X-ray Structure of a Truncated Form of Cytochrome *f* from *Chlamydomonas reinhardtii*[†,‡]

Young-In Chi,[§,||] Li-Shar Huang,[⊥] Zhaolei Zhang,[⊥,@] Javier G. Fernández-Velasco,*[,#] and Edward A. Berry*[,⊥]

E. O. Lawrence Berkeley National Laboratory, Berkeley, California 94720, and Graduate Group of Biophysics, Department of Chemistry, and Department of Plant and Microbial Biology, University of California, Berkeley, California 94720

Received January 14, 2000; Revised Manuscript Received April 3, 2000

ABSTRACT: A truncated form of cytochrome *f* from *Chlamydomonas reinhardtii* (an important eukaryotic model organism for photosynthetic electron transfer studies) has been crystallized (space group $P2_12_12_1$; three molecules/asymmetric unit) and its structure determined to 2.0 Å resolution by molecular replacement using the coordinates of a truncated turnip cytochrome *f* as a model. The structure displays the same folding and detailed features as turnip cytochrome *f*, including (a) an unusual heme Fe ligation by the α-amino group of tyrosine 1, (b) a cluster of lysine residues (proposed docking site of plastocyanin), and (c) the presence of a chain of seven water molecules bound to conserved residues and extending between the heme pocket and K58 and K66 at the lysine cluster. For this array of waters, we propose a structural role. Two cytochrome *f* molecules are related by a noncrystallographic symmetry operator which is a distorted proper 2-fold rotation. This may represent the dimeric relation of the monomers in situ; however, the heme orientation suggested by this model is not consistent with previous EPR measurements on oriented membranes.

Biological systems couple redox energy to processes, such as biosynthesis and transport, through electron transport chains and proton electrochemical gradients across membranes (*1*−*4*). The electron transport chains include a series of membrane-bound redox enzymes which couple their redox reactions to translocation of protons across the membrane, generating a high-energy proton electrochemical potential gradient.

The plastoquinol:plastocyanin oxidoreductase (cytochrome b_6f complex, EC 1.10.99.1) (*5*, *6*) plays such a role in the eukaryotic and prokaryotic oxygenic photosynthetic electron transfer chain. Its role is analogous to that of the ubiquinol: cytochrome *c* oxidoreductase (cytochrome bc_1 complex) in mitochondrial (*1*, *2*) or prokaryotic (*4*) respiration, and prokaryotic anoxygenic photosynthesis (*4*), and indeed, the sequences of the two proteins are significantly homologous.

To understand the detailed mechanism of electron transfer and proton translocation, high-resolution structural information is useful. Atomic structural models of the mitochondrial cytochrome bc_1 complex (*7*, *8*) have shed light on its mechanism; however, the structure of the cytochrome b_6f complex is not known. At present, only the general dimeric nature of the cytochrome b_6f complex (*9*−*12*) and a projection map at 8−9 Å resolution (*13*, *14*) are available. Crystals of a cyanobacterial b_6f complex have been reported (*15*), but considerable improvement in the order will be required before an atomic-resolution structure can be obtained.

A useful strategy for gaining structural information about multisubunit proteins is to divide the large complexes into smaller components that are more amenable to structure determination. Thus far, elucidation of the structure of individual subunits has been successful for a truncated form of cytochrome *f* (a component of the chloroplast cytochrome b_6f complex) from turnip (*16*, *17*), as well as for truncated forms of the Rieske iron−sulfur subunit of the cytochrome bc_1 (*18*) and b_6f (*19*, *20*) complexes.

Cytochrome *f* is one of the four redox centers in the cytochrome b_6f complex of the thylakoid membrane in oxygenic photosynthetic organisms and is analogous in function to the mitochondrial/bacterial cytochrome c_1 of the cytochrome bc_1 complex. Cytochrome *f* is the electron donor to the Cu-containing protein plastocyanin (*21*) [or to cytochrome c_6 (*c*552) in some eukaryotic algae and cyanobacteria, when growing in a Cu-deficient environment (*22*−*24*)] and

[†] This work was supported in part by a grant from the National Institutes of Health (GM-20571) to Prof. R. Malkin (Department of Plant and Microbial Biology, University of California, Berkeley) and by the Office of Health and Environmental Research, U.S. Department of Energy, under Contract DE-AC03-76SF00098 to the E. O. Lawrence Berkeley National Laboratory. Financial support to Y.-I.C. is also acknowledged.

[‡] The coordinates of the structure presented here have been deposited in the Protein Data Bank under accession number 1CFM.

* To whom correspondence should be addressed. J.G.F.-V.: telephone, (510) 642-9425; fax, (510) 642-4995; e-mail, jfernand@nature.berkeley.edu. E.A.B.: telephone, (510) 486-4336; fax, (510) 486-5272; e-mail, eaberry@LBL.gov.

[§] Department of Chemistry, University of California.

[||] Present address: Joslin Diabetes Center and Department of Medicine, Harvard Medical School, Boston, MA 02215.

[⊥] E. O. Lawrence Berkeley National Laboratory.

[@] Graduate Group of Biophysics, University of California.

[#] Department of Plant and Microbial Biology, University of California.

10.1021/bi000090k CCC: $19.00
Published on Web 06/06/2000

is thought to be reduced directly by the Rieske iron−sulfur protein (*3*). The protein has a predicted molecular mass of ca. 32 kDa, contains one covalently bound heme C group per molecule, and has a midpoint redox potential of ca. 350 mV (*5*, *25*). Hydropathy profiles have predicted that the protein has one transmembrane-spanning region, near its C-terminus, and that most of the protein is localized in the thylakoid lumen (*25*). When a lumen-side water-soluble fragment of cytochrome *f* was isolated from turnip, it lacked the C-terminus transmembrane-spanning region (33 residues) (*16*, *26*), but was still redox-active and possessed the same midpoint redox potential as the intact cytochrome *f* (*17*). The structure of this truncated form has been reported by Martinez et al. (*16*, *17*). The protein consists predominatly of β-sheet, has a unique ligation of the heme by the α-amino N of the N-terminal residue, and is divided into a large domain and a small domain. There is a cluster of basic residues ("lysine patch") at the junction of the two domains, and there is a chain of five buried water molecules.

In recent years, the eukaryotic unicellular green alga *Chlamydomonas reinhardtii* has become the preferred model organism for genetic studies of eukaryotic photosynthesis. This alga has a photosynthetic electron transfer chain similar to the one found in higher plants but with an advantage over them in that the chloroplast genome of *C. reinhardtii* is easily transformable (*27*, *28*). Thus, it allows the preparation of site-directed mutants of chloroplast-encoded proteins. Among these proteins, cytochrome *f* (chloroplast petA gene) is of particular interest because in *C. reinhardtii* it is 72% identical to the higher-plant cytochrome (*25*) and because the electron transfer reactions of cytochrome *f* can be studied by kinetic spectrophotometry in situ in whole algal cells (*29*−*32*). Thus, the algal cytochrome *f* is a very good model of the higher-plant cytochrome *f* but has the advantage of being mutationally manipulatable. For these reasons, and to complement functional descriptions with specific structural data, it is of great interest to know the actual structure of the *C. reinhardtii* cytochrome *f*.

For this work, the nonphotosynthetic strain F283ST of *C. reinhardtii* constructed by Kuras et al. (*30*) was used to purify a truncated soluble form of cytochrome *f* that is naturally overexpressed. In the petA gene of this strain, a stop codon replaces the codon for I252 (I283 numbered from the M initiation codon in the preprocessed protein). Thus, the truncated cytochrome lacks the last 35 residues that would involve the membrane-anchoring α-helix. Consequently, this modified protein is normally processed and translocated into the thylakoid lumen where it remains soluble, being redox-active in vivo and having spectral characteristics very similar to those of the wild-type form (*30*). The transformation of this intrinsic membrane protein into a soluble one, which represents the peripheral portion of the native cytochrome, facilitates its purification and, in combination with the conserved physicochemical properties, provides an attractive system for structural studies.

In this paper, we present the crystal structure at 2.0 Å resolution of a truncated form of cytochrome *f* from *C. reinhardtii* and compare this structure with that previously described from turnip. We find that the structures are similar but also contain some significant differences. An abstract of this work has been previously communicated (*33*).

MATERIALS AND METHODS

Protein Preparation and Crystallization. C. reinhardtii strain F283ST was kindly provided by Francis-Andrè Wollman (Institut de Biologie Physico-Chimique, Paris, France). The microalga was grown in carboys containing 12 L of TAP medium (*34*) at 28 °C with strong aeration (2 L min^{-1} per carboy) and moderate magnetic stirring. Dim incandescent light (10 μeinstein m^{-2} s^{-1}) was provided. Typically, the culture grew with a duplication time of 16−18 h and was harvested with a continuous flow centrifuge at 29000*g* when a density equivalent to $A_{750} = 0.6-0.8$ was reached. The cell paste was stored at −20 or −80 °C.

For cytochrome purification (all steps in ice or at 4 °C), the cells (80 g wet weight) were thawed and suspended in 800 mL of 50 mM MOPS/KOH (pH 7.2), 100 mM KCl, 3 mM NaN_3, 0.5 mM Na-EDTA, and 0.1 mM phenylmethanesulfonyl fluoride, sonicated [5 min, Heat Systems Ultra Sonics Inc., model W185, setting 9 (ca. 110 W), 1 cm diameter probe], and then centrifuged for 2 h at 11300*g*. The supernatant was diluted to 6 L with 0.5 mM Na-EDTA and loaded onto a 300 mL column of DEAE-Sepharose CL6B (Pharmacia) equilibrated with 20 mM KP_i (pH 7.5) and 0.5 mM EDTA. The column was washed with 200 mL of the equilibration buffer and then eluted with an 800 mL linear gradient from 20 to 75 mM KP_i (pH 7.5) and 0.5 mM EDTA. The pink cytochrome *f*-containing fractions were pooled and diluted with an equal volume of distilled water, and ion exchange chromatography was repeated in the absence of EDTA. Pooled fractions from the second DEAE column were applied to a 75 mL column of ceramic hydroxyapatite type II, 0.08 mm (Bio-Rad). The column was washed with 200 mL of 50 mM KP_i (pH 7.5) and then eluted with an 800 mL linear gradient from 50 to 120 mM KP_i. The pooled fractions were concentrated to ≤5 mL by ultrafiltration (Amicon YM-10 membrane) and loaded onto a 600 mL column of Sephacryl S-200 HR (Pharmacia) equilibrated with 50 mM KP_i (pH 7.5) and 0.5 mM EDTA, to finally elute cytochrome *f* with the same buffer. The protein is obtained in its reduced state.

The crystals were initially screened by the sparse-matrix method (*35*), and one promising set of conditions was optimized. For crystallization, the pooled fractions from the Sephacryl column were concentrated by ultrafiltration (Centricon P-10, Amicon) to about 2 mM on the basis of an extinction coefficient of 34 mM^{-1} cm^{-1} at 554 nm. Portions of 5−10 μL were mixed with an equal volume of precipitant containing 180 g/L PEG-8000, 0.1 M sodium cacodylate (pH 6.5), and 0.2 M $Zn(OAc)_2$ and allowed to equilibrate at 4 °C by vapor diffusion against the same precipitant until crystal growth stopped.

Data Collection. The crystal, cryoprotected in the mother liquor with 30% (v/v) glycerol, was flash-frozen in the nitrogen cold stream for data collection at 108 K on a R-AXISIIC imaging plate detector coupled with a Rigaku Rotaflex X-ray generator. The data set was processed with DENZO (*36*) and scaled with the CCP4 program suite (*37*), resulting in the data statistics shown in Table 1.

Structure Determination. The structure was determined by the molecular replacement method using the AMORE program (*38*) with a preliminary 2.3 Å dataset and the coordinates of the truncated cytochrome *f* from turnip

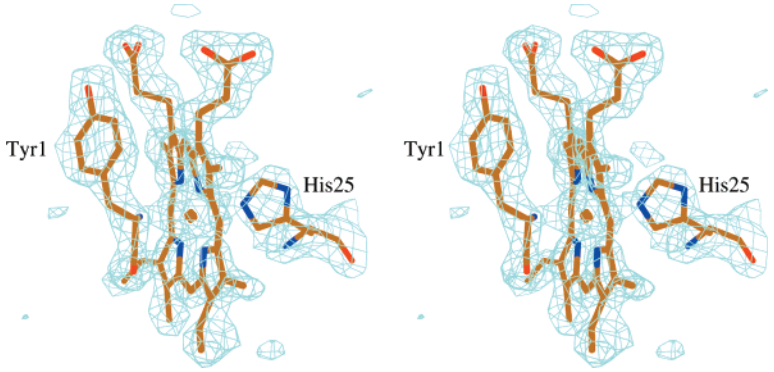

FIGURE 1: Omit map of the heme region. The electron density map is made with coefficients of $F_o - F_c$ and Φ_c, where F_c and Φ_c are amplitudes and phases, respectively, calculated from the final refined model after deleting the heme and residues Y1 and H25 to avoid model bias in the heme region. The map is contoured at the 3σ level.

Table 1: Data Processing and Refinement Statistics[a]

data statistics	
resolution (Å)	2.0
no. of reflections	47961
redundancy	3.3 (2.8)
data coverage (%)	91.6 (68.7)
I/σ	14.8 (3.5)
R_{sym} (%)	4.2 (17.6)
refinement statistics	
resolution range (Å)	25.0−2.0
no. of non-H protein atoms	5877
no. of solvent molecules	1001
R-factor for 44561 reflections ($F > 2\sigma$, 95%)	21.4% (24.1%)
R_{free} for 2346 reflections ($F > 2\sigma$, 5%)	27.1% (31.7%)
rms deviation from ideal geometry	
bond lengths (Å)	0.006
bond angles (deg)	1.5
dihedral angles (deg)	24.3
average temperature factors (Å^2)	
main chain atoms only	23.8
all protein atoms (excluding water)	24.8
water molecules only	36.6

[a] Values in parentheses correspond to the highest-resolution shell (2.13−2.0 Å).

(Protein Data Bank entry 1CTM) as the search model. After rigid-body refinement in the program X-PLOR (*39*), the *R*-factor for data between 10.0 and 2.3 Å was reduced to 39.4%. An F_o[1] $- F_c$ map calculated with phases from this rigid-body-refined model (still lacking the heme) clearly showed electron density for the heme group (Figure 1). The heme was then added; the sequence was changed to that of *C. reinhardtii*, and those residues that differ between the two species or that did not fit the electron density were rebuilt using the graphics program O (*40*). This model was refined against the final 2.0 Å data set using X-PLOR 3.1. Solvent molecules (all regarded as water) were added conservatively with due regard for their environment, including potential interactions with hydrogen-bond partners. Inclusion of individual atomic temperature factors and removal of the ncs restriction during the final stages were validated by a substantial decrease in the value of R_{free}. At the end of the X-PLOR refinement, the crystallographic *R*-factor was 21.5% and R_{free} was 29.3%. Bulk solvent correction and anisotropic *B*-factor refinement were then carried out in CNS 0.2[2] using all data with $|F| > 2\sigma_F$ between 2.0 and 25 Å, which further reduced the *R*-factor and R_{free} to 21.4 and 27.1%, respectively. Root-mean-square deviations from ideal values were 0.006 Å for bond lengths and 1.5° for angles. The standard deviation in coordinate values estimated from cross-validated Sigma-A treatment was 0.23 Å. Representations of the structure were prepared using the progams O (*40*), Molscript (*41*), Raster 3D (*42*), GRASP (*43*), and Rasmol (*44*).

Other Methods. Absorption spectra were taken with a computerized Aminco DW-2 spectrophotometer in dual beam mode with a spectral band-pass of 1 nm, a scanning rate of 0.8 nm/s, and a response time of 50 ms. The wavelength scale was calibrated using the 656.1 nm peak of the deuterium lamp. SDS−PAGE (*45*) was performed after denaturing samples in the presence of 5% (w/v) SDS with or without 50 mM dithiothreitol at 80 °C for 5 min; samples were not centrifuged before electrophoresis was carried out. Nondenaturing electrophoresis was performed with Tris-glycine gels (pH 8.8). In all cases, precast gels with 14% or 4 to 20% acrylamide (with equivalent results), all buffers, and MW standards (MWs of 200−2.5 kDa) as well as running conditions were from NOVEX (San Diego, CA). Protein determination, using BSA as a standard, and heme-chrome extraction and measurements were performed as described in ref *46*. Analytical size exclusion chromatography through a Pharmacia LKB Superdex 75 HR 10/30 column was performed with a Pharmacia FPLC system at 22 °C and 1 mL/min. The column was calibrated with 12 standards (MWs of 66−3.5 kDa).

RESULTS

Description of the Purified Protein. The truncated form of *C. reinhardtii* cytochrome *f*, which is purified here for the first time, has spectral characteristics almost identical to those of cytochrome *f* from a wild-type strain (*30*) (see results below). It is also redox-active; it can be rapidly photooxidized in intact cells (*30*), and in vitro it can be reduced and oxidized by ascorbate and ferricyanide, respectively. In addition, it can bind to and react with homologous plastocyanin (*47*, *48*).

After the last step of the purification of cytochrome *f* (554.5 nm/280 nm absorbance ratio of 0.72, for the fully

[1] Abbreviations: F_o, observed structure factor amplitudes; F_c, calculated structure factor amplitudes; rms, root-mean-square; ncs, noncrystallographic symmetry; W, water molecule; Φ_c, phases of calculated structure factors; σ, rms deviation from the mean value of an electron density map.

[2] A preliminary release of the program CNS was made available to the S.-H. Kim lab for testing purposes by A. Brünger.

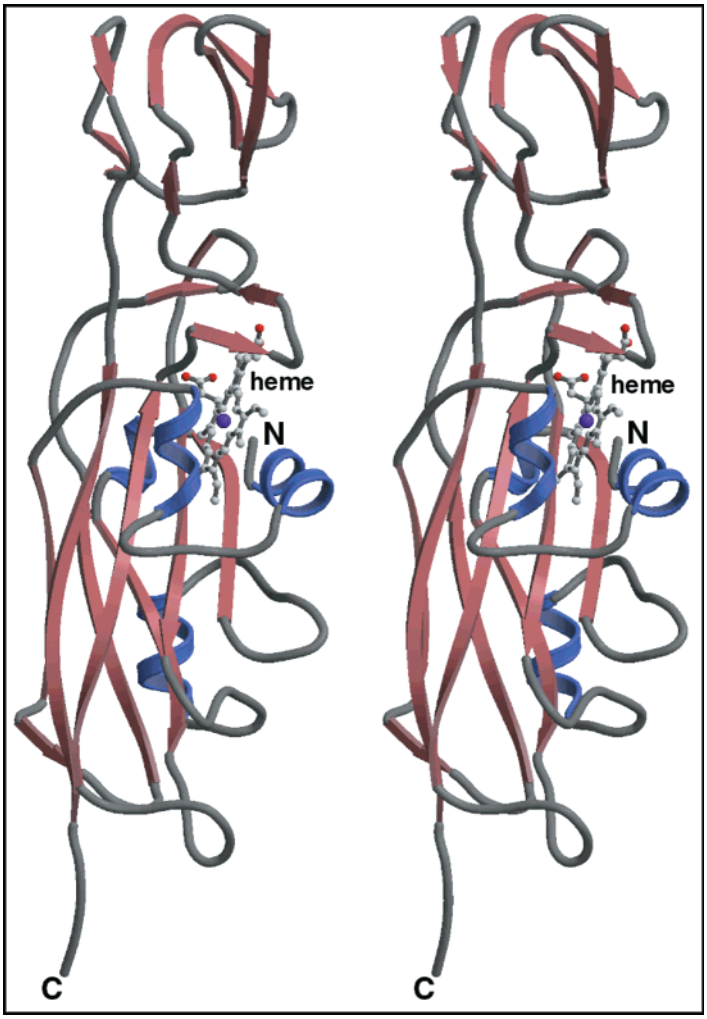

FIGURE 2: Stereo ribbon diagram of cytochrome *f* from *C. reinhardtii* [drawn with Raster 3D (*42*)]. In the wild type, the C-terminus would be continuous forming a membrane-spanning α-helix and a short cytoplasmic segment.

reduced form), an overloaded SDS−PAGE gel (30 μg of protein/lane) stained with Coomasie blue shows only one band and a mobility equivalent to a MW of 31 ± 1 kDa, with the sample prepared under reducing conditions (dithiothreitol) or under nonreducing conditions. The same mobility is also observed with a load of 1 μg of protein. FPLC size exclusion chromatography of the cytochrome under nondenaturing conditions [4 mM NaP_i (pH 6.0) and 250 mM KCl] yields a slightly asymmetrical tailing peak and a molecular mass of 32 ± 2 kDa. Nondenaturing electrophoresis shows (Coomasie) a main band, with little tailing, and a minor component (10%) with a slightly higher mobility.

The spectrum of the purified cytochrome in the reduced state [in 50 mM KP_i (pH 7.4)] exhibits peaks at 421.4, 523.9, 531.1, and 554.2 nm, whereas in the oxidized state, the maxima are at 410.8 and ∼530 nm. The maxima of the reduced-minus-oxidized spectrum are at 421.8, 524.1, 531.1, and 554.2 nm [this last value is 0.5 nm higher than the one found through the light-induced differential spectrum in situ (*30*)]. The isosbestic points are at 414.9, 432.3, ∼509, ∼533, ∼543, and ∼560 nm.

Using the pyridine hemechrome quantification, the extinction coefficients were 34 mM^{-1} cm^{-1} at 554 nm for the reduced protein, whereas in the reduced-minus-oxidized spectrum, the coefficients are 26 mM^{-1} cm^{-1} at 554 nm and 28 mM^{-1} cm^{-1} at 554−538 nm. These values are similar to the ones for purified spinach and turnip cytochromes *f* (*49*).

The properties of this purified protein indicate that the nature of the truncated cytochrome *f* is very similar to that of the native form and that, therefore, it is a good model for studying structural−functional relationships.

Structural Determination. Crystals of the type used for this study have maximal and minimal dimensions of 1 and 0.3 mm, respectively. The cytochrome is crystallized in the reduced form. The crystals belong to space group $P2_12_12_1$ (a = 75.62 Å, b = 94.95 Å, and c = 120.20 Å). The asymmetric unit contains three monomers of the cytochrome.

Diffraction data were phased by molecular replacement using the coordinates of Brookhaven Protein Data Bank entry 1CTM for a truncated form of turnip cytochrome *f* (*16*) as a search model. A model of the three cytochrome molecules was built and refined against 2.0 Å diffraction data to a crystallographic *R*-factor of 21.4% and to an R_{free} of 27.1%. The detailed refinement statistics are listed in Table 1.

Figure 1 shows an "omit" map of the heme region, phased from a model in which the heme and residues Y1 and H25, which provide the fifth and sixth ligands to the heme iron in turnip (*16*) (see below), were omitted. This eliminates phase bias with respect to these residues. The electron density clearly shows heme axial ligation by the N-terminal amino group of Y1 and by Nδ of H25, as indicated in the model.

Gray's alignment:

```
                180         190         200
                 |           |           |
Turnip -  G-IISKILRKE-KGGYEITIVDASNERQVID
Chlamy -  GKIVAITALSEKKGGFEVSIEKA-NGEVVVD
                 |      *  |           |
               180        190         200
```

Structure-based: Turnip[a] and cyanobacterial[c] vs *Chlamydomonas reinhardtii*[b] cytochrome f:

```
                      180         190         200
                       |      *    |           |
Turnip[a]  - 177 GII**S**KILR-KE-KGG**Y**EITIVDASN**E**-RQVID 205
                 |||||||×-××-×|||||||||×--×-×||||
Chlamy[b]  - 177 GKI**V**AITALSEKK**G**GFEVSIEKA--**N**GEVVVD 206
                 |||||||×-×-××||||||||||--|||||||
Phormid[c] - 178 GV**I**TAIAK-A-DDG**S**AEVKIRTE--D**G**TTIVD 205
                       |      *  |           |
                      180         190         200
```

FIGURE 3: Sequence alignment of *C. reinhardtii* (Chlamy) with turnip cytochrome *f* based on Gray's alignment (*25*), compared with the alignment based on the structure (this work). In the latter, symbols between the lines indicate the closeness of superposition of the aligned residues, using the *C. reinhardtii* monomer B structure[b], turnip structure (entry 1CTM)[a], and *Phormidium laminosum* structure (entry 1CI3)[c] superimposed as described in the legend of Figure 4B. (|) Cα positions differ by less than 1.0 Å after alignment. (×) Cα positions differ by 1.0 Å or more. Insertions are placed to minimize the distance between the Cα's of aligned residues. The asterisk indicates the totally conserved carboxylate E186 (turnip), E187 (*C. reinhardtii*), and D187 (*25*).

Description of the Structure. The final model (PDB entry 1CFM) contains in the asymmetric unit three hemes, all the residues of three monomers, and 1001 water molecules. As we expect from sequence homology (72% sequence identity), the overall folding is very similar to that of *Brassica campestris* (turnip) (*16*). The protein is made of two distinct domains in an elongated form where the heme group sits within the large domain near the interface between two domains (Figure 2). The large domain has an immunoglobulin-like fold flanked by small α-helices, and the small domain has a jellyroll fold made up of four major and three minor β-strands. A schematic drawing of the secondary structure arrangement, which is very similar to that of the turnip cytochrome (*16*), is available as Supporting Information.

C. reinhardtii cytochrome *f* contains one more residue than the turnip protein, 286 amino acids rather than 285. Gray (*25*) aligned the sequences (Figure 3) using two insertions in *C. reinhardtii* and one in turnip, all in the region with a relatively low degree of homology between residues 169 and 204 (turnip).

Using this scheme, the backbone atoms of the large domain of the algal cytochrome *f* monomers and that of the turnip cytochrome *f* can be superimposed accurately (with rms deviations of ≤0.5 Å), with the exception of the truncated C-terminus (R251) and the loop after helix 1 (around residue 15). When this superposition is done (Figure 4a), it is seen that the small domain has a slightly different orientation in the four structures. This implies some flexibility of the hinge between the two domains. The difference in the orientation of the small domains in the three independent molecules can be described as a rotation of 2.3−3.2°. The rotations do not share a common axis, so the movement is not a simple hinge motion. When these angles between the two species are compared, larger rotation angles of 6.0−7.5° are observed. A similar small change in the relative orientation of the two folding domains was also observed when two turnip cytochrome *f* structures at different temperatures were compared (*17*).

In Gray's alignment,[3] residues 1−177 of *C. reinhardtii* align with residues 1−177 of turnip, *C. reinhardtii* residues 179−187 align with turnip residues 178−186, *C. reinhardtii* residues 189−199 align with turnip residues 187−197, and *C. reinhardtii* residues 201−251 align with turnip residues 200−250. However, when the backbone atoms of the small domain are superimposed using Gray's alignment, residues 179−183 of the algal protein do not superimpose on residues 178−182 of the turnip protein. A much better fit is obtained by superimposing *C. reinhardtii* residues 178−182 onto turnip residues 178−182 (Figure 4b). This implies that β-strand 13 of the small domain aligns with same numbered residues of turnip and that both insertions in *C. reinhardtii* should be located in the loop between β-strands[3] SC1 and SC2 in entry 1CFM (loop 1 in Figure 4b). This loop is relatively disordered, as indicated by high temperature factors (especially in monomer A) and large conformational differences between these residues in the three different *C. reinhardtii* structures (Figure 4b). Besides the conserved E186 (turnip; E187 in *C. reinhardtii*), this loop contains K187 (turnip; K188 and K189 in *C. reinhardtii*). These lysines are believed to be involved in plastocyanin binding (also see Lysine Cluster). The turnip insertion is in the loop between the β-strands[3] SC2 and SC3 in the small domain (loop 2 in Figure 4b). This loop also appears to be flexible and disordered in *C. reinhardtii*, especially in monomer B. The correct alignment based on the structure is compared with Gray's alignment in Figure 3.

Although the asymmetric unit contains three molecules of cytochrome *f*, in vivo the cytochrome would be part of a dimeric b_6f complex (*13*). The functionally homologous cytochrome c_1 forms part of the dimer interface in the bc_1 complex (*8*). The three monomers in the asymmetric unit are related by improper symmetry. The choice of asymmetric unit is arbitrary, and was made to give a relatively compact asymmetric unit. Because the crystal packing contacts affecting the three monomers are different, those features that are the same in all three monomers can be considered intrinsic features of the cytochrome and not artifacts of packing forces. Those parts of the structure that are different in the three monomers can be assumed to be flexible, and the differences probably are due to packing forces. To understand the crystal packing, and account for the different effects of crystal contacts on the three monomers, a list of all intermonomer distances that are less than 3.5 Å was made.[4] The analysis indicates that residues in the front face of the heme binding pocket (Figure 5), Y1 of monomer A and F4 of monomers B and C, have contacts with other monomers inside or outside the asymmetric unit. For residues of the Lys patch (see below), only one residue in each monomer has molecular contacts, specifically K188 of

[3] All residue numbers in this work refer to the sequence of the mature protein. For comparison with the sequence of the precursor that is available in databases, the length of the leader sequence (31 residues) must be added. Secondary structure elements of the protein are denoted as in Figure A of the Supporting Information, which corresponds to the helix and sheet anotation in PDB entry 1CFM.

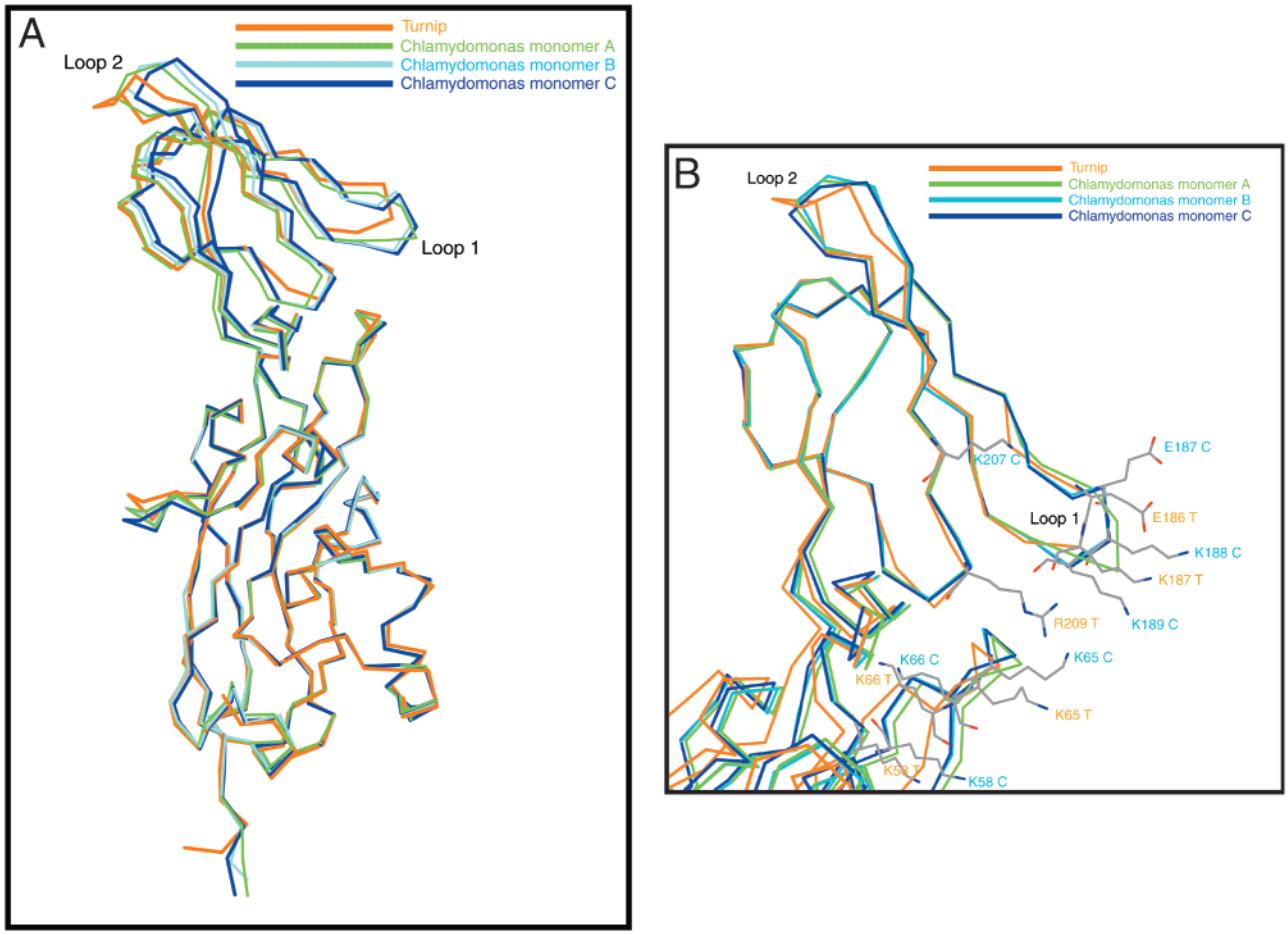

FIGURE 4: Comparison of the backbone representations of the three ncs-related monomers of *C. reinhardtii* (A−C) and of turnip cytochrome *f*. The molecules were superimposed so as to minimize the rms distance between corresponding Cα's. The "loop 1" and "loop 2" regions are discussed in the text. (a) The backbone of the full protein. Residues 1−13, 18−169, and 231−248 of the large domain of *C. reinhardtii* were matched with residues 1−13, 18−169, and 230−247 of turnip, respectively. For these 183 residues, the rms deviation in Cα positions between the turnip cytochrome and monomer B of *C. reinhardtii* cytochrome was 0.5 Å. The corresponding deviations of monomer A and monomer C with monomer B in *C. reinhardtii* were 0.3 Å, in both cases. (b) Detail of the small domain. Residues 176−182, 191−197, and 204−222 of the small domain of *C. reinhardtii* were matched with residues 176−182, 189−195, and 203−221 of turnip, respectively. For these 33 residues, the rms deviation in Cα position between the turnip cytochrome and the monomer B of *C. reinhardtii* cytochrome was 0.4 Å. The corresponding deviations of monomer A and monomer C with monomer B in *C. reinhardtii* were 0.2 Å, in both cases. Side chains for residues in the basic patch (putative plastocyanin binding site) are shown (with CPK coloring code) for the turnip structure and for monomer B of *C. reinhardtii* identified by T and C, respectively.

monomer A, K58 of monomer B, and K189 of monomer C. Structural differences between the three monomers are observed in the loop 1 region of the small domain (residues 185−190) (see Figure 4b). Here monomers B and C are nearly the same (Figure 4B), as expected from the involvement of this region in similar interactions (contacts 1 and 2 of footnote 4) which involve S186 and (in monomer C) K189, while monomer A is somewhat different due to the involvement of K188 in contact 3. The turnip structure is different in this region due to insertion of residues L185 and K188 in *C. reinhardtii*.

Dimeric Relation between Two Cytochrome f Monomers. Although the three monomers in the chosen asymmetric unit are related by improper symmetry, the relation between monomer B in one asymmetric unit and monomer A in another is a distorted (176°) proper 2-fold rotation (described as contact 5 in footnote 4). The possibility that this represents a physiological dimeric state will be approached in the Discussion. The putative dimer is depicted in Figure 6, with the likely position of the membrane shown assuming (for lack of other information) that the cytochrome *f* dimer 2-fold is perpendicular to the plane of the membrane. The binding mode of *C. reinhardtii* palstocyanin based on the work of

[4] The individual cytochrome monomers pack together in the crystal forming seven different intermonomer interfaces. (1) The interface between monomers A and B in the asymmetric unit involves H-bonds between E34, N37, and K145 of monomer A and S186, A21, A260, and K165 of monomer B, and close approach of A250 and Y143 of monomer A to P117 and F4 of monomer B. (2) Monomer B has an interface with monomer C in a different asymmetric unit which is nearly the same as the interface of A and B, respectively, described above with the addition of two H-bonds: one between R251 of B and P117 of C and one between E34 of B and K189 of C. (3) The interface between monomers A and C in the chosen asymmetric unit involves H-bonds between Y1, Q7, N63, and K188 in A and N16, N23, and E241 in C. (4) K178 in monomer A forms H-bonds with T43 and D42 in monomer C of another asymmetric unit. (5) Monomer B is related by a distorted 2-fold rotation of 176° to monomer A of another asymmetric unit. This brings the side chain N of N231 in the two monomers into H-bonding distance. Also, K217 of each monomer ion pairs with E14 of the other, and N231 N of each H-bonds with L26 O of the other. The guanidino group of R18 in monomer A H-bonds with the carbonyl of L214 in B, but the complementary interaction does not occur. (6) K58 of monomer B H-bonds with V203 of monomer B in another asymmetric unit. (7) N99 of monomer C H-bonds with D87 of monomer A in another asymmetric unit.

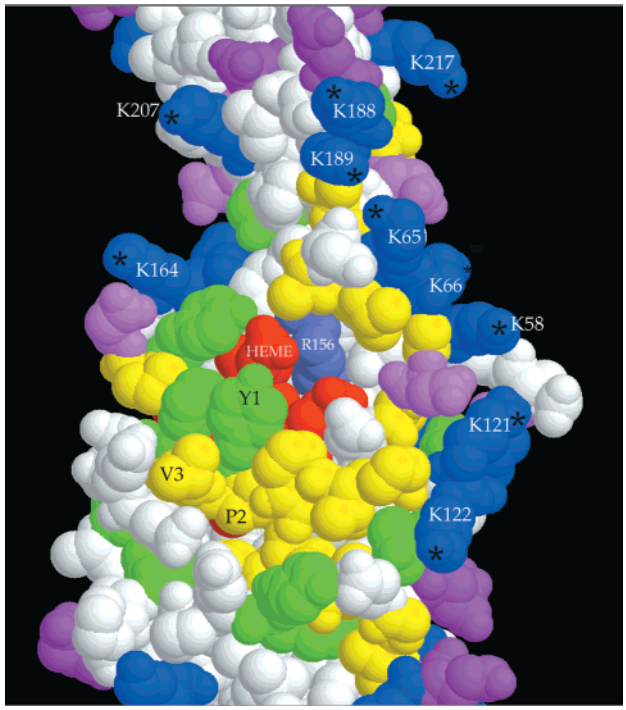

FIGURE 5: Space-filling Rasmol (*44*) representation of the central part of the *C. reinhardtii* truncated form of cytochrome *f* (monomer B). The hydrophobic outer face of the heme binding pocket (involving the N-terminal region) and the Lys cluster (putative plastocyanin binding site) are shown (see the text). The top and base of the figure are areas distal and proximal to the membrane, respectively. Color code: blue, Lys; gray-blue, Arg; pink, Glu; magenta, Asp; yellow, 12 aliphatic, 4 Gly, and 4 Pro residues; green, aromatic; white, polar; and red, heme (only the two propionate chains are clearly seen in this view). Due to the molecule orientation, exposed water molecules W5−W7 are not seen. The Lys ζN groups are identified with asterisks. Notice that K66 ζN is not in the same plane with the other ζN's in the cluster and is almost not seen in this view. This figure encompasses approximately two-thirds of the large domain (bottom) and one-half of the small domain (top).

Ubbink et al. (*50*) is also depicted. This model is further described in the Discussion.

Heme Ligation. As seen in the cytochrome *f* structure from turnip (*16*), the heme is covalently bound to the protein by thioether bonds through strictly conserved cysteine residues (Cys21 and Cys24, in the fingerprint peptide Cys-X-Y-Cys-His sequence in all *c*-type cytochromes). The heme Fe is ligated by the four tetrapyrrole nitrogens, by a proximal histidine group of H25, and by the α-amino group of the N-terminus residue Y1 (Figure 1). This type of coordination was shown for the first time in turnip cytochrome *f*. The distances between the Fe atom and Nδ of H25 or N of Y1 are 2.7 Å, in both cases. In all three monomers, the aromatic ring of Y1 lies approximately parallel to the plane of the heme ring (Figures 1 and 5). The heme is almost completely buried in a hydrophobic pocket (established by highly conserved residues such as Y1, P2, F4, F237, G72, V74, I19, V20, A5, and P161) that shields the redox center from the solvent (Figure 5). The only portions of the heme that are exposed to the surface are the two propionic acid side chains (whose negative charges are partially counterbalanced by the fully conserved residue R156; see Figure 5) and one of its saturated vinyl groups, barely seen at the left of Figure 5.

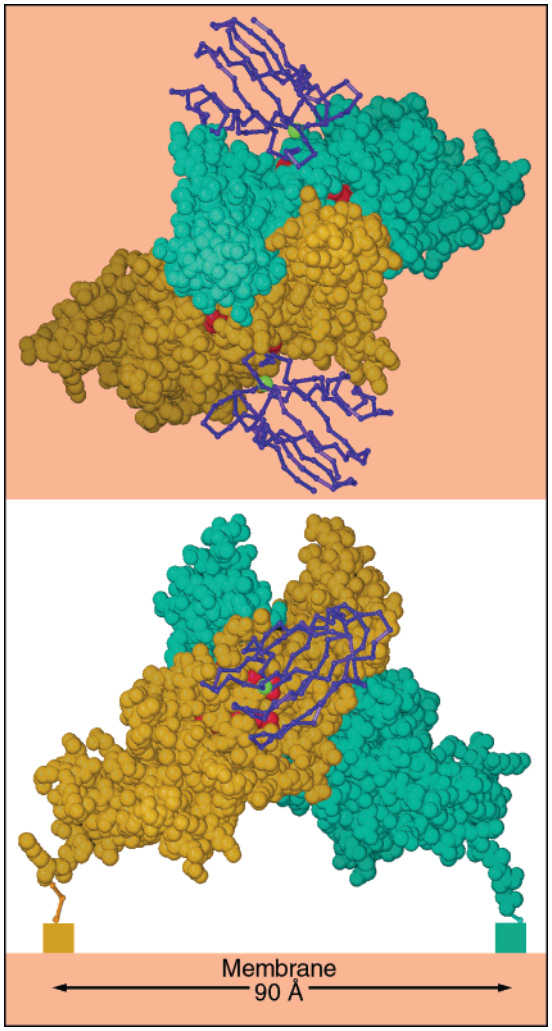

FIGURE 6: Dimer-like noncrystallographic symmetry observed in the crystal of *C. reinhardtii* cytochrome *f*, and interaction with plastocyanin modeled from ref *50*. Above is a top view, looking down the 2-fold axis of symmetry. Below is a side view, looking parallel to a hypothetical membrane plane (orange) constructed perpendicular to the 2-fold axis. The blue and yellow space-filling models are monomers A and B from the crystal of *C. reinhardtii* cytochrome *f*, respectively. The hemes are shown as red space-filling models. The rectangles below the C termini in the lower panel represent the lumenal ends of the transmembrane helices. The magenta backbone drawings are *C. reinhardtii* plastocyanin (*66*) oriented as in the NMR structure of the plastocyanin−cytochrome *f* complex (2PCF), oriented with the operators that best superimpose cytochrome *f* of that structure on each monomer of the *C. reinhardtii* dimer. The copper atom of plastocyanin is shown as a green sphere. The interactions betwen these two monomers in the crystal may be similar to those involved in stabilizing a physiological dimer in vivo (see the text). In this figure, monomer B is taken directly from structure 1CFM; monomer A is transformed by crystallographic symmetry operator [$^1/_2 + x$, $^1/_2 - y$, $1 - z$] because the chosen asymmetric unit does not include this dimer. For positioning plastocyanin, the rotation translation operator giving the best superposition of *C. reinhardtii* plastocyanin in entry 1PLT with poplar plastocyanin in entry 2PCF (over residues 5−45 and 65−95) was multiplied on the left by the operator giving the best superposition of turnip cytochrome *f* with *C. reinhardtii* cytochrome *f* monomers A and B oriented as described above (superimposing residues 1−4, 27−30, 45−52, 68−76, 111−117, and 142−159 of turnip with the same residues in *C. reinhardtii* and residues 234−238 of turnip with residues 235−239 of *C. reinhardtii*).

Lysine Cluster. Figure 7 compares the electrostatic potentials at the molecular surfaces of cytochrome *f* from *C.*

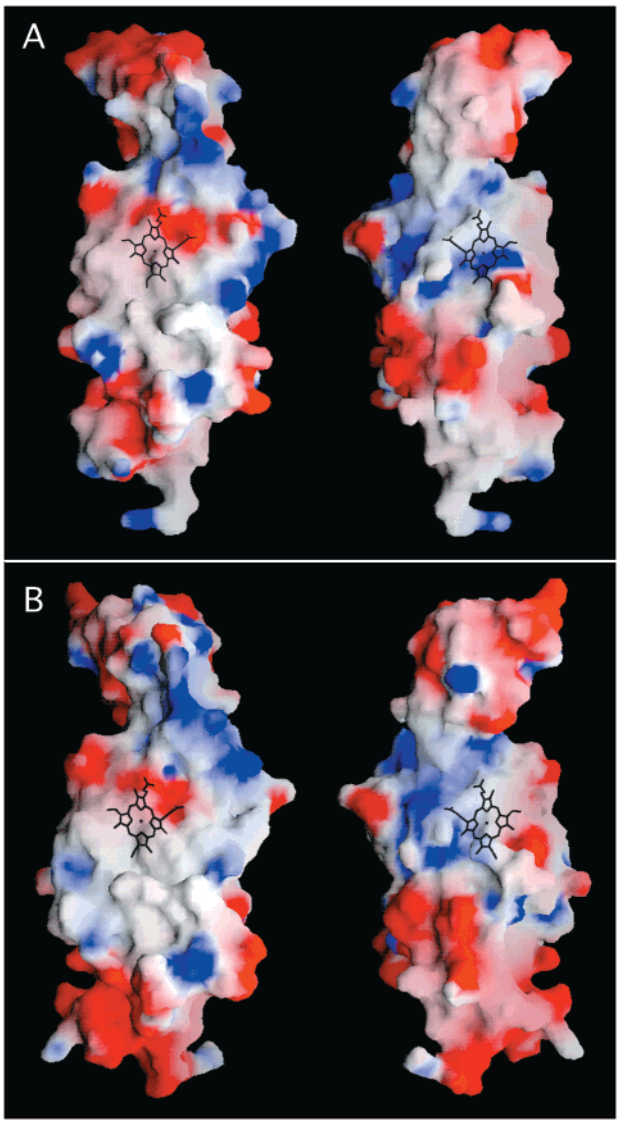

FIGURE 7: Electrostatic potentials at the molecular surface of the reduced forms of cytochromes *f* from *C. reinhardtii* (monomer B) (A) and from turnip (coordinates from 1CTM) (B). Two views rotated 180° around the longest dimension of the molecules are shown. The orientation of the molecules at the left in each panel is similar to the one shown in Figure 5. The surface electrostatic potentials were calculated with the program GRASP (*43*) and displayed as a color gradient from red (electronegative, ≤ -5 kT/charge) to blue (electropositive, ≥ 5 kT/charge). The uncharged hydrophobic environment around the heme pocket and the overall positive potentials of the lysine cluster, with some differences between cytochromes *f* from two species, are seen. The dark stick figures indicate the location and orientation of the heme moiety, mainly buried in the protein (see Figure 5) with only the propionate groups exposed, if observed in the left view (see Figure 5), and one vinyl group exposed, if observed in the right view. In the calculation of electrostatic potentials, hydrogen atoms were omitted and their charges included in the charge of the atom to which they are bound (the default for GRASP). A charge of 1 was used for the Nζ atoms of Lys, 0.5 for Nη1 and Nη2 of Arg, and −0.5 for both carboxylate oxygens of Asp, Glu, and the heme propionates. All other atoms of the heme were considered neutral, which is appropriate for the reduced heme.

reinhardtii (A) and from turnip (B) in their reduced forms. Both basic and acidic patches are seen. These charge clusters are natural candidates for sites involved in the stabilization of the quaternary structure of the b_6f complex and in the (transient) association with the intermediary electron transfer carriers (e.g., plastocyanin or cytochrome c_6).

In the case of turnip cytochrome *f*, a prominent positively charged region at the interface between the large and small domains that contains basic residues K58, K65, K66, K187, and R209 (*16*) has been proposed by many in vitro studies, including kinetic, biochemical, and cross-linking studies (see the references in refs *31* and *50*), electrostatic analysis (*51*−*53*), and NMR studies (*50*), to be the site of electrostatic interaction with plastocyanin. The first four of these residues are highly conserved in eukaryotic cytochromes *f* (*25*). On the other hand, most of the acidic residues in two structurally conserved, closely associated, and negatively charged patches on the surface of eukaryotic plastocyanins are also conserved. These last patches are proposed to be the counterpart binding site for cytochrome *f* (*21*, *54*, *55*).

In *C. reinhardtii* cytochrome *f*, the equivalent cluster of basic residues, also contributed by both the large and small domains, would be composed of K58, K65, K66, K188, and K189 (Figure 5). Notice that the Lys patch and the heme binding pocket are separated from each other through both a translation and a ca. 60° rotation along the molecule main axis. The temperature factors for the side chains of those residues in all three monomers are relatively high, indicating conformational flexibility.[5]

The largest difference in the electrostatic landscape around the heme face and lysine patch is due to the presence of aspartate 68 in *C. reinhardtii* where the turnip protein has alanine and most other cytochromes *f* have a neutral residue. This results in the negative (red) patch to the right of the negative patch around the heme propionates in Figure 7A. It also results in the complete masking (at the surface) of the positive charge due to arginine 156, seen as a blue spot above the heme propionates in the turnip (Figure 7B) which is absent in the *C. reinhardtii* electrostatic surface. Recalculation of the *C. reinhardtii* surface charge after removing the charge from the carboxylic oxygens of D68 resulted in essentially the same charge distribution as seen for turnip in this region (not shown).

The lysine cluster, as defined originally for turnip, could be extended to K121, K122, and K217 in the microalga (Figure 5). The ζN atoms of K58, K65, K188, K189, and K121 are approximately in the same plane, whereas that of K66 is not, as its lateral chain points out of the cluster (Figures 4b and 5). This characteristic of K66 is also seen in the turnip structure (entry 1CTM) (see Figure 4b).

Water Chain. An internal chain of buried water molecules extending from the heme ligand H25 to a position close to K66 has been found in the cytochrome *f* structure from turnip (*17*). The authors proposed that it could constitute a "proton wire" (*56*, *57*) and that, as such, it could be involved in a

[5] The ranges of temperature factors of Lys Nζ in the cluster detected in the three monomers are 36−40, 30−40, 29−40, 45−78, and 37−78 Å^2 for K58, K65, K66, K188, and K189, respectively. Monomer C exhibits the highest temperature values for ζN in K58, K65, and K66, whereas monomer A exhibits the highest ones for ζN in K188 and K189. Apart from the cases of ζN in K188 and K189 in monomer A (which exhibit also the highest temperature values for ζN in the entire structure, possibly indicating multiple conformations), those temperature factors are not different from the ones corresponding to all other ζN's outside the cluster. For the three monomers, they are close to the median of the broad distributions shown by the total of 21 lysines per monomer. The lowest ζN temperature factor is 25 Å^2.

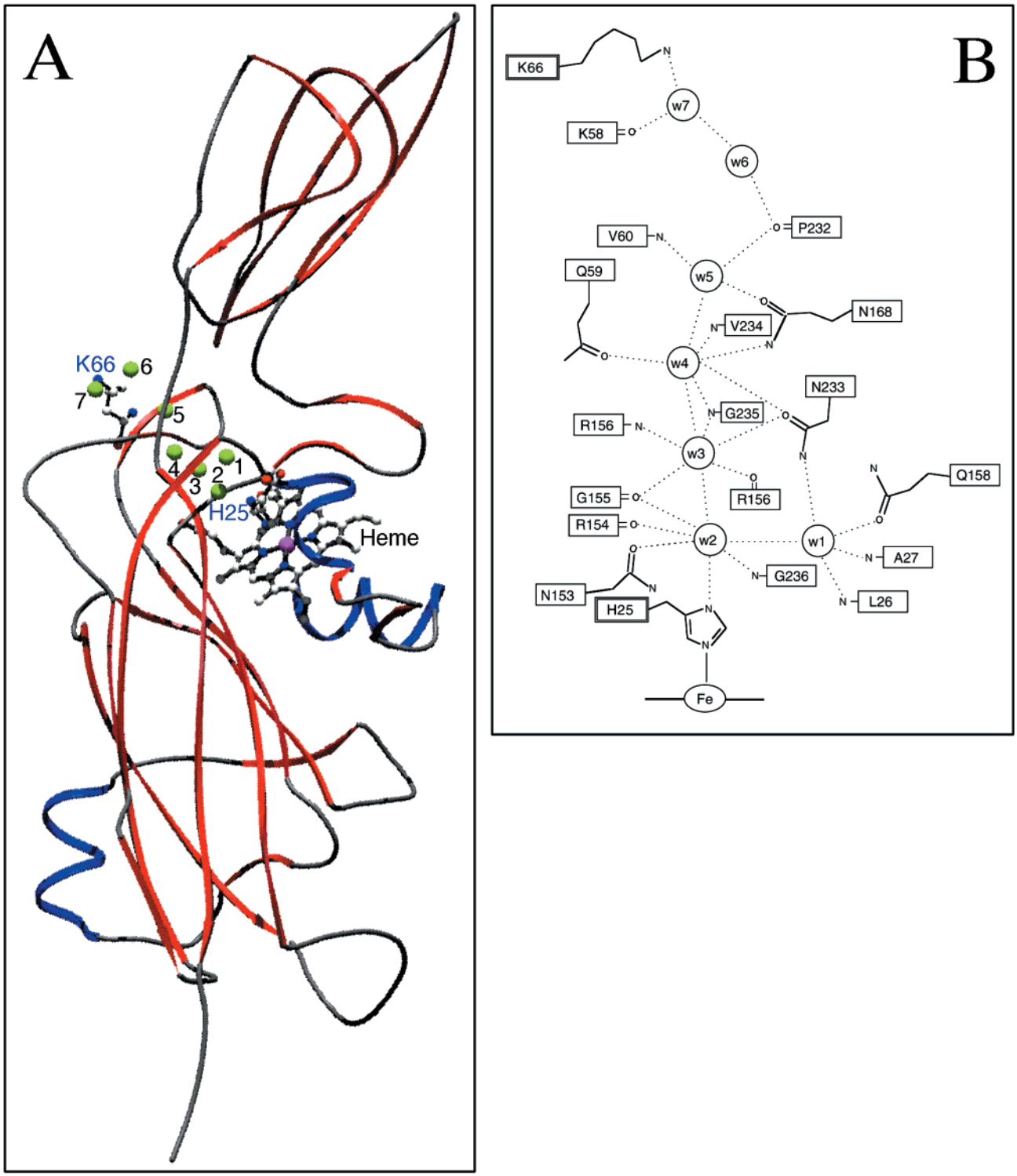

FIGURE 8: "Water chain". (A) Locations of the seven water molecules forming the water chain (W1−W7, numbered green spheres; see the text) are shown relative to the whole molecule. Residues H25 and K66, at the ends of the water chain, as well as the heme moieties are indicated in a ball-and-stick representation. (B) Schematic representation of the hydrogen-bonding network directly involving the buried (W1−W4) and exposed (W5−W7) water molecules of cytochrome *f* from *C. reinhardtii*. Side chains are not shown for those residues whose backbone atom is forming a hydrogen bond. All the residues shown in this network are strictly conserved in cytochromes *f* from *C. reinhardtii* and higher plants (*25*).

pathway of proton transfer from the Rieske iron−sulfur protein, through cytochrome *f*, and into the chloroplast lumen, following the oxidation of plastoquinol at the quinol oxidation site.

In our structure for *C. reinhardtii*, the chain of five buried water molecules previously reported in the turnip cytochrome *f* structure (*17*) is seen in all three independent monomers. In addition, two more water molecules (W6 and W7 in Figure 8 and Table A in the Supporting Information) can be identified in each monomer, which extend the water chain all the way up to ζN of K66, a surface residue that, in situ, is believed to be in contact with bulk water and interacting with plastocyanin (Figure 8A). The locations of the five buried water molecules are identical, and their hydrogen-

bonding network is almost identical to the respective information reported for the turnip cytochrome *f* structure (*17*). The two additional waters correspond to water residues 107 and 135 of the turnip structure. The hydrogen-bonding network of the water chain is shown in Figure 8B, and the refined parameters and residue numbers in the submitted coordinates file (1CFM) are listed in Table A in the Supporting Information. In comparison with the turnip structure 1HCZ, the OD1 and ND2 atoms of water ligands N153 and N168 are interchanged in the present structure as determined by H-bond analysis by the verification program What-Check; however, as these atoms cannot be distinguished by the electron density, the difference may not be real.

The exposure of these molecules to the solvent is partial for W5 and complete for W6 and W7, all three of them being less ordered than W1−W4, which are completely buried. There are no other completely buried waters in the cytochrome. Water molecules W5−W7 are visible on the surface in a space filling model; however, they are not seen in the orientation of Figure 5. The inter-water hydrogen-bonding chains are broken between W5 and W6 (distances of 4.0, 3.8, and 4.1 Å in monomers A−C, respectively); however, the hydrogen bond chain is complete through hydrogen bonding with the P232 carbonyl oxygen.

These water sites have relatively low temperature factors (a parameter expressing the dynamic vibrations of an atom as well as its static disorder in the crystal), and each water's position is within 0.2 Å of the average position for that water in the three independent monomers. To compare the degree of order of W1−W7 with the rest of the water molecules in the asymmetric unit, we analyzed the distribution of temperature factors of the whole population (1001 waters in identified positions).[6] This analysis indicates that the degree of localization of all these seven waters, although exhibiting levels lower than the mean and median, is also shared by many other water molecules in the crystal.

DISCUSSION

Cytochrome *f* is a subunit of the cytochrome b_6f complex, a fundamental redox enzyme in oxygenic photosynthetic electron transport. The folding and structural features described for turnip cytochrome *f* (*16*, *17*) were confirmed here for three noncrystallographically related monomers of cytochrome *f* purified from the microalga *C. reinhardtii* that, due to its genetic amenability, is the preferred model organism for eukaryotic photosynthesis. While significant differences are observed between the turnip and the algal cytochromes in loops and at points of insertion or deletion in the sequence, they are not much greater than the ones found between the structures of the three monomers for the algal cytochrome which, being from the same crystal, were obtained under identical conditions of ionic strength, pH, and buffer composition. The two insertions and one deletion in the amino acid sequence of the algal protein occur at turns between β-strands and do not affect the overall folding. The insertion of one Lys does affect the composition of the Lys cluster (see below), which might be functionally significant in modifying specificity for the different electron acceptor partners.

The main structural characteristics of cytochrome *f* that are currently discussed with regard to function are (a) the conserved residues of the heme pocket and the unusual heme Fe ligation by α-amino group of tyrosine 1, (b) the cluster of lysine residues that is proposed to be the docking site of plastocyanin, and (c) the chain of seven water molecules bound to conserved residues that extends between the heme pocket and K58 and K66 at the lysine cluster.

Front Face of the Heme Binding Pocket. The significance of the conserved residues in the N-terminal end of the protein (*25*) has been tested with site-directed mutagenesis in *C. reinhardtii* (*32*, *58*). Mutants of Y1 or P2 exhibited normal cytochrome *f* photooxidation but with its reduction inhibited to various degrees. This indicated that residues 1 and 2 in the front face of the heme pocket play a direct or indirect role in the redox equilibration with the Fe−S cluster in the Rieske subunit, but not with plastocyanin (*32*). These results are in contradiction with simulations (*51*, *52*), and the NMR resolution (*50*) of the plastocyanin−cytochrome *f* complex configuration. All these works show that Y1 is involved in the electron transfer path between both molecules. The paradox will be solved with the structural resolution of the cytochrome b_6f complex.

Flexibility in the Orientation of the Small and Large Domains and of the Lateral Chains at the Conserved Lysine Patch. The lysine patch in *C. reinhardtii* involving K58, K65, K66, K188, and K189 is very similar to the one in turnip. Nevertheless, the following differences are found between both proteins. Considering that in turnip the span from residue 202 to 221 contains a single basic residue (R209) while in *C. reinhardtii* in the span from residues 190 to 216 there is a single basic residue (K207), it can be proposed that K207 is the charged residue counterpart of R209. Nevertheless, K207 in *C. reinhardtii* is not associated with the Lys cluster described above (Figures 4b and 5), in contrast with the case of R209 for turnip (*16*). Besides, K121 and K217 which could be proposed as participants in the Lys cluster in *C. reinhardtii* (see Figure 5) are replaced by Q121 and S216 in turnip, respectively. These subtle differences in the Lys cluster in terms of charges and positions due to mutations, as well as insertions and deletions (Figures 3 and 4b), could be responsible for some differences in the electron partner specificity of the two species. For example, the *C. reinhardtii* cytochrome *f* can interact with both plastocyanin and cytochrome c_6 as mobile electron acceptors (*22*, *31*), whereas higher-plant cytochrome *f* interacts in situ with only plastocyanin (*22*).

The role of these conserved residues in the redox reaction has been studied in *C. reinhardtii* by site-directed mutagenesis. With a series of mutants, where K58, K65, K66, K188, and K189 have been replaced mainly with polar residues or Glu, it has been shown that, in contrast with what was expected, the complete neutralization of the charges in the Lys cluster does not affect the kinetics of photooxidation of cytochrome *f* at pH 7 in uncoupled intact cells (*31*, *48*), although reactions in permeabilized cells (*48*, *59*) and in vitro (*60*) are inhibited. The in vivo results are also in disagreement

[6] The range of values detected (2.4−75.4 Å^2) was divided in intervals of 1.25 Å^2. The resulting distribution is broad and, probably, multimodal. The global mean, median, and standard deviation of the whole distribution are 36.6, 35.7, and 13.8 Å^2, respectively. In comparison, and for all three monomers, temperature factors of W1−W5 lie before the first 13% of the whole distribution, whereas W6 and W7 can have values extending before the median.

with theoretically predicted (*51*−*53*) and NMR-resolved (*50*) structures of the reactive complex. The reason for the discrepancy is unsettled.

The flexibility shown by these Lys lateral chains in the crystal is still consistent with this cluster being involved in the plastocyanin binding site, because a conformationally mobile area may be instrumental in allowing an "induced fit" to different electron acceptors (plastocyanin or cytochrome c_6) or in guiding the approach of them from different orientations (see below). Decreased mobility of these residues upon complexation could be expected, as seen in other protein complexes (*61*).

The three monomers show flexibility in the relative position of the constituent large and small domains. The interface between these domains involves a patch of Lys residues (putative binding site for plastocyanin) with relatively mobile side chains. This flexibility might be needed to accommodate different electron transfer partners, such as plastocyanin or cytochrome c_6, and to achieve a conformational rearrangement after the initial docking (*62*). It can also be interpreted in terms of the "Velcro" model of a binding site (*63*−*65*). According to this model, there is not one unique specific lock-and-key interaction, but rather the two reacting molecules can interact through multiple different configurations with similar energies. Such a dynamic configuration of the binding between plastocyanin and cytochrome *f* has been deduced by NMR (*50*).

Since high-resolution structures are available for both plastocyanin (*66*) and cytochrome c_6 (*67*) from *C. reinhardtii*, this is a particularly good system for studying the interaction of cytochrome *f* with these redox partners. In the model of Figure 6, we have used *C. reinhardtii* cytochrome *f* and plastocyanin, oriented by superposition of the molecules with poplar plastocyanin and turnip cytochrome *f* from the NMR structure of the complex (*50*). Cytochrome *f* monomer A makes a particularly good interaction with plastocyanin. Lysines 188 and 189 have their Cα positions about 5.7 Å from the Cα of plastocyanin D53 and E43, respectively. Asn 63 is 6.2 Å from D59 of plastocyanin [*C. reinhardtii* plastocyanin sequence numbering as in entry 2PLT (*66*) to match spinach plastocyanin]. Note that K188 (cytochrome *f*) and D53 (plastocyanin) are not conserved in the turnip and poplar proteins from which the complex structure was determined.

The interdomain flexibility could be important in the mechanism of interaction of cytochrome *f* not only with its oxidant partner(s) but also, in alternation, with the Rieske protein, involving an oscillatory catalytic mechanism analogous to the one proposed in ref *8*. Finally, the interdomain flexibility could have a role in the correct association of cytochrome *f* with other interfacing subunits of the cytochrome b_6f complex. In the model depicted in Figure 6, the interface between the two cytochrome *f* molecules involves mainly interlocking of the interdomain notches.

An Alternative Structural Role for the Water Array. As in turnip, an array of buried waters is also present in *C. reinhardtii* cytochrome *f*. A proton wire function has been proposed for this chain (*17*). The residues participating in side chain hydrogen bonds such as Q59, N153, Q158, N168, and N233 (N232 in turnip) are strictly conserved in all cytochromes *f* whose sequences are known (*25*). Furthermore, those residues participating in backbone hydrogen bonds are also strictly conserved in all cytochromes *f* with only one exception, K58 which is replaced with Q in cyanobacteria (*25*).

Similar chains of hydrogen-bonded water molecules have been previously observed in bacteriorhodopsin (*68*) and in a bacterial photosynthetic reaction center (*69*−*70*) and have been proposed to function as proton wires. Nevertheless, the proposed proton translocation function of this conserved water array in cytochrome *f* has not yet been demonstrated or, in principle, expected, as protons generated at the quinol oxidizing site should have shorter pathways by which to reach the lumen than one involving, in series, the Rieske and cytochrome *f* proteins. Mutational analysis of the residues associated with the cytochrome *f* water array has not yielded a unique answer yet (*71*).

In addition to the proton wire function, another purely structural (hydrogen bonding) role can be postulated. The analysis of the structural data shown in Figure 8B and Table A (Supporting Information) is consistent with the possibility of a structural role in the large domain. Most significantly, due to their complete or partial enclosure in the protein body and low temperature factors, W1−W5 make hydrogen bonds between some of themselves and between stretches of residues that are far away in sequence. In one case, they relate at the same time with residues placed 80 positions apart, i.e., residues 153−156, 158, and 168 with residues 232−236. Consequently, as can be seen from Figure 8B and Table A,[3] one end of β-strands SB3 and SB4, four residues immediately upstream of SB4, two residues immediately downstream of SB3, and Q158 and N168 in a loop between the small and large domains are connected by an abundant network of water-dependent hydrogen bonds. In the other case, W1 and W2 also establish bonds with the loop between SA1-H2 that is close to the N-terminal portion, more than 100 and 200 residues apart from SB3 and SB4, respectively. We can consider this structure as a "core" contributed by different secondary structure elements of the protein that is stabilized, at least in part, by hydrogen bonds mediated by waters. Probably less significant, due to their higher temperature factors and exposure, W6 and W7, with W4 and W5, connect to the core the loop between strands BB1 and BB2, which bear the conserved K58, K65, and K66. Other cases of buried water molecules, either single or in arrays, are known (*72*−*75*). Even water molecules bound to the protein surface can have a role in protein folding, stability, recognition, and activity (*76*).

Another explanation for the presence of the internal water array would consider these waters as resulting from their entrapment during the folding process. Those waters, without any functional role, would be stabilized by circumstantial hydrogen bonds to the indicated residues, which would have been conserved for other reasons.

Putative in Situ Dimeric Configuration of Cytochrome f. The physiological state of the membrane-embedded cytochrome b_6f complex is a dimer (*9*−*12*, *14*), and the functionally homologous cytochrome c_1 is involved in the dimer contact in the mitochondrial cytochrome bc_1 complex (*8*). If cytochromes *f* of the dimeric b_6f complex also contribute to the dimer interface, cytochrome *f* may have an affinity for itself which could be revealed in the crystal packing. Of the seven different types of intermonomer contacts observed in the crystal, the only one which

approximates 2-fold symmetry is that described as contact 5 in footnote 4. This association is the most intimate of the intermonomer contacts, with 3033 Å^2 buried in dimer formation (about 1500 Å^2, or more than 10% of the monomer surface area, is buried on each monomer). It is thus tempting to speculate that this association between two monomers in the crystal represents the arrangement of monomers in the dimeric cytochrome b_6f complex, as suggested by Figure 6.

This contact involves both large and small domains; in fact, the "necks" between the domains are interlocked in a way that provides considerable rigidity to the putative dimer. The lysine patch and the hydrophobic heme face (likely to be involved in the interaction with plastocyanin) are exposed. A solution structure for the complex of poplar plastocyanin with turnip cytochrome *f* is available (*50*) (PDB entry 2PCF). As illustrated in Figure 6, it is possible for plastocyanin to bind either monomer, in the manner revealed by this NMR structure, without interference from the other monomer. We assume that the dimer 2-fold axis would be perpendicular to the membrane as indicated in Figure 6. Both C-termini of the truncated proteins are close to the membrane surface in this model, as expected if the truncated portion is a transmembrane anchor.

A functional dimer of cytochrome c_6, which like plastocyanin mediates between cytochrome *f* and PS I under some conditions, has been postulated by Kerfeld and co-workers (*67*) and is based on crystal contacts. Our proposed dimeric cytochrome *f* would not be consistent with one of the physiological roles suggested for the cytochrome c_6 dimer, efficient intercomplex transfer of electrons from dimeric cytochrome *f*, because the heme irons of dimeric cytochrome c_6 are only separated by about 18 Å. It would be consistent with three other possible roles proposed by those authors. (a) Dimeric cytochrome c_6 could allow efficient delivery of electrons to trimeric PSI complexes (although here again one suspects the receiving centers will be rather too far apart). (b) Dimerization of cytochrome c_6 after reduction by cytochrome *f* could occlude the C corner of the heme to inhibit electron return to cytochrome *f*. (c) Dimerization of cytochrome c_6 in response to ionic composition and pH could regulate electron transfer based on the electrochemical potential gradient accross the thylakoid lumen (*67*).

In the model depicted in Figure 6, the normal to the heme plane is 79° from the dimer 2-fold axis; therefore, the heme plane is 79° from the plane of the membrane. This is similar to the value (74°) calculated for cytochrome c_1 from entry 1BCC. However this is inconsistent with the values of ~25−30° estimated by EPR spectroscopy of oriented membranes (*77*−*79*). Thus, unless there is some artifact in the EPR measurements, our dimer from the crystal is probably not a good model for the cytochromes *f* in the dimeric cytochrome b_6f complex.

Perspectives. Despite functional similarities, the structure of cytochrome *f* is different from that of mitochondrial cytochrome c_1 which is mainly made of α-helices and has different heme ligation (*8*). Because they share a low degree of sequence identity and have different cellular origins (cytochrome *f* is organelle-encoded and cytochrome c_1 is nucleus-encoded in eukaryotes), they might represent a convergent functional evolution (*80*).

As cytochrome *f* sequences are highly conserved across eukaryotes and cyanobacteria (*25*), the structures presented here together with those from turnip (*17*) and cyanobacteria (*81*) should serve as a prototype for the main structural features of all cytochromes *f*. Small differences could be attributed to functional specificities or evolutionary drift.

Free subunits of macromolecular complexes, although easier to crystallize than the entire complex, may have different structural and physicochemical characteristics than they do in situ, where a variety of interactions with other domains exists. Nevertheless, these results will allow structure−function studies with solid structural foundation, and will help in the resolution of the full cytochrome b_6f complex structure, once suitable crystals are obtained.

ACKNOWLEDGMENT

We thank Dr. F.-A. Wollman (Institut de Biologie Physico-Chimique) for providing strain F283ST, Prof. S.-H. Kim (Calvin lab, University of California, Berkeley) for the use of his crystallographic facilities, Ms. A. Reyes and Mr. R. Solem for assistance in growing the microalga, Mr. T. Linden for assistance in cytochrome *f* purification and crystallization, and Mr. L. R. Comolli for helpful discussions. J.G.F.-V. thanks Prof. R. Malkin (Department of Plant and Microbial Biology, University of California, Berkeley) for generously providing laboratory facilities and for critically reading the manuscript.

SUPPORTING INFORMATION AVAILABLE

Figure A, a schematic diagram defining the secondary structure elements of cytochrome *f*, and Table A giving details of the hydrogen-bonding network of the buried water chain. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Degli Esposti, M., DeVries, S., Crimi, M., Ghelli, A., Patarnello, T., and Meyer, A. (1993) *Biochim. Biophys. Acta 1143*, 243−271.
2. Trumpower, B. L., and Gennis, R. B. (1994) *Annu. Rev. Biochem. 63*, 675−716.
3. Hope, A. B. (1993) *Biochim. Biophys. Acta 1143*, 1−22.
4. Gray, K. A., and Daldal, F. (1995) in *Anoxygenic Photosynthetic Bacteria* (Blankenship, R. E., Madigan, M. T., and Bauer, C. E., Eds.) pp 747−774, Kluwer Academic Publishers, Dordrecht, The Netherlands.
5. Malkin, R. (1992) *Photosynth. Res. 33*, 121−136.
6. Hauska, G., Schutz, M., and Buttner, M. (1996) in *Oxygenic Photosynthesis: The Light Reactions* (Ort, D. R., and Yocum, C. F., Eds.) pp 377−398, Kluwer Academic Publishers, Dordrecht, The Netherlands.
7. Xia, D., Yu, C. A., Kim, H., Xia, J. Z., Kachurin, A. M., Zhang, L., Yu, L., and Deisenhofer, J. (1997) *Science 277*, 60−66.
8. Zhang, Z., Huang, L., Shulmeister, V. M., Chi, Y.-I., Kim, K. K., Hung, L.-W., Crofts, A. R., Berry, E. A., and Kim, S.-H. (1998) *Nature 392*, 677−684.
9. Chain, R. K., and Malkin, R. (1991) *Photosynth. Res. 28*, 59−68.
10. Huang, D., Everly, R. M., Cheng, R. H., Heymann, J. B., Schagger, H., Sled, V., Ohnishi, T., Baker, T. S., and Cramer, W. A. (1994) *Biochemistry 33*, 4401−4409.
11. Chain, R. K., and Malkin, R. (1995) *Photosynth. Res. 46*, 419−426.
12. Breyton, C., Tribet, C., Olive, J., Dubacq, J. P., and Popot, J. L. (1997) *J. Biol. Chem. 272*, 21892−21900.
13. Mosser, G., Breyton, C., Olofsson, A., Popot, J. L., and Rigaud, J. L. (1997) *J. Biol. Chem. 272*, 20263−20268.

14. Bron, P., Lacapere, J. J., Breyton, C., and Mosser, G. (1999) *J. Mol. Biol. 287*, 117−126.
15. Huang, D., Zhang, H., Soriano, G. M., Dahms, T. E. S., Krahn, J. M., Smith, J. L., and Cramer, W. A. (1998) in *Photosynthesis: Mechanisms and Effects* (Garab, G., Ed.) pp 1577−1581, Kluwer Academic Publishers, Dordrecht, The Netherlands.
16. Martinez, S. E., Huang, D., Szczepaniak, A., Cramer, W. A., and Smith, J. L. (1994) *Structure 2*, 95−105.
17. Martinez, S. E., Huang, D., Ponomarev, M., Cramer, W. A., and Smith, J. L. (1996) *Protein Sci. 5*, 1081−1092.
18. Iwata, S., Saynovits, M., Link, T. A., and Michel, H. (1996) *Structure 4*, 567−579.
19. Zhang, H., Carrel, C. J., Huang, D., Sled, V., Ohnishi, T., Smith, J. L., and Cramer, W. A. (1996) *J. Biol. Chem. 271*, 31360−31366.
20. Carrell, C. J., Zhang, H., Cramer, W. A., and Smith, J. L. (1997) *Structure 5*, 1613−1625.
21. Gross, E. L. (1996) in *Oxygenic Photosynthesis: The Light Reactions* (Ort, D. R., and Yocum, C. F., Eds.) pp 413−429, Kluwer Academic Publishers, Dordrecht, The Netherlands.
22. Wood, P. M. (1978) *Eur. J. Biochem. 87*, 9−19.
23. Merchant, S., and Bogorad, L. (1986) *Mol. Cell. Biol. 6*, 462−469.
24. Kerfeld, C. A. (1997) *Photosynth. Res. 54*, 81−98.
25. Gray, J. C. (1992) *Photosynth. Res. 34*, 359−374.
26. Gray, J. C., Rochford, R. J., and Packman, L. C. (1994) *Eur. J. Biochem. 223*, 481−488.
27. Rochaix, J.-D. (1995) *Annu. Rev. Genet. 29*, 209−230.
28. Davies, J. P., and Grossman, A. R. (1998) *J. Phycol. 34*, 907−917.
29. Delosme, R. (1991) *Photosynth. Res. 29*, 45−54.
30. Kuras, R., Wollman, F. A., and Joliot, P. (1995) *Biochemistry 34*, 7468−7475.
31. Soriano, G. M., Ponamarev, M. V., Tae, G. S., and Cramer, W. A. (1996) *Biochemistry 35*, 14590−14598.
32. Zhou, J., Fernandez-Velasco, J. G., and Malkin, R. (1996) *J. Biol. Chem. 271*, 6225−6232.
33. Berry, E. A., Huang, L. S., Chi, Y., Zhang, Z., Malkin, R., and Fernandez-Velasco, J. G. (1997) *Biophys. J. 72*, A125.
34. Harris, E. H. (1989) *The Chlamydomonas Source Book. A comprehensive Guide to Biology and Laboratory Use*, Academic Press, San Diego.
35. Jancarik, J., and Kim, S.-H. (1991) *J. Appl. Crystallogr. 24*, 409−411.
36. Otwinowski, Z. (1993) in *Proceedings of the CCP4 Study Weekend: "Data Collection and Processing"* (Sawyer, L., Isaacs, N., and Bailey, S., Eds.) pp 56−62, SERC Daresbury Laboratory, Warrington, U.K.
37. CCP4: The SERC (UK) Collaborative computational project No. 4 (1995) *Acta Crystallogr. D50*, 760−763.
38. Navaza, J. (1994) *Acta Crystallogr. A50*, 157−163.
39. Brünger, A. T. (1992) *X-PLOR manual version 3.1*, Yale University Press, New Haven, CT.
40. Jones, T. A., Zhou, J.-Y., Cowan, S. W., and Kjeldgaard, M. (1991) *Acta Crystallogr. A47*, 110−119.
41. Kraulis, P. J. (1991) *J. Appl. Crystallogr. 24*, 946−950.
42. Merrit, E. A., and Murphy, M. E. P. (1994) *Acta Crystallogr. D50*, 869−873.
43. Nicholls, A., Sharp, K. A., and Honig, B. (1991) *Proteins: Struct., Funct., Genet. 11*, 281−296.
44. Sayle, R. A., and Milner-White, E. J. (1995) *Trends Biochem. Sci. 20*, 374−376.
45. Laemli, U. K. (1970) *Nature 227*, 680−685.
46. Berry, E. A., and Trumpower, B. L. (1987) *Anal. Biochem. 161*, 1−15.
47. Fernandez-Velasco, J. G., Linden, T., Hung, T., Flores, J., Huang, L.-s., and Malkin, R. (1998) *Biophys. J. 74* (2), A78.
48. Comolli, L. R., Zhou, J., Linden, T., Breitling, R., Flores, J., Hung, T., Jamshidi, A., Huang, L.-s., and Fernandez-Velasco, J. G. (1998) in *Photosynthesis: Mechanisms and Effects* (Garab, G., Ed.) pp 1589−1592, Kluwer Academic Publishers, Dordrecht, Netherlands.
49. Metzger, S. U., Cramer, W. A., and Whitmarsh, J. (1997) *Biochim. Biophys. Acta 1319*, 233−241.
50. Ubbink, M., Ejdeback, M., Karlsson, B. G., and Bendall, D. S. (1998) *Structure 6*, 323−335.
51. Pearson, D. C., Jr., Gross, E. L., and David, E. S. (1996) *Biophys. J. 71*, 64−76.
52. Soriano, G. M., Cramer, W. A., and Krishtalik, L. I. (1997) *Biophys. J. 73*, 3265−3276.
53. Ullmann, G. M., Knapp, E.-W., and Kostic, N. M. (1997) *J. Am. Chem. Soc. 119*, 42−52.
54. Gross, E. L. (1993) *Photosynth. Res. 37*, 103−116.
55. Redinbo, M. R., Yeates, T. O., and Merchant, S. (1994) *J. Bioenerg. Biomembr. 26*, 49−66.
56. Nagle, J. F., and Morowitz, H. (1978) *Proc. Natl. Acad. Sci. U.S.A. 75*, 298−302.
57. Pomes, R., and Roux, B. (1996) *Biophys. J. 71*, 19−39.
58. Baymann, F., Zito, F., Kuras, R., Minai, L., Nitschke, W., and Wollman, F.-A. (1999) *J. Biol. Chem. 274*, 22957−22967.
59. Fernandez-Velasco, J. G., Zhou, J., and Malkin, R. (1997) *Biophys. J. 72*, A126.
60. Soriano, G. M., Ponamarev, M. V., Piskorowski, R. A., Krishtalik, L. I., and Cramer, W. A. (1998) *Biophys. J. 74*, A41.
61. Edwards, K. J., Ollis, D. L., and Dixon, N. E. (1997) *J. Mol. Biol. 271*, 258−265.
62. Qin, L., and Kostic, N. M. (1993) *Biochemistry 32*, 6073−6080.
63. McLendon, G. (1991) in *Metal Ions in Biological Systems* (Sigel, H., and Sigel, A., Eds.) pp 183−198, Marcel Dekker, New York.
64. McLendon, G. (1991) *Struct. Bonding 75*, 159−174.
65. McLendon, G., and Hake, R. (1992) *Chem. Rev. 92*, 481−490.
66. Redinbo, M. R., Cascio, D., Choukair, M. K., Rice, D., Merchant, S., and Yeates, T. O. (1993) *Biochemistry 32*, 10560−10567.
67. Kerfeld, C. A., Anwar, H. P., Interrante, R., Merchant, S., and Yeates, T. O. (1995) *J. Mol. Biol. 250*, 627−647.
68. Cao, Y., Chang, V. M., Ni, B., Needleman, R., and Lanyi, J. K. (1991) *Biochemistry 30*, 10972−10979.
69. Ermler, U., Fritzsch, G., Buchanan, S. K., and Michel, H. (1994) *Structure 2*, 925−936.
70. Fritzch, G., Kampmann, L., Kapaun, G., and Michel, H. (1998) *Photosynth. Res. 55*, 127−132.
71. Ponamarev, M. V., and Cramer, W. A. (1998) *Biochemistry 37*, 17199−17208.
72. James, M. N. G., and Sielecki, A. R. (1983) *J. Mol. Biol. 163*, 299−361.
73. Oefner, C., and Suck, D. (1986) *J. Mol. Biol. 192*, 605−632.
74. Jeffrey, G. A. (1991) *Hydrogen Bonding in Biological Structure*, Springer-Verlag, Berlin, Germany.
75. Meyer, E. (1992) *Protein Sci. 1*, 1543−1562.
76. Schoenborn, B. P., Garcia, A., and Knott, R. (1995) *Prog. Biophys. Mol. Biol. 64* (2−3), 105−119.
77. Bergström, J., and Vänngard, T. (1982) *Biochim. Biophys. Acta 682*, 452−456.
78. Crowder, M. S., Prince, R. C., and Bearden, A. (1982) *FEBS Lett. 144*, 204−208.
79. Riedel, A., Rutherford, A. W., Hauska, G., Müller, A., and Nitschke, W. (1991) *J. Biol. Chem. 266*, 17838−17844.
80. Prince, R. C., and George, G. N. (1995) *Trends Biochem. Sci. 20*, 217−218.
81. Carrell, C. J., Schlarb, B. G., Bendall, D. S., Howe, C. J., Cramer, W. A., and Smith, J. L. (1999) *Biochemistry 38*, 9590−9599.

BI000090K