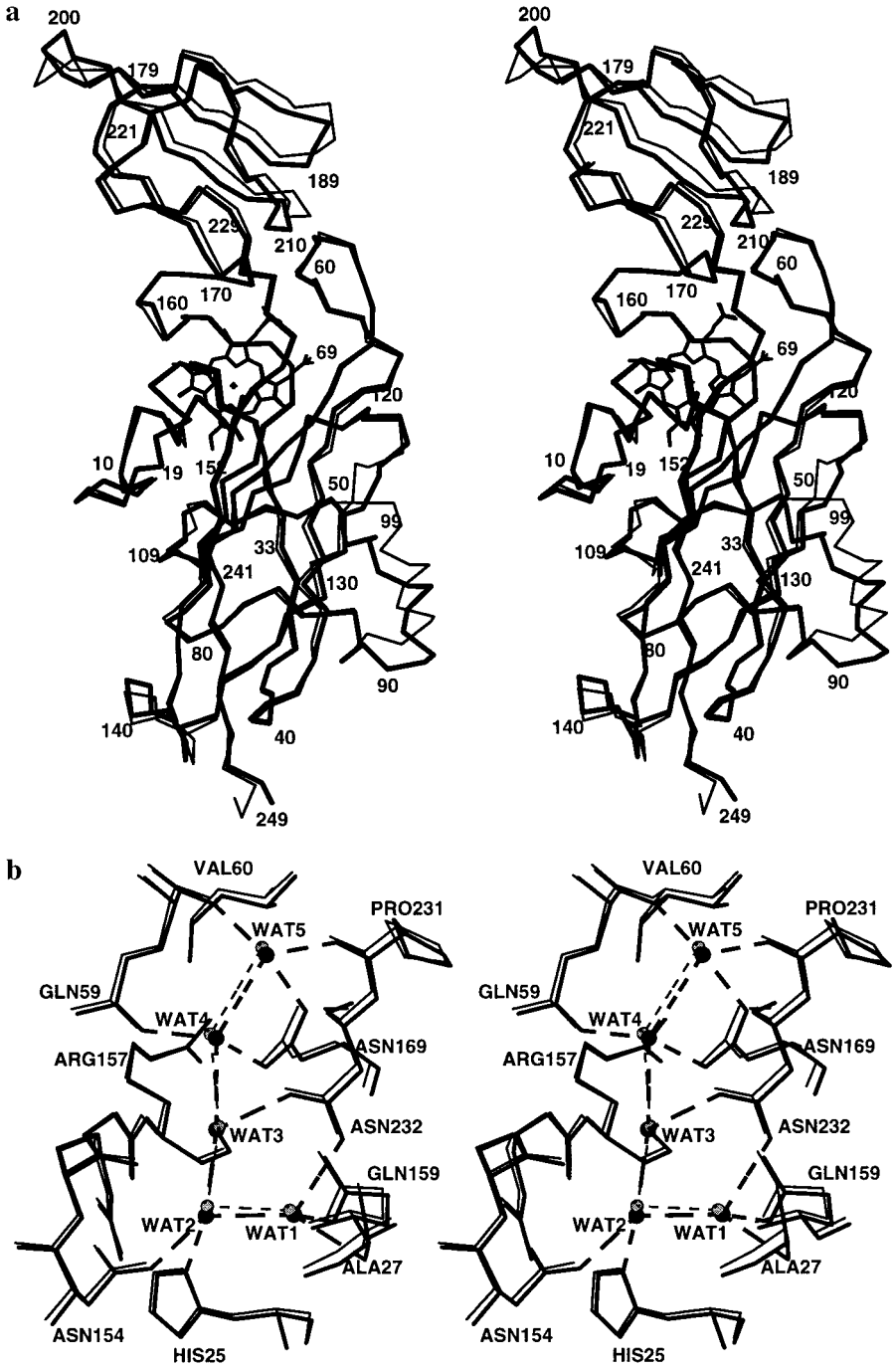

FIGURE 3: Comparison of cyanobacterial and plant cytochromes *f*. Both figures are based on superposition of the 188 Cα atoms of the large domain (RMSD = 0.85 Å), with *P. laminosum* cytochrome *f* in thick lines and turnip cytochrome *f* in thin lines. (a) Stereo Cα trace of the polypeptide. Residue numbers pertain to *P. laminosum* cytochrome *f*. (b) Close-up stereoview of the buried water chains and surrounding amino acids in cytochrome *f* of *P. laminosum* (waters as black spheres) and turnip (gray spheres). Hydrogen bonds are shown as dashed lines. Labels pertain to *P. laminosum* cytochrome *f*. The diagram was drawn with MOLSCRIPT (*57*).

strongest water peaks in electron density maps (Figure 1). The buried water chains in the cyanobacterial and plant cytochromes are identical to the limit of detection in the high-resolution crystal structures. Based on the superposition of Cα atoms in the large domain, the five water molecules in the buried chain have an RMSD of 0.28 Å (Figure 3b). This structural deviation is similar to the estimated error of atomic positions in the crystal structure (0.1−0.2 Å). We also compared the water chain equivalence with the agreement of common water sites on the surfaces of the cyanobacterial and plant cytochromes. A total of 92 common water sites have an RMSD of 0.81 Å. Thus, the buried water chains are more similar than either the Cα atoms or the other common water sites.

The buried five-water chain occurs throughout the biological range of cytochrome *f*, from cyanobacteria to higher plants and algae, as it is also found in the algal cytochrome *f* from *C. reinhardtii* (*21*). Conservation of the water chain also includes its protein-binding site. A group of invariant side chains surrounds the water chain, much like the group surrounding the heme. Of the 10 residues that form hydrogen bonds with the water chain, 7 are invariant among the 22

reported cytochrome *f* sequences. Of the three variable residues, two are hydrogen-bonded to the water chain through backbone NH groups and not side chains (Ala27 and Val60). The other variable side chain is Asn153, which is invariant in 21 of 22 cytochrome *f* sequences. In the *Vicia faba* sequence, this residue is reported to be Thr from a codon (ACC) that differs from Asn (AAC) at a single position.

A cluster of 5 hydrogen-bonded water molecules buried inside a 188-residue protein domain is indeed very rare (*37*). The buried water chain is unlikely to be conserved throughout the biological range of cytochrome *f* for reasons of "structural stability". In general, stability for the folded structure of soluble proteins comes from the hydrophobic effect through desolvation of hydrophobic side chains. However, hydrophobic interactions have little specificity, which is provided by hydrogen bonds of backbone atoms in secondary structure and of polar side chains. Buried water is seen only occasionally, apparently due to the high entropic cost of ordering a water molecule.

The conservation and rarity of the buried water chain provide strong circumstantial evidence that it has a role in cytochrome *f* function. Experimental evidence comes from recent mutagenesis studies of cytochrome *f* from the alga *C. reinhardtii*. Mutation of side chains interacting with the water chain results in altered function in vivo and altered spectra for the soluble form expressed in *E. coli* (*38*). We have suggested that the water chain may be a wire for protons exiting the cytochrome b_6f complex (*20*), although an exit port through a protein domain in the aqueous lumen is not required. Alternatively, the water chain may bind a proton for electrostatic stabilization upon reduction of the heme, thus acting as a dielectric well. In either case, electron transfer from the Rieske protein to cytochrome *f* may be thought of as "proton-coupled". Thus, while the exact function of the buried water chain in proton and electron transfer to cytochrome *f* remains uncertain, both mutagenesis and structural data point to an important function.

Berry and co-workers have proposed that two surface water molecules (W6 and W7) are also part of the water chain (*21*). However, W6 and W7 are on the cytochrome surface whereas the water chain (W1−W5) is completely buried. Neither W6 nor W7 is in contact with any water molecule in the buried water chain of either cyanobacterial or plant cytochrome *f*. In cyanobacterial cytochrome *f*, W6 is 3.8 Å from W5, and bridged to it through the backbone carbonyl of Pro231. W7 is 4.7 Å from W1, and connected through the side chain of Asn232. W6 and W7 are typical of surface water molecules hydrogen-bonded to conserved regions of the protein. For these reasons, we do not consider W6 and W7 to be part of the buried water chain.

Electrostatic Surface Potential. The importance of an electrostatic interaction between plant cytochrome *f* and its oxidant, plastocyanin, has been widely discussed. Electron-transfer rates are dependent on ionic strength in vitro (*39*). Chemical cross-linking has been achieved (*40*) between Lys187, part of a basic surface patch on cytochrome *f* (*15*), and Asp44, which is in a highly acidic patch on plastocyanin (*41*). Electrostatic interactions between spinach plastocyanin and the soluble fragment of turnip cytochrome *f* have been confirmed in studies of the complex structure in solution by paramagnetic NMR and restrained rigid-body molecular dynamics (*42*). However, recent mutagenesis studies imply that such interactions are not necessary in vivo. Mutant forms of algal cytochrome *f*, in which the prominent basic surface was eliminated by up to five amino acid substitutions, lost the in vitro ionic strength dependence, as anticipated, but retained wild-type rates of cytochrome *f* oxidation in vivo (*43, 44*). The nature of the selection pressure that leads to conservation of the charge patterns of the plant proteins therefore remains uncertain.

Cyanobacterial and plant cytochrome *f* have strikingly different electrostatic surface potentials (Figure 4A,B). The surface of *P. laminosum* cytochrome *f* is highly acidic with no basic patch, and the same is true for the surfaces predicted from sequences of three other cyanobacterial cytochromes. The plant cytochrome surface is about 2 pH units less acidic, and includes a basic patch surrounding Lys187, the site of chemical cross-linking. The redox partner plastocyanin also has a different electrostatic potential surface in cyanobacteria than in plants (Figure 4C,D). The difference is roughly complementary to the difference in cytochrome *f* surfaces. Plant plastocyanins have a highly acidic surface (*41, 45*−*47*), whereas the surface of cyanobacterial plastocyanin is about 2 pH units less acidic (*48*−*50*). The surfaces of the cyanobacterial and plant proteins are consistent with an electrostatic interaction between cytochrome *f* and plastocyanin, when considered outside the context of the cytochrome b_6f complex in a photosynthetic membrane. Another possibility is that the strongly acidic surfaces of cyanobacterial cytochrome *f* and of plant plastocyanin make significant contributions to buffering the lumen phases of cyanobacteria and chloroplasts, respectively.

The charge on plastocyanin should also be compatible with the charge on its electron acceptor. In chloroplasts, as in cyanobacteria, the major acceptor is the P700 reaction center of photosystem I. In chloroplasts, binding is thought to be strongly influenced by interaction between the acidic patch of plastocyanin and basic residues present in an amphipathic helix near the N-terminus of the small subunit PsaF (*51*). Sequences of cyanobacterial PsaF, on the other hand, suggest that such charge−charge interactions are unimportant in cyanobacteria. We have previously suggested that the more basic character of cyanobacterial plastocyanin, and a correspondingly more acidic cytochrome *f*, may be determined by the need to interact with acidic receptor groups in cytochrome oxidase (*52*), but this suggestion may need to be modified if the immediate donor to cytochrome oxidase is a *c*-type cytochrome known as cytochrome M (*53, 54*). In any event, such a restraint is inadequate to explain the strongly acidic character of cyanobacterial cytochrome *f*.

The electrostatic potential for one surface of the cytochrome *f* large domain is similar in cyanobacteria and plants (Figure 4B). Given the strong conservation of cytochrome b_6 and subunit IV throughout all oxygenic photosynthetic organisms (*55*), this surface may pack against other protein subunits of the cytochrome b_6f complex.

Conservation of Structure and Function in Cytochrome f and Cytochrome b_6f. Three critical regions of the cytochrome *f* structure were highlighted by mapping invariant amino acids onto the three-dimensional structure (Figure 2). Among all of the 22 reported cytochrome *f* sequences, including cyanobacteria, algae, and higher plants, 63 residues in the 249-residue extrinsic portion are strictly invariant. Virtually all of these invariant residues are confined to the heme-binding large domain. The first cluster of invariant residues

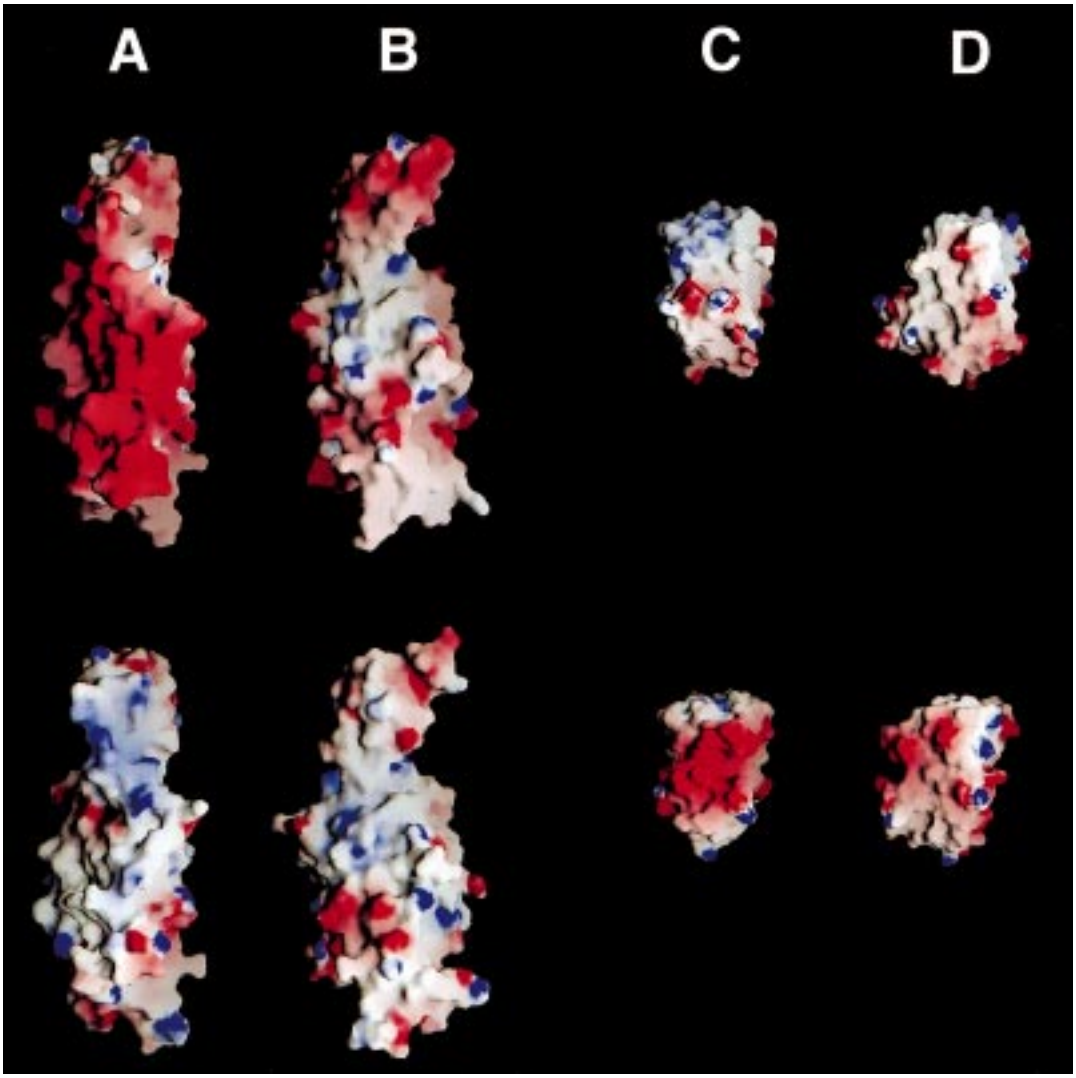

FIGURE 4: Electrostatic potential surfaces of cyanobacterial (top row) and plant (bottom row) cytochrome *f* and plastocyanin. *P. laminosum* cytochrome *f* (A,B) and plastocyanin (C,D; PDB code 1baw; *50*) are shown in the top row. Equivalent views of turnip cytochrome *f* (PDB code 1hcz; *20*) and poplar plastocyanin (PDB code 1pnc; *41*) are shown in the bottom row. (A) Analogous surfaces include a strongly acidic region in *P. laminosum* cytochrome *f* (top) and a basic patch in turnip cytochrome *f* (bottom). (B) The most similar surfaces of *P. laminosum* (top) and turnip (bottom) cytochrome *f* show the greatest similarity in their large domains. (C) The strongly acidic patch on the plant plastocyanin surface (bottom) does not exist on the analogous surface of the cyanobacterial protein (top). (D) Other surfaces of cyanobacterial (top) and plant (bottom) plastocyanins do not have highly acidic or basic patches. The views in (A) are from the right with respect to Figure 2 and differ from the views in (B) by 135° rotation about the vertical axis; views in (C) are perpendicular to those in (D). For all images, the color ramp for positive (blue) or negative (red) surface potential saturates at 10 kT. The potentials were calculated and the surfaces rendered with GRASP (*59*). Formal charges only were assigned.

surrounds the heme and may be responsible for the unusually positive midpoint potential of cytochrome *f* (290−360 mV). The second cluster of invariant residues is near the transmembrane helix and was unexpected. This region of cytochrome *f* may interact with other proteins of the cytochrome b_6f complex in a manner that is conserved throughout cyanobacteria, plants, and algae. The third cluster of invariant residues binds the water chain and provides strong evidence in support of a critical function for the buried waters, as discussed above. In contrast to these invariant residues, the core of the large domain is conserved in hydrophobic character but is not invariant. The small domain has only three invariant glycine residues in reverse turns of the polypeptide, suggesting that all cytochromes *f* have a small domain with a conserved fold, but with variable interior and surface.

Several important conclusions about cytochrome b_6f structure and function can be drawn from available structural data and from the sequence database. The structural data include crystal structures of intact cytochrome bc_1 (*5*−*8*) and of extrinsic domains for cytochrome *f* (*15*) and the Rieske protein (*14*) from the b_6f complex. The sequences of cytochrome b_6f and bc_1 complexes are a patchwork of highly conserved, highly diverged, and unrelated domains. Cytochrome b_6 and subunit IV of the cytochrome b_6f complex are homologous to the N- and C-terminal halves, respectively, of cytochromes *b* in the bc_1 complex (*3*). The Rieske Fe_2S_2 proteins of cytochromes b_6f and bc_1 are homologous (*13, 14*), although sequence invariance is confined to their cluster-binding subdomains. The dissimilar structures of cytochrome *f* and cytochrome c_1 demonstrate unequivocally that the *c*-type cytochromes of the b_6f and bc_1 complexes

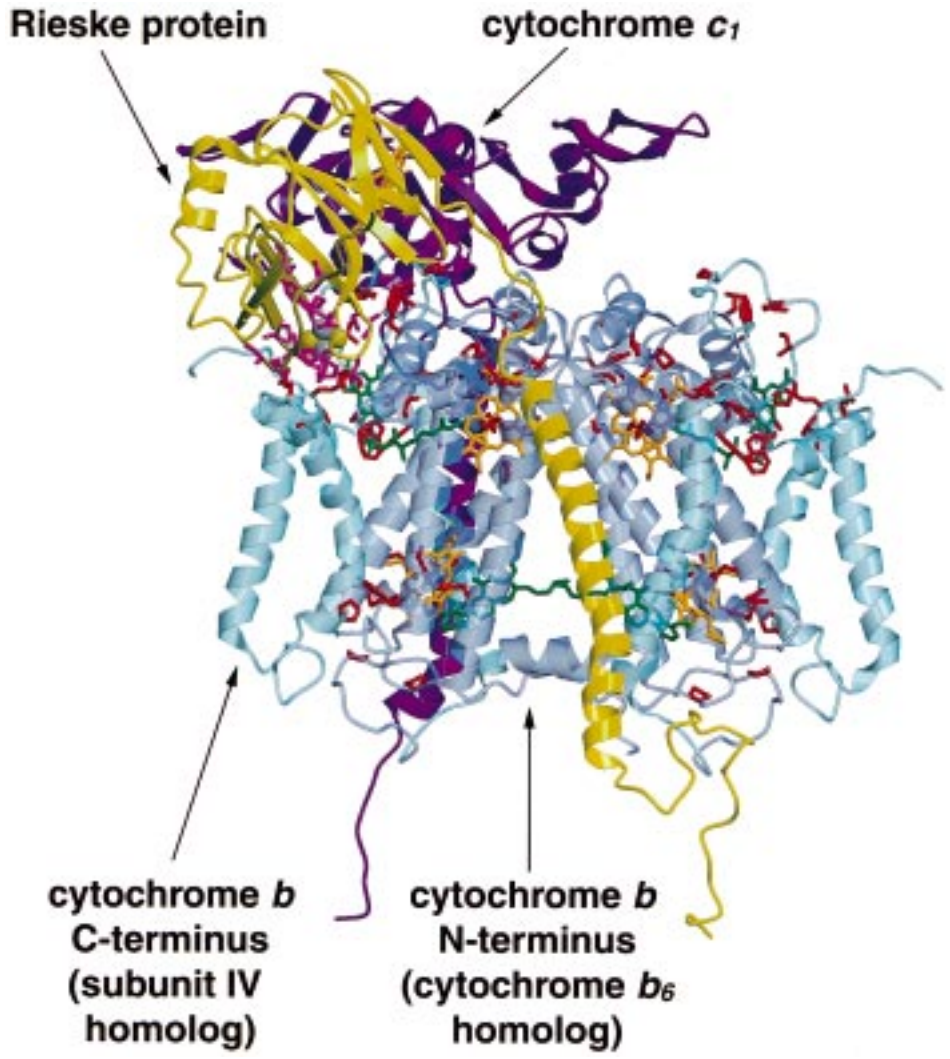

FIGURE 5: Structural conservation in the cytochrome b_6f/bc_1 family. Relevant portions of the chicken cytochrome bc_1 structure (*6*) are shown (PDB code 3bcc for the stigmatellin complex). In this view, the *p*-side of the membrane is at the top of the figure and the *n*-side at the bottom. In the transmembrane core, the N-terminal segment of the cytochrome *b* dimer that is homologous with cytochrome b_6 is shown in medium blue, and the C-terminal segment homologous with subunit IV in light blue. The cytochrome *b* C-terminal helix, which has no homologue in cytochrome b_6f, is omitted. The *p*-side Rieske protein (yellow) and cytochrome c_1 (purple) are shown for only one monomer of the dimeric complex. Heme b_p at the top of cytochrome *b*, heme b_n at the bottom of cytochrome *b*, and heme *c* in cytochrome c_1 are in gold. The inhibitors stigmatellin in the Q_p site and antimycin in the Q_n site are in green. Invariant residues throughout the b_6f/bc_1 family are drawn in red for cytochrome *b* and in magenta for the Rieske protein (*14*). The invariant residues in cytochrome *b* are identified under Materials and Methods.

are unrelated. Three small, hydrophobic, transmembrane subunits are also present in each complex. None of the small subunits nor the transmembrane domains of the Rieske protein and cytochrome *f* in cytochrome b_6f have detectable sequence similarity with any of their counterparts in cytochrome bc_1. Five additional subunits of the cytochrome bc_1 complex (subunits 1, 2, 6, 8, and 9) are not present in cytochrome b_6f. We infer a common structure and function for those parts of the complexes that are conserved across the entire cytochrome b_6f/bc_1 family.

Regions of common structure and function emerge from this patchwork of conservation and variability. A common architecture for the b_6f/bc_1 transmembrane core is apparent in the sequence identity of cytochrome *b* of the bc_1 complex with cytochrome b_6 and subunit IV of the b_6f complex. Important binding sites are remarkably conserved between cytochromes b_6f and bc_1 on the *p*-side, but not the *n*-side, of the transmembrane core (Figure 5). This includes the heme b_p, Q_p, and Rieske Fe_2S_2 binding sites in the *b* cytochrome, as well as the cluster-binding domain of the Rieske protein itself. In contrast, no residues in the bc_1 Q_n site are conserved between cytochromes b_6f and bc_1. The cluster of conserved residues around the proximal *p*-side binding sites in the *b*-cytochrome is consistent with the current understanding of *p*-side function. The Fe_2S_2-binding region of the Rieske protein docks into cytochrome *b* at the Q_p site during part of the electron-transfer turnover cycle (*6*, *7*), with b_6f/bc_1-conserved residues in contact (Figure 5). The structures and the trend of *p*-side conservation, which was detected in earlier sequence comparisons (*55*), imply common mechanisms for *p*-side electron-transfer events near the Q_p site, and differences in mechanism elsewhere. Details of electron transfer at the *n*-side may be quite different in the b_6f and bc_1 complexes, and events in the *p*-side aqueous phase may also differ, based on the different structures of the *c*-type cytochromes *f* and c_1 and of their oxidant proteins.

Much larger regions of conservation result when only cytochrome b_6f sequences are compared. Conservation of cytochrome *b* and subunit IV in all cytochrome b_6f complexes includes a large cluster around the heme b_n and the putative Q_n site. Clearly these are sites of common mechanism among cytochrome b_6f complexes, but are rather different than the situation in the cytochrome bc_1 complex. The area of b_6f-only conservation on the *p*-side surface of cytochrome b_6 and subunit IV is also larger than the area conserved throughout the cytochrome b_6f/bc_1 family. This surface may interact with the conserved region of cytochrome *f* near the membrane-anchor helix.

Evolution of Cytochrome bc Complexes. The results described here show that the structure of the major soluble portion of cytochrome *f* has been strongly conserved since chloroplasts originated from the endosymbiotic uptake of oxygenic photosynthetic bacteria by a primitive eukaryote over a billion years ago. Together with the recently obtained structure of cytochrome c_1 (*6*−*8*), these data show that cytochromes *f* and c_1 are structurally distinct, unrelated proteins. It is remarkable that extant *bc* complexes employ completely different membrane-bound *c* cytochromes to transfer electrons to the shuttle redox partner, plastocyanin or cytochrome *c*, although both photosynthetic and respiratory *bc* complexes perform essentially identical overall functions during energy transduction. The ability of cytochromes *f* and c_1 to attain a similar functional status implies a process of convergent evolution from independent origins. Sequence data for *bc* complexes in a few other, bacterial branches of the phylogenetic tree provide some information on the evolutionary changes relevant to the appearance of cytochromes *f* and c_1, although it is difficult to construct a pathway of c_1/f acquisition from so few examples.

The Rieske Fe-S protein may be a predictive connection between variable and conserved components of *bc* complexes. The Rieske extrinsic portion includes a highly conserved, small, cluster-binding domain and a variable large domain. However, large domains are closely related for Rieske proteins with a common *c* cytochrome redox partner (*14*). Sequences of Rieske large domains from *Bacillus*, *Chlorobium*, and *Heliobacillus* are unlike Rieske sequences from either cytochrome b_6f or cytochrome bc_1, just as their apparent *c* cytochromes are unlike.

Cytochrome c_1 shares a fold, and presumably a common ancestor, with several soluble cytochromes. No source has been identified for cytochrome *f*. Sequence searches failed to reveal any cytochrome *f* homologues in cyanobacteria or plants other than cytochrome *f* itself. A search of the structure database produced only one plant or cyanobacterial protein with the same β-sheet topology as the large domain of cytochrome *f*. This is plastocyanin, the cytochrome *f* oxidant protein. In the absence of more obvious similarities in sequence or structure, it cannot be proposed that plastocyanin and cytochrome *f* have a common ancestor. It is nevertheless interesting that these two proteins, which have very different redox centers but are functionally interactive, have the same three-dimensional topology.

ACKNOWLEDGMENT

We thank J. M. Guss and H. C. Freeman for sharing their atomic model of *P. laminosum* plastocyanin prior to publication, and J. Hollister and C. Greski for skilled preparation of the manuscript.

REFERENCES

1. Cramer, W. A., Soriano, G. M., Ponomarev, M., Huang, D., Zhang, H., Martinez, S. E., and Smith, J. L. (1996) *Annu. Rev. Plant Physiol. 47*, 477−508.
2. Cramer, W. A., and Knaff, D. B. (1991) *Energy Transduction in Biological Membranes*, Chapters 4, 7, 579 pp, Springer Study Edition, Springer-Verlag, New York.
3. Widger, W. R., Cramer, W. A., Herrmann, R. G., and Trebst, A. (1984) *Proc. Natl. Acad. Sci. U.S.A. 81*, 674−678.
4. Saraste, M. (1984) *FEBS Lett. 166*, 367−372.
5. Xia, D., Yu, C.-A., Kim, H., Xia, J.-Z., Kachurin, A. M., Zhang, J., Yu, L., and Deisenhofer, J. (1997) *Science 277*, 60−66.
6. Zhang, Z., Huang, L., Shulmeister, V. M., Chi, Y.-I., Kim, K. Y., Hung, L.-W., Crofts, A. R., Berry, E. A., and Kim, S.-H. (1998) *Nature 392*, 677−684.
7. Kim, H., Xia, D., Yu, C.-A., Xia, J.-Z., Kachurin, A. M., Zhang, L., Yu, L., and Deisenhofer, J. (1998) *Proc. Natl. Acad. Sci. U.S.A. 95*, 8026−8033.
8. Iwata, S., Lee, J. W., Okada, K., Lee, J. K., Iwata, M., Rasmussen, B., Link, T. A., Ramaswamy, S., and Jap, B. K. (1998) *Science 281*, 64−71.
9. Willey, D. L., Auffret, A. D., and Gray, J. C. (1984) *Cell 36*, 555−562.
10. Furbacher, P. N., Tae, G.-S., and Cramer, W. A. (1996) in *Origin and Evolution of Biological Energy Conversion* (Baltscheffsky, H., Ed.) Chapter 9, pp 221−253, VCH Publishers, New York.
11. Degli Esposti, M., de Vries, S., Crimi, M., Ghelli, A., Paternello, T., and Meyer, A. (1993) *Biochim. Biophys. Acta 1143*, 243−271.
12. Nitschke, W., Mühlenhoff, U., and Liebl, U. (1998) in *Photosynthesis: a Comprehensive Treatise* (Raghavendra, A. S., Ed.) pp 285−304, Cambridge University Press, Cambridge, U.K.
13. Iwata, S., Saynovits, M., Link, T. A., and Michel, H. (1996) *Structure 4*, 567−579.
14. Carrell, C. J., Zhang, H., Cramer, W. A., and Smith, J. L. (1997) *Structure 5*, 1613−1625.
15. Martinez, S. E., Huang, D., Szczepaniak, A., Cramer, W. A., and Smith, J. L. (1994) *Structure 2*, 95−105.
16. Anemüller, S., Lübben, M., and Schäfer, G. (1985) *FEBS Lett. 193*, 83−87.
17. Zhang, H., Carrell, C. J., Huang, D., Sled, V., Ohnishi, T., Smith, J. L., and Cramer, W. A. (1996) *J. Biol. Chem. 271*, 31360−31366.
18. Schmidt, C. L., Anemüller, S., Texeira, M., and Schäfer, G. (1995) *FEBS Lett. 359*, 239−243.
19. Castresana, J., Lübben, M., and Saraste, M. (1995) *J. Mol. Biol. 250*, 202−210.
20. Martinez, S., Huang, D., Ponomarev, M., Cramer, W. A., and Smith, J. L. (1996) *Protein Sci. 5*, 1081−1092.
21. Chi, Y.-I., Huang, L.-S., Zhang, Z., Fernández-Velasco, J. G., and Berry, E. A. (1998) in *ESF Workshop on Coupling of Electrons and Protons in Photosynthesis*, Seregélyes, Hungary, August 14−17, 1998, p 45.
22. Kerfeld, C. A., Krogmann, D. W., and Yeates, T. O. (1997) in *Abstracts of the American Crystallographic Association*, Series 2, Vol. 24, p 129, Abstract P134.
23. Wagner, M. J., Packer, J. C., Howe, C. J., and Bendall, D. S. (1996) *Biochim. Biophys. Acta 1276*, 246−252.
24. Schlarb, B. G., Wagner, M. J., Vijgenboom, E., Ubbink, M., Bendall, D. S., and Howe, C. J. (1999) *Gene 234*, 275−283.
25. Otwinowski, Z., and Minor, W. (1997) *Methods Enzymol. 276*, 307−326.
26. Navaza, J. (1994) *Acta Crystallogr. A50*, 157−163.
27. Brünger, A. T. (1992) *X-PLOR Version 3.1, a System for X-ray Crystallography and NMR*, Yale University Press, New Haven, CT.

28. Jones, T. A., Zou, J.-Y., Cowan, S. W., and Kjeldgaard, M. (1991) *Acta Crystallogr. A47*, 110−119.
29. Collaborative Computational Project Number 4 (1994) *Acta Crystallogr. D50*, 760−763.
30. 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 Warrem, G. L. (1998) *Acta Crystallogr. D54*, 905−921.
31. Holm, L., and Sander, C. (1993) *J. Mol. Biol. 233*, 123−138.
32. Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. (1997) *Nucleic Acids Res. 25*, 3389−3402.
33. Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994). *Nucleic Acids Res. 22*, 4673−4680.
34. Read, R. J. (1986) *Acta Crystallogr. A42*, 140−149.
35. Luzzati, P. V. (1952) *Acta Crystallogr. 5*, 802−810.
36. Chothia, C., and Lesk, A. M. (1986) *EMBO J. 5*, 823−826.
37. Williams, M. A., Goodfellow, J. M., and Thornton, J. M. (1994) *Protein Sci. 3*, 1224−1235.
38. Ponamarev, M. V., and Cramer, W. A. (1998) *Biochemistry 37*, 17199−17208.
39. Niwa, S., Ishikawa, H., Nikai, S., and Takabe, T. (1980) *J. Biochem. 88*, 1177−1183.
40. Morand, L. Z., Frame, M. K., Colvert, K. K., Johnson, D. A., Krogmann, D. W., and Davis, D. J. (1989) *Biochemistry 28*, 8039−8047.
41. Guss, J. M., and Freeman, H. C. (1983) *J. Mol. Biol. 169*, 521−563.
42. Ubbink, M., Ejdeback, M., Karlsson, B. G., and Bendall, D. S. (1998) *Structure 6*, 323−335.
43. Soriano, G. M., Ponamarev, M. V., Tae, G.-S., and Cramer, W. A. (1996) *Biochemistry 35*, 14590−14598.
44. Soriano, G. M., Ponamarev, M. V., Piskorowski, R., and Cramer, W. A. (1998) *Biochemistry 37*, 15120−15128.
45. Moore, J. M., Lepre, C. A., Gippert, G. P., Chazin, W. J. Case, D. A., and Wright, P. E. (1991) *J. Mol. Biol. 221*, 533−555.
46. Bagby, S., Driscoll, P. C., Harvey, T. S., and Hill, H. A. O. (1994) *Biochemistry 33*, 6611−6622.
47. Xue, Y., Okvist, M., Hansson, O., and Young, S. (1998) *Protein Sci. 7*, 2099−2105.
48. Badsberg, U., Jorgensen, A. M., Gesmar, H., Led, J. J., Hammerstad, J. M., Jespersen, L. L., and Ulstrup, J. (1996) *Biochemistry 35*, 7021−7031.
49. Romero, A., de la Cerda, B., Varela, P. F., Navarro, J. A., Hervás, M., and de la Rosa, M. A. (1998) *J. Mol. Biol. 275*, 327−336.
50. Bond, C. S., Bendall, D. S., Freeman, H. C., Guss, J. M., Howe, C. J., Wagner, M. J., and Wilce, M. C. J. (1999) *Acta Crystallogr. D55*, 414−421.
51. Hippler, M., Reichert, J., Sutter, M., Zak, E., Altschmied, L., Schröer, U., Herrmann, R. G., and Haehnel, W. (1996) *EMBO J. 15*, 6374−6384.
52. Bendall, D. S., Wagner, M. J., Schlarb, B. G., Söllick, T.-R., Ubbink, M., and Howe, C. J. (1999) in *The Phototrophic Prokaryotes* (Peschek, G. A., Löffelhardt, W., and Schmetterer, G., Eds.) pp 315−328, Kluwer Academic/Plenum Publishers, New York.
53. Pils, D., Gregor, W., and Schmetterer, G. (1997) *FEMS Microbiol. Lett. 152*, 83−88.
54. Manna, P., and Vermaas, W. (1997) *Plant Mol. Biol. 35*, 407−416.
55. Cramer, W. A., Martinez, S. E., Huang, D., Tae, G.-S., Everly, R. M., Heymann, J. B., Cheng, R. H., Baker, T. S., and Smith, J. L. (1994) *J. Bioenerg. Biomembr. 26*, 31−47.
56. Evans, S. V. (1993) *J. Mol. Graph. 11*, 134−138.
57. Kraulis, P. J. (1991) *J. Appl. Crystallogr. 24*, 946−950.
58. Merritt, E. A., and Bacon, D. J. (1997) *Methods Enzymol. 277*, 505−524.
59. Nicholls, A., Sharp, K., and Honig, B. (1991) *Proteins: Struct., Funct., Genet. 11*, 281−296.

BI9903190