

Published in final edited form as:

*J Bioenerg Biomembr.* 1994 February ; 26(1): 31–47.

## Structural Aspects of the Cytochrome *b<sub>6</sub>f* Complex; Structure of the Lumen-Side Domain of Cytochrome *f*

W. A. Cramer<sup>1</sup>, S. E. Martinez<sup>1</sup>, D. Huang<sup>1</sup>, G.-S. Tae<sup>1</sup>, R. M. Everly<sup>1</sup>, J. B. Heymann<sup>1</sup>, R. H. Cheng<sup>1</sup>, T. S. Baker<sup>1</sup>, and J. L. Smith<sup>1</sup>

<sup>1</sup>Department of Biological Sciences, Purdue University, West Lafayette, Indiana 47907

### Abstract

The following findings concerning the structure of the cytochrome *b<sub>6</sub>f* complex and its component polypeptides, cyt *b<sub>6</sub>*, subunit IV and cytochrome *f* subunit are discussed:

1. Comparison of the amino acid sequences of 13 and 16 cytochrome *b<sub>6</sub>* and subunit IV polypeptides, respectively, led to (a) reconsideration of the helix lengths and probable interface regions, (b) identification of two likely surface-seeking helices in cyt *b<sub>6</sub>* and one in SU IV, and (c) documentation of a high degree of sequence invariance compared to the mitochondrial cytochrome. The extent of identity is particularly high (88% for conserved and pseudo-conserved residues) in the segments of cyt *b<sub>6</sub>* predicted to be extrinsic on the *n*-side of the membrane.
2. The intramembrane attractive forces between *trans*-membrane helices that normally stabilize the packing of integral membrane proteins are relatively weak.
3. The complex isolated in dimeric form has been visualized, along with isolated monomer, by electron microscopy. The isolated dimer is much more active than the monomer, is the major form of the complex isolated and purified from chloroplasts, and is inferred to be a functional form in the membrane.
4. The isolated cyt *b<sub>6</sub>f* complex contains one molecule of chlorophyll *a*.
5. The structure of the 252 residue lumen-side domain of cytochrome *f* isolated from turnip chloroplasts has been solved by X-ray diffraction analysis to a resolution of 2.3 Å.

### Keywords

Cytochrome *bc<sub>1</sub>*; electron transfer; energy transduction; membrane protein; structure

## INTRODUCTION AND GENERAL CONSIDERATIONS

The cytochrome<sup>2</sup> *b<sub>6</sub>f* complex is one of the three integral oligomeric membrane protein complexes involved in linear or noncyclic electron transport in the chloroplast thylakoid and

© 1994 Plenum Publishing Corporation

<sup>2</sup>Abbreviations: Cyt, cytochrome; cyt *b*(*bc<sub>1</sub>*), cytochrome *b* of the *bc<sub>1</sub>* complex; ESMS, electrospray mass spectrometry; MOA-stilbene, *E*-β-methoxyacrylate-stilbene; MW, molecular weight; NQNO, 2-*n*-nonyl-4-hydroxyquinoline-*N*-oxide; SDS-PAGE, sodium dodecyl sulfate-polyacrylamide gel electrophoresis; SU IV, subunit IV; TM, transmembrane.

cyanobacterial membrane systems that participate in oxygenic photosynthesis. It is located electrochemically and in the pathways of electron transport between the two reaction center complexes (Cramer *et al.*, 1991), and its photosynthetic electron transfer reactions can occur in the dark. It is phylogenetically related to the cytochrome  $bc_1$  complex of mitochondria and photosynthetic bacteria (Widger *et al.*, 1984), with which there are many structure-function similarities.

Reviews have recently appeared on sequence-structure-function of the mitochondrial  $bc_1$  complex (Degli Esposti *et al.*, 1993), mutational and mutagenesis studies of the bacterial  $bc_1$  complex (Gennis *et al.*, 1993), and aspects of the function of the cytochrome  $b_6f$  complex of oxygenic photosynthesis (Hope, 1993). The present article concerns the cytochrome  $b_6f$  complex, new structural information on the complex, reconsideration of the larger number of compiled sequences of cytochrome  $b_6$  and subunit IV, and a discussion of the crystal structure determination of the cytochrome  $f$  subunit. It does not consider questions of structure-function related to the Rieske iron-sulfur protein subunit of the complex, aspects of which have been recently considered for the mitochondrial protein (Graham *et al.*, 1993; Link *et al.*, 1993). The present article has the underlying viewpoint that the set of data obtained on cytochrome  $bc_1$  complexes is not sufficient to describe all important properties related to the function of the  $b_6f$  complex in oxygenic photosynthesis.

The four major ( $M_r > 15,000$ ) subunit polypeptides of the cytochrome  $b_6f$  complex that are readily detected using SDS-PAGE are cytochrome  $f$  (285 residues in spinach chloroplasts; MW = 31,372), cytochrome  $b_6$  (214 residues; MW = 24,038), the Rieske iron-sulfur protein (180 residues; MW = 18,922), and subunit IV (160 residues; MW = 17,444) (MW data for subunits from different sources compiled in Widger and Cramer, 1991). There are also 4–5 small ( $M_r < 5,000$ ) subunits in the complex, one of which has been identified as a 37-residue *pet* G subunit V (Haley and Bogorad, 1989), whose function is not known.

## SEQUENCE COMPARISONS AND STRUCTURAL INFERENCES FOR CYT $b_6$ AND SUBUNIT IV

The assumption of major subunit and prosthetic and redox group similarity of the cytochrome  $b_6f$  complex of oxygenic photosynthetic membranes to the cytochrome  $bc_1$  complex of mitochondria and purple photosynthetic bacteria (Hauska *et al.*, 1983) was extended by the demonstration of significant amino acid sequence identity and hydrophobic segment alignment of the cyt  $b$  polypeptides (Widger *et al.*, 1984). Hydropathy graphs derived from the amino sequence of the long hydrophobic putative membrane-spanning segments in the N-terminal half of the approximately 400-residue cyt  $b(bc_1)$  could be aligned by cross-correlation analysis (Shiver *et al.*, 1989) with a similar graph derived from the 214-residue spinach chloroplast cyt  $b_6$  sequence to (i) identify the four histidine residues apparently conserved and involved in heme coordination, and (ii) infer that cytochrome  $b_6$  is analogous to the hemebinding domain of cytochrome  $b$  of the  $bc_1$  complex, and subunit IV to its C-terminal half (Widger *et al.*, 1984). This also implied that the information in the single cyt  $b$  gene product of the  $bc_1$  complex is contained in the split gene products, cyt  $b_6$  and subunit IV, of oxygenic photosynthesis. These genes are contiguous or close in higher

plant chloroplasts and cyanobacteria, but separated by a large distance in the green alga *Chlamydomonas* (Büschlen *et al.*, 1991).

The involvement of the same four histidine residues in ligation of the two hemes was also inferred from their conservation in six mitochondrial sequences (Saraste, 1984). The original models for folding of the cytochrome *b* heme binding domain across the membrane bilayer contained five *trans*-membrane helices, with heme coordination involving the second and fifth helices from the N-terminus (Widger *et al.*, 1984; Saraste, 1984). It was subsequently proposed that the fourth helix in this model was not *trans*-membrane, but rather a surface-bound amphiphilic helix (Crofts *et al.*, 1987). Support for the revised four-helix model of the cytochrome *b* heme-binding domain was provided for cyt *b*(*bc*<sub>1</sub>) by: (i) the consistent mapping pattern of cytochrome *b* mutants resistant to inhibitors known to act primarily at the quinone binding site on the *n*- (e.g., antimycin, funiculosin) and *p*- (e.g., stigmatellin, myxothiazol) sides (cf. discussion of notation in legend to Fig. 1) of the membrane (diRago and Colson, 1988; Howell and Gilbert, 1988; Daldal *et al.*, 1989; Gennis *et al.*, 1993); (ii) the location of both polypeptide termini of cyt *b*<sub>6</sub> on the *n*-side of the membrane using protease accessibility of epitopes for peptide-directed antibodies (Szczepaniak and Cramer, 1990).

The original calculations of the approximate location and identity of the hydrophobic *trans*-membrane helices (Widger *et al.*, 1984; Saraste, 1984; Crofts *et al.*, 1987) could not predict the orientation of the cyt *b* polypeptide in the membrane. For cyt *b*(*bc*<sub>1</sub>), the orientation was subsequently established through (i) the distribution of sites of resistance for the *n*- and *p*-side inhibitors mentioned above, and (ii) for cyt *b* of the complex in the bacterium *Rb. sphaeroides*, the sites of fusion to alkaline phosphatase (Yun *et al.*, 1991). (iii) For cyt *b*<sub>6</sub>, the orientation shown in Fig. 2A was determined by protease accessibility to specific epitopes (Szczepaniak and Cramer, 1990), also mentioned above, and is consistent with the statistical *cis*-positive rule of Von Heijne (1992) [cf. Gavel *et al.* (1991) for a specific application to thylakoid membrane proteins]. This rule states that for intrinsic membrane proteins the number of positively charged residues located on the side of the membrane (*cis*) from which the protein is imported is larger than the number on the opposite (*trans*) side (i.e., the translocation of positively charged residues across biological membranes is energetically costly).

The number of (Arg + Lys) residues on the *n*-(*cis*) and *p*-(*trans*) sides of cyt *b*<sub>6</sub> is 8 and 5 (not including the intramembrane Arg-86), and for SU IV it is 8 and 3. This identity of the basic residues participating in this asymmetric distribution is essentially conserved in both cyt *b*<sub>6</sub> and SU IV. All of the basic residues are conserved in the 13 cyt *b*<sub>6</sub> sequences shown in Table I with the exception of the substitution of Asn (N) for Lys (K) at position 111 in *Proto-chorothrix hollandica*. In the case of SU IV, the identity of 8 of the 11 residues contributing to the (+) charge bias is conserved (Table IB), *P. hollandica* has the two changes R-15 → L and K-20 → Q, the cyanobacterium *A. quadruplicatum* is also changed at the latter position, and 4 of 14 sequences are changed at position 367. [Note that the *cis*-positive rule does not apply to cytochrome *f* and the Rieske iron-sulfur protein of the cyt *b*<sub>6</sub>*f* complex because in each of these subunits the peripheral segment of the protein exceeds 60 residues (Von Heijne, 1992); in the case of the Rieske protein, there is also a question as to

whether it has a membrane span or is totally extrinsic (Gonzalez-Halphen *et al.*, 1988; Breyton *et al.*, 1994).]

The four-helix model applied to cytochrome  $b_6$  is shown in Fig. 1A along with the folding model for SU IV. The helices of cyt  $b_6$  (Fig. 1A; in the rectangular boxes) and subunit IV (Fig. 1B) are designated A–D and E–F, respectively, and the *n*- and *p*-side extrinsic linking peptide segments “ab,” “bc,” and “cd” (segments of cyt  $b_6$ ), and “ef,” “fg” (segments of SU IV) according to the notation of Crofts *et al.* (1990). The models of Fig. 1A, B differ from earlier models of cyt  $b_6$  and SU IV (Widger and Cramer, 1991) by (i) inclusion of the tendency of the *p*-side surface segments “ab,” “cd” (cyt  $b_6$ ), and “ef” (SU IV) to form amphiphilic helices. The hydrophobic moments of helices “ab,” “cd,” and “ef” are 0.39, 0.53, and 0.45, respectively, using a segment length of 15 residues and the amino acid hydrophobicity data base of Eisenberg (1984), as in Shiver *et al.* (1989). This amphiphilic character has been noted previously for the “cd” helix of cyt  $b(bc_1)$  (Crofts *et al.*, 1987), and the latter amphiphilic helix has been explicitly included in models of the *p*-side quinone binding site (Robertson *et al.*, 1990). Asp-155 or Glu-166 in the “cd” segment have been proposed to bind DCCD and to facilitate H<sup>+</sup> translocation to the membrane surface associated with *p*-side quinol oxidation (Beattie, 1993). The “ef” segment in SU IV contains seven proline residues, three of which have been included in an “ef” helix because of its pronounced amphiphilicity. Precedents for such a helix are provided by surface helices in the photosynthetic bacterial reaction center: the “ab” helix with three or two (*Rb. capsulatus*), and the “cd” helix with three or two (*C. aurantiacus*) proline residues. (ii) The length of the F, G helices of SU IV has been extended one turn in order to take into account the tendency of basic residues Lys-333 and R-339, R-340 to provide punctuation for TM helices. (iii) The tendency of aromatic residues, and tryptophans in particular, to be found near the aqueous interface (Deisenhofer and Michel, 1989; Jacobs and White, 1989) was taken into account by placing Trp-79 on the *p*-side of helix B, and Trp-246 on the *n*-side of helix E in contact with the bilayer phase.

The orientation of the “fg” helix of subunit IV inferred by Li *et al.* (1991) from trypsin accessibility studies was opposite to that shown in Fig. 1B. The distribution of (Arg+Lys) in a SU IV with an orientation reversed from that shown in Fig. 1B would be decidedly opposite to that predicted by the *cis*-positive rule.

Two striking features of the comparison of sequences of the 13 cyt  $b_6$  and 16 SU IV polypeptides are (a) the high degree of overall sequence identity, as noted in a comparison of an initial set of three complete, and two partial  $b_6$ -SU IV sequences (Hauska *et al.*, 1988), and (b) the higher extent of identity on the *n*-side compared to the *p*-side of the membrane. The degree of conservation of the 13 sequences of the 214 cyt  $b_6$  polypeptide (Table IA) is 69% and 79% for identical and pseudo-identical (K↔R, D↔E, and S↔T) residues, and is 53% and 63%, respectively, for the 160 residue subunit IV (Table IB). These levels of identity appear to far exceed those seen in cyt  $b$  of the mitochondrial  $bc_1$  complex (Hauska *et al.*, 1988; Degli Esposti *et al.*, 1993). In the 18 mitochondrial sequences compiled by Hauska *et al.* (1988), 10% of the residues are invariant (Degli Esposti *et al.*, 1993). Degli Esposti *et al.* (1993) have also noted that the result of sequence analysis of the cytochrome  $b$

protein from more than 800 species, the most extensive analysis for any integral membrane protein, is that there are only 9 invariant amino acids (Table II).

Table II summarizes (a) the position in the sequence of cyt  $b_6$ -SU IV, and (b) the likely or proposed function of the 24 most highly conserved residues in the mitochondrial cytochrome  $b$  (Degli Esposti *et al.*, 1993). Sixteen of the 24 residues are in cyt  $b_6$ , according to Degli Esposti *et al.* (1993). Eight are in SU IV, of which three are the consecutive Pro 291-Glu 292-Trp 293. Five of the residues listed as most highly conserved by Degli Esposti *et al.* (1993) are different although conserved in cyt  $b_6$ , where they are Val-132 ( $p$ -side, C helix), Gly-143 ( $p$ -side, “cd” loop), Lys-148 ( $p$ -side, “cd” helix), Ile-149 ( $p$ -side, “cd” helix), and Arg-206 ( $n$ -side, C-terminal segment). No such residues occur in SU IV. Thus, four of the 16 most conserved residues in cyt  $b_6$ , and four of the five that break the invariance of the superfamily, are in the “cd” segment of cyt  $b_6$ . The fifth residue that breaks the invariance, Arg-206, is analogous to His-202 that is otherwise completely conserved in the mitochondrial and photosynthetic bacterial cytochrome. In the latter, it is His-217. The mutation His 217  $\rightarrow$  Arg in the bacteria results in an inhibition of the rate of heme  $b_n$  oxidation by a factor of 3–4 (Gray and Daldal, 1993; Hacker *et al.*, 1993). The fact that the Arg residue that results in inhibition of bacterial heme  $b_n$  (cf. discussion of notation in legend of Fig. 1) oxidation is conserved in cyt  $b_6$  can be conservatively interpreted as indicating that the putative quinone binding site on the  $n$ -side of the membrane ( $Q_n$  site) has a different structure compared to that in the mitochondria and bacteria. Arg 206-Lys 207 have also been shown to be accessible to trypsin in thylakoid membranes (Szczepaniak and Cramer, 1990). This implies that the heme coordinated by His-201 is close (ca. 7 – 9 Å) to the  $n$ -side interface, in contrast to heme  $b_n$  of the mitochondrial cytochrome that was inferred from EPR experiments with external paramagnetic complexes (Ohnishi *et al.*, 1989) and electrochromic msec “slow” phase measurements (Glaser and Crofts, 1984; Robertson and Dutton, 1988) to be appreciably removed (approximately 25 Å) from the  $n$ -side interface. One other Arg residue that is uniquely conserved in the cyt  $b_6$  sequences is Arg-86 adjacent to His-85 in the B helix. It is the only conserved residue that carries a nominal charge in the designated *trans*-membrane helices of the cytochrome  $b$  polypeptides.

The difference in sequence and probable structure of the  $n$ -side of cyt  $b_6$  compared to cyt  $b(bc_1)$  is further illustrated by the extremely high level of sequence identity of the  $n$ -side peripheral segments of cyt  $b_6$ , 78% and 87% for invariance and pseudo-invariance (Table IA, Fig. 1). These levels are higher than the corresponding values of 61% and 73% for the  $p$ -side domains. Thirteen of the last fifteen residues in cyt  $b_6$ , Met-200 to Leu-214 on the  $n$ -side of helix D, are invariant, with Arg-206 approximately in the middle of this segment. The greater invariance on the  $n$ -side of the membrane is a bias opposite to that in cytochrome  $b$  of the  $bc_1$  complex (Degli Esposti *et al.*, 1993). The  $n$ -side of the  $b_6f$  complex is known to be characterized by the lack of a high-affinity binding site for quinone-analogue inhibitors (Widger and Cramer, 1991; Hope, 1993). The lack of such a site for a potent  $n$ -side inhibitor such as antimycin A is one of the obvious differences between the  $b_6f$  and  $bc_1$  complexes, and has been a major point in discussions about the possibility of different properties of the  $b_6f$  complex compared to cyt  $b(bc_1)$  (Cramer *et al.*, 1991; Rich *et al.*, 1992). [The inhibitors NQNO (Selak and Whitmarsh, 1982; Rich *et al.*, 1991; Kramer and

Crofts, 1992) and MOA-stilbene (Rich *et al.*, 1992) may be weakly binding *n*-side inhibitors of cyt  $b_6$ ; see Hope (1993) for a discussion of some inconsistencies in this interpretation regarding NQNO.] However, when contrasted with mitochondria, the higher extent of sequence identity of the *n*-side of cyt  $b_6$  and SU IV implies that (i) a much higher degree of sequence invariance is required to establish a binding site of plastoquinone to allow the same Q-cycle mechanism that has been documented for mitochondria and bacterial chromatophores (Crofts, 1985; Trumpower, 1990); (ii) alternatively, the *n*-side extrinsic regions of cyt  $b_6$  and SU IV are responsible for essential functions in addition to, or instead of, the Q cycle function. One such function might be the docking of redox proteins [e.g., the “G” protein component of Joliot and Joliot (1988)] that can supply reducing equivalents to heme  $b_n$  from the *n*-side aqueous phase.

## WEAKNESS OF HELIX–HELIX INTERACTIONS

The  $b_6f$  complex is extracted from the membrane as an integral membrane protein complex in detergents such as cholate, MEGA-9, or  $\beta$ -D-dodecylmaltoside. As a membrane protein complex, the  $b_6f$  complex has the unusual property that the subunits can be separated from the complex and from each other at alkaline pH (ca. 10.5–11.5) [Fig. 2; Szczepaniak *et al.*, 1991; Cramer *et al.*, 1992]. This implies that, unlike membrane proteins such as bacteriorhodopsin (Popot and Engelman, 1990) and the glycoporphin dimer (Lemmon *et al.*, 1992), the attractive forces arising from intramembrane hydrophobic, hydrogen bond, and van der Waals interactions between the subunits are relatively weak. As a consequence, the electrostatic repulsive force exerted between the subunits at alkaline pH, arising from net excess negative charge on the peripheral polypeptide segments, is sufficient to cause the subunits of the complex to separate in the membrane. The fact that the subunits of the spinach chloroplast  $b_6f$  complex do undergo lateral separation in the membrane at high pH is inferred from the separate and sequential release of the subunits of the complex as a function of increasing pH. A satisfying explanation does not exist as to why the subunits are not just separated, but also extruded from the membrane. One may speculate that the release from the membrane is a consequence of the partial deprotonation and neutralization of the basic residues that are proposed to form the defined “stops” for the predicted *trans*-membrane  $\alpha$ -helices A, B, and D of cyt  $b_6$  (Fig. 1A) and E, F, and G of SU IV (Fig. 1B). Once the “stops” are lost, the positions of the helices in the membrane would have fewer thermodynamic constraints, and more freedom to move in the direction normal to the plane of the membrane.

The relative weakness of the attractive intramembrane forces between the subunits of the  $b_6f$  complex may be partly a consequence of prosthetic groups such as the helix-bridging heroes (Fig. 1C) providing spacers between these helices that prevent close packing between some of the *trans*-membrane helices.

## THE $b_6f$ COMPLEX AS A STRUCTURAL AND FUNCTIONAL DIMER

The relatively weak intramembrane forces between subunits could also facilitate interconversion of monomer and dimer forms of the complex. The existence of a dimer form of isolated cyt  $b_6f$  reconstituted *in vitro* was indicated from its geometric cross-section in

electron micrographs (Mörschel and Staehelin, 1983). The presence of both monomer and dimer forms in isolated cyt *b<sub>6</sub>f* complex was inferred from (a) the presence in a sucrose density gradient of two fractions of unknown, but different, relative molecular weight, (b) the conversion of the higher to the lower  $M_r$  form after longer centrifugation times or in higher Triton detergent concentrations, (c) inhibition of this conversion by cross-linking reagents, and (d) the generation of different cross-linked  $M_r$  forms from the high- and low- $M_r$  fractions (Chain and Malkin, 1991).  $M_r$  values and distinct sizes and geometries corresponding to monomer and dimer forms of cyt *b<sub>6</sub>f* complex isolated from spinach thylakoids were determined by (a) FPLC molecular chromatography in  $\beta$ -D-dodecylmaltoside, (b) native gel electrophoresis according to the method of Schagger and von Jagow (1991), and (c) averaged images obtained by electron microscopy of negatively stained single particles (Huang *et al.*, 1993). The dimer was shown to be much more abundant (80–90% of the total complex) than the monomer when assayed by native gel electrophoresis (Huang *et al.*, 1993). When isolated by FPLC molecular sieve chromatography which, compared to the native gel technique, resulted in a higher fraction (30% of the total) of monomer, the dimer was also much more (~ 5-fold) active in electron transfer from plastoquinol-2 to plastocyanin. The level of activity of the monomer was low enough that it could be attributed to residual dimer (Huang *et al.*, 1993). Single-particle images of the monomer and dimer are shown in Fig. 3. A discussion of structure determination by electron microscopy is presented elsewhere in this volume by Boekema *et al.*

The existence of the dimeric form as the predominant form of the purified mitochondrial cytochrome *bc<sub>1</sub>* complex is well established (Weiss and Kolb, 1979; Von Jagow and Sebald, 1980; Leonard *et al.*, 1981; Nalecz and Azzi, 1985), although the functional significance of this dimer is still unclear (de Vries *et al.*, 1983; Schmitt and Trumpower, 1990; Bechmann *et al.*, 1992; Nieboer and Berden, 1992). It is known that oligomeric structures of membrane proteins are statistically likely in the two-dimensional space of the membrane (Grasberger *et al.*, 1986), and that a dimeric structure and dimer–monomer transitions of the respective receptors are important in the mechanism of *trans*-membrane signaling in at least several systems (Pakula and Simon, 1992; Bormann and Engelman, 1992). Two of the more thoroughly studied examples include the bacterial aspartate chemoreceptor (Milburn *et al.*, 1991), and the epidermal growth factor receptor-tyrosine kinase (Canals, 1992) signaling systems. We make the admittedly speculative suggestion that the cytochrome *b<sub>6</sub>f* and *bc<sub>1</sub>* complexes are primitive *trans*-membrane signaling systems. In the case of the *b<sub>6</sub>f* complex, this is suggested by its apparent association with a protein kinase, the LHCP kinase, on the *n*-side of the membrane, and the ability of *p*-side quinone analogue inhibitors to block kinase activity (Gal *et al.*, 1990a, b).

## MOLECULE OF CHLOROPHYLL *a* IN *b<sub>6</sub>f* COMPLEX

The presence of chlorophyll in the spectrum of the cytochrome *b<sub>6</sub>f* complex isolated from spinach chloroplasts was noted in a benchmark paper on the purification and properties of the complex (Hurt and Hauska, 1983). The cyt *b<sub>6</sub>f* complex from the cyanobacterium *Synechocystis* PCC 6803, which was isolated in the monomeric form, was found to contain a single chlorophyll *a* molecule that was determined by linear dichroism analysis to be

oriented in the complex (Bald *et al.*, 1992). Electron microscopic analysis of this complex was carried out by Bald *et al.* (1992) and is reviewed in the present volume of this journal by Boekema *et al.* Analysis of the isolated  $b_6f$  complex from spinach chloroplasts showed a chlorophyll *a*:cytochrome *f* stoichiometry  $> 1$ , but a smaller value close to 1 when the dimer was further purified from the complex by FPLC molecular sieve chromatography (Huang *et al.*, 1993). The latter stoichiometry appeared to be the minimum value because it was not decreased further by an additional chromatographic step, implying that the one chlorophyll *a* molecule was not bound adventitiously.

## CYTOCHROME *f* STRUCTURE: BIOCHEMICAL AND CELL BIOLOGICAL CONSEQUENCES

The understanding of structure/function of cytochrome *f* was advanced recently by the determination of its X-ray structure at a resolution of 2.3 Å (Martinez *et al.*, 1991, 1992, 1993). The structure that was solved was that of the soluble 252-residue lumen-side domain of the 285-residue (MW = 31, 298 in turnip chloroplasts) mature polypeptide that spans the membrane once (Fig. 4C).

The use of the lumen-side polypeptide of the cytochrome arose from attempts to prepare intact cytochrome *f* from spinach chloroplast thylakoid membranes for crystallization. These attempts were abandoned because the cytochrome tended to aggregate. This had been noted previously by Ho and Krogmann (1980) who determined a molecular weight of 285,000 for the aggregate that approximately corresponded to an octamer. Preparations of cytochrome *f* from cruciferous plants, however, were reported to be soluble and monomeric with an  $M_r$  value of 27,000 instead of 32,000–33,000 (Gray, 1978). It was suggested that the decrease in apparent molecular size might arise from proteolysis at the C-terminus because the N-terminal sequence of cytochrome *f* from the cruciferous plant, rape, was found to be identical to that of spinach and pea (Willey *et al.*, 1984). It was also inferred that the tendency of the spinach cytochrome *f* to aggregate in aqueous solution might arise from the 20-residue hydrophobic segment (residues Nos. 251–270, underlined in Table III) near the C-terminus, and that removal of the hydrophobic domain near the C-terminus might explain the water-soluble monomeric nature of the cruciferous cytochrome *f* fragment. Therefore, it was decided to purify cruciferous cytochrome *f* for structural analysis.

Cytochrome *f* was purified from turnip leaves and found to display an  $M_r$  value of 31,000 on SDS-PAGE, compared to 33,800 in the turnip thylakoid membrane (Fig. 5). After promising crystals had been obtained, truncation of the turnip cytochrome *f* near the C-terminus during purification was confirmed by C-terminal sequencing of the turnip cytochrome polypeptide with carboxypeptidases Y and A. Three of the last four residues were found to be Leu 249-Arg 250-Val 251 (Martinez *et al.*, 1991). The fourth C-terminal residue was inferred to be a serine at position 252, instead of the glutamine that was conserved in the several chloroplast cytochrome *f* sequences known at that time. It was therefore thought that the explanation of the susceptibility of cytochrome *f* to lumen-side cleavage by an endogenous protease was a Gin → Set substitution at residue 252 (Martinez *et al.*, 1991). Determination of the nucleotide sequence of the turnip cytochrome *f* gene by J. C. Gray showed that residue 252 is, in fact, a glutamine in the turnip (*Brassica campestris*) cytochrome *f* (Gray, 1992).



Therefore, the explanation of the “natural” proteolysis in the *Cruciferae* cytochrome *f* is the existence of a protease in chloroplasts from these plants that is activated or made accessible to the lumen-side of the cytochrome *f* during its extraction from the thylakoid membranes. The same cleavage was subsequently reported for cytochrome *f* from charlock (Gray, 1992). It is perhaps interesting that the cleavage predominantly occurs two residues into the 20-residue, hydrophobic *trans*-membrane  $\alpha$ -helix. This suggests that the protease might be related to a processing or leader peptidase. The thylakoid processing peptidase of cytochrome *f* is known to be membrane-bound (Johnson *et al.*, 1991), and the active site of that for plastocyanin is on the lumen side of the membrane (Kirwin *et al.*, 1991). On the other hand, the cleavage site of a processing protease should be unique, whereas analysis of the charlock (Gray, 1992) cytochrome *f* polypeptide by ESMS showed that the C-terminus is somewhat ragged, with small populations of fragments resulting from cleavage at residues adjacent to Glu-252. Preliminary ESMS of the turnip cyt *f* fragment indicated that it is 90% of the fragment population (D. Huang, D.L. Smith, and W.A. Cramer, unpublished data). The residue closest to the C-terminus for which electron density can be resolved in the structure is Leu-249 adjacent to Arg-250 (Fig. 6).

The successful crystallization of the soluble turnip cyt *f* polypeptide that contains a small part (Val 251-Gln 252) of the putative hydrophobic membrane-spanning helical domain suggests a strategy for the systematic genetic engineering for purposes of crystallization of redox proteins such as cytochrome *c*<sub>1</sub> or the Rieske iron-sulfur protein that have the same kind of general topography: generate a polypeptide fragment of the peripheral soluble fragment that is as long as possible to allow proper folding of the protein. This fragment could include a few residues of the hydrophobic helical domain, in order to increase the chances for correct folding of the fragment without affecting solubility.

The turnip cytochrome *f* polypeptide was crystallized at 5°C in the reduced form in the presence of dithioerythritol using 40–42% acetone as the precipitant (Martinez *et al.*, 1992). The structure has been solved by multiple isomorphous replacement and anomalous scattering to a resolution of 2.3 Å (Martinez *et al.*, 1993). Some of the major features of the structure are:

1. Unlike the prototypical structure of soluble cytochrome *c* that has one mostly  $\alpha$ -helical domain, the elongate ( $25 \times 35 \times 75$  Å) cytochrome *f* structure is made of two domains whose major secondary structure motif is a  $\beta$ -sheet.
2. The heme is near the interdomain interface in the larger domain, the Fe 45 Å from the C-terminus of Arg-250 that is connected to the TM  $\alpha$ -helix.
3. The smaller domain, consisting of residues 169–231, contains Lys-187, previously shown to cross-link to Asp-44 of plastocyanin (Morand *et al.*, 1989; cf. Redinbo *et al.*, this volume), that is solvent-exposed and 28 Å from the heme Fe (Fig. 6). Lys-187 is in a positively charged region that includes Lys-185 and Arg-209 on the small domain. Lys-66, Lys-65, and Lys-58 are neighboring basic residues on the large domain.
4. The axial sixth heme ligand is the  $\alpha$ -amino group of the N-terminal tyrosine residue. A number of spectroscopic studies had noted similarities between spectra

of cytochrome *f* and soluble mammalian cytochrome *c* at alkaline pH (> 10) in which the methionine ligand is exchanged for a lysine (Siedow *et al.*, 1980; Davis *et al.*, 1988; Rigby *et al.*, 1988; Simpkin *et al.*, 1989). The conserved Lys-145 was proposed as the axial sixth ligand (Davis *et al.*, 1988). However, the  $\epsilon$ -amino group of Lys-145 is 33 Å from the heme iron. The amino group function inferred from the spectroscopic studies can be fulfilled by the tyrosine  $\alpha$ -amino group, and in one respect is fulfilled even more readily because the pK of the latter is close to neutrality.

5. Unlike mammalian cytochrome *c*, the heme is insulated from the aqueous phase by a shield including three aromatic residues of which one is Tyr-1. This may explain why the midpoint redox potential of cyt *f* is approximately 100 mV more positive than that of mammalian cyt *c*.

Because the structure of plastocyanin is known (Gross, 1993; Redinbo *et al.*, this volume), the question immediately arises as to its competent binding configuration with cytochrome *f*. Information on the sites of interaction was provided by covalent cross-linking of plastocyanin to cytochrome *f* using the water-soluble 1-ethyl-3-[3-(dimethyl-amino)propyl]carbodiimide (EDC). One site of linkage involved Asp-44 of plastocyanin and Lys-187 of cyt *f*. A second linkage involved Glu-59 and/or 60 to an unknown, presumably basic, residue on cyt *f*. As noted above, Lys-187 is in a basic “patch” of residues (Fig. 6). The transfer of electrons from cytochrome *f* to plastocyanin is thought to be mediated as well by the Tyr-83 residue of plastocyanin which is exposed on the surface of PC and whose ring is approximately 9–10 Å and 10 Å from Asp-44 and the Cu atom, respectively. The distance of Asp-44 from the copper is 20 Å. With cyt *f* in the orientation shown in Fig. 6, PC would bind through its acidic regions (residues 42–45 and 59–61 of spinach plastocyanin) on the side of the barrel to the basic region of cyt *f* around Lys-187, the copper would be at the bottom of the PC barrel in this binding geometry, and Tyr-83 would then be located closer to the heme than Asp-44. It should be noted that covalent cross-linking of the PC–cyt *f* pair renders it incompetent for intermolecular electron transfer (Qin and Kostic, 1993), and the PC inactive in the reduction of photosystem I (Morand *et al.*, 1989).

### Orientation of Cytochrome *f* Relative to the Membrane Surface

It would seem that interaction of cytochrome *f* with the Rieske protein would be facilitated if the long axis of the large domain of cyt *f* was oriented with not too large an angle relative to the plane of the membrane. Interaction between the proposed surface “cd” helix of cytochrome *b* and the analogous cytochrome *c*<sub>1</sub> is implied by the existence of a second site suppressor mutation in cyt *c*<sub>1</sub> to a primary inhibitor-resistant mutation in the “cd” loop of cyt *b*(*bc*<sub>1</sub>) of *Rb. capsulatus* (F. Daldal, personal communication). It has been observed in EPR studies on magnetically oriented chloroplast membranes (Bergstrom and Vanngard, 1982; Crowder *et al.*, 1982) and dehydrated two-dimensional lattices of cytochrome *b*<sub>6</sub>*f* complex oriented after centrifugation on Mylar sheets (Riedel *et al.*, 1991) that the cyt *f* heme has a broad (Bergstrom and Vanngard, 1982; Riedel *et al.*, 1991) or narrow (Crowder *et al.*, 1982) distribution of orientations relative to the membrane surface. The most probable orientation of the heme plane in both kinds of distributions would have it sharply tilted at an angle of approximately 25–30° relative to the plane of the membrane. Although there may be some

flexibility in the orientation of cyt *f*, the most probable orientation of the protein appears to be favorable for interaction with surface-bound or extrinsic segments of the Rieske iron-sulfur protein and cyt *b*<sub>6</sub>.

### Sixth Heme Ligand of Cytochrome *f*: Consequences for Protein Translocation and Assembly

The consequence of the sixth heme ligand being the  $\alpha$ -amino group of the N-terminal Tyr residue of the mature protein is that the complete ligation of the heme that results in formation of the low-spin coordination state, and assembly of the final folded state of the protein, cannot occur until the signal peptide of the intermediate translocated preprotein has been cleaved by the processing peptidase. This peptidase is believed to be located on the lumen side of the membrane (Johnson *et al.*, 1991; Kirwin *et al.*, 1991). This implies that the protein folding could not be completed until translocation has proceeded at least to the point where a significant part of the cytochrome including Tyr-1 and His-25 has crossed the membrane. In addition, the ligation of the heme and associated folding, as in the translocation pathway of apocytochrome *c* (Planner and Neupert, 1990), may provide additional free energy needed to complete the translocation process.

The use of the Tyr-1  $\alpha$ -amino group as the sixth heme ligand can be viewed as a control mechanism to ensure a proper delay in the timing of protein refolding, and is an example of the general principle of membrane protein translocation that the protein to be translocated be unfolded (Verner and Schatz, 1988; Pfanner and Neupert, 1990). Although the use of amino terminal ligation thus seems to be a rational mechanism, it appears that it is not necessarily ubiquitous for metalloproteins even for the closely related cytochrome *c*<sub>1</sub> protein: (i) in the nitrogen-fixing microaerobic endosymbiont, *Bradyrhizobium japonicum*, cytochromes *b* and *c*<sub>1</sub> can be made by mutagenesis as a single two-domain polyprotein connected by a 31-residue internal signal peptide (Thony-Meyer *et al.*, 1991). Although it was shown in the latter work that covalent binding of the *c*<sub>1</sub> heme was required for formation of the *bc*<sub>1</sub> polyprotein, cleavage of the internal signal peptide was not. The use of a sixth ligand different from the N-terminal  $\alpha$ -amino group, for at least some cytochromes *c*<sub>1</sub>, is also suggested by the appearance of a normal reduced  $\alpha$ -band spectrum in a mutant of yeast cytochrome *c*<sub>1</sub> that is blocked in cleavage of the signal peptide (Yang and Trumpower, 1993).

### Acknowledgments

This research was supported by grants NIH GM-18457 (WAC), USDA 9101624 (JLS and WAC), and NSF DMB-8905062 (TSB). We thank P.N. Furbacher for helpful discussions, B. L. Trumpower for communication of a manuscript before publication, and J. Hollister and V. Livingston for careful work in preparing the manuscript.

### References

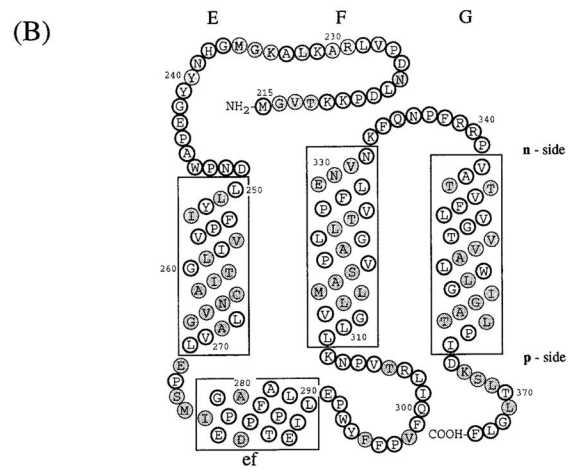
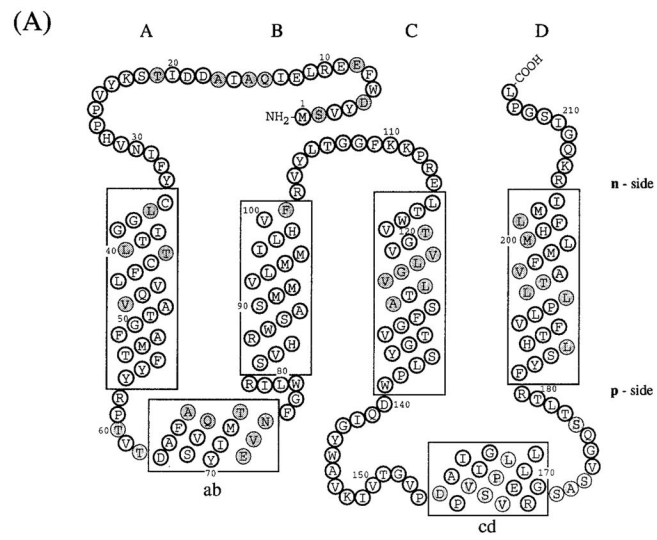
- Bald, D.; Kruip, J.; Boekema, E.J.; Rögner, M. Research in Photosynthesis. Murata, N., editor. Vol. I. Kluwer; Dordrecht: 1992. p. 629-632.
- Beattie DS. J Bioenerg Biomembr. 1993; 25:233–244. [PubMed: 8394319]
- Bechmann G, Weiss H, Rich PR. Eur J Biochem. 1992; 208:315–325. [PubMed: 1325904]
- Bergstrom J, Vanngard T. Biochim Biophys Acta. 1982; 682:452–456.

- Bormann BJ, Engelman DE. *Annu Rev Biophys Biophys Chem.* 1992; 19:369–403.
- Brand SN, Tan X, Widger WR. *Plant Mol Biol.* 1992; 20:481–491. [PubMed: 1421151]
- Breyton C, de Vitry C, Popot J-L. *J Biol Chem.* 1994 in press.
- Büschlen S, Choquet Y, Kuras R, Wollman FA. *FEBS Lett.* 1991; 284:257–262. [PubMed: 2060646]
- Canals F. *Biochemistry.* 1992; 31:4493–4501. [PubMed: 1316148]
- Chain R, Malkin R. *Photosynth Res.* 1991; 28:59–68. [PubMed: 24414859]
- Cote JC, Wu NH, Wu R. *Plant Mol Biol.* 1988; 11:873–874. [PubMed: 24272638]
- Cramer, WA.; Black, MT.; Widger, WR.; Girvin, ME. *The Light Reactions.* Barber, J., editor. Elsevier; Amsterdam: 1987. p. 446-493.
- Cramer, WA.; Furbacher, PN.; Szczepaniak, A.; Tae, G-S. *Current Topics in Bioenergetics.* Lee, CP., editor. Vol. 16. Academic Press; Orlando: 1991. p. 179-222.
- Cramer WA, Engelman DM, von Heijne G, Rees DC. *FASEB J.* 1992; 6:3397–3402. [PubMed: 1464373]
- Crofts, AR. *The Enzymes of Biological Membranes.* 2. Martinosi, A., editor. Vol. 4. Plenum Press; New York: 1985. p. 347-382.
- Crofts, AR.; Robinson, H.; Andrews, K.; Van Doren, S.; Berry, E. *Cytochrome Systems: Molecular Biology and Bioenergetics.* Papa, S.; Chance, B.; Ernster, L., editors. Plenum Press; New York: 1987. p. 617-624.
- Crofts, AR.; Wang, Z.; Chen, Y.; Mahalingam, S.; Yun, CH.; Gennis, RB. *Highlights in Ubiquinone Research.* Lenaz, G., editor. Taylor and Francis; London: 1990. p. 98-103.
- Crowder MS, Prince RC, Bearden A. *FEBS Lett.* 1982; 144:204–208.
- Daldal F, Tokito MK, Davidson E, Faham M. *EMBO J.* 1989; 8:3951–3961. [PubMed: 2556259]
- Davis DJ, Frame MK, Johnson DA. *Biochim Biophys Acta.* 1988; 936:61–66. [PubMed: 2846050]
- Degli Esposti M, DeVries S, Crimi M, Ghelli A, Patarnello T, Meyer A. *Biochim Biophys Acta.* 1993; 1143:243–271. [PubMed: 8329437]
- Deisenhofer J, Michel H. *EMBO J.* 1989; 8:2149–2169. [PubMed: 2676514]
- de Vries S, Albracht SPJ, Berden JA, Marres CAM, Slater EC. *Biochim Biophys Acta.* 1983; 723:91–103. [PubMed: 6299337]
- diRago JP, Colson AM. *J Biol Chem.* 1988; 263:12564–12570. [PubMed: 2842335]
- Eisenberg D. *Annu Rev Biochem.* 1984; 53:595–623. [PubMed: 6383201]
- Fukuzawa H, Yoshida T, Kohchi T, Okumura T, Sawano Y, Ohyama K. *FEBS Lett.* 1987; 220:61–66.
- Furbacher PN, Girvin ME, Cramer WA. *Biochemistry.* 1989; 28:8990–8998. [PubMed: 2605237]
- Gal, A.; Mets, LJ.; Ohad, I. *Current Research in Photosynthesis.* Baltscheffsky, M., editor. Vol. II. Kluwer Academic; Dordrecht: 1990a. p. 779-781.
- Gal, A.; Mor, TS.; Hauska, G.; Herrmann, R.; Ohad, I. *Current Research in Photosynthesis.* Baltscheffsky, M., editor. Vol. II. Kluwer Academic; Dordrecht: 1990b. p. 783-785.
- Gavel Y, Steppuhn J, Herrmann R, von Heijne G. *FEBS Lett.* 1991; 282:41–46. [PubMed: 2026263]
- Gennis RB, Barquera B, Hacker B, Van Doren SR, Arnaud S, Crofts AR, Davidson E, Gray KR, Daldal F. *J Bioenerg Biomembr.* 1993; 25:195–209. [PubMed: 8394316]
- Girvin ME, Cramer WA. *Biochim Biophys Acta.* 1984; 767:29–38. [PubMed: 6487614]
- Glaser EA, Crofts AR. *Biochim Biophys Acta.* 1984; 766:322–333. [PubMed: 6087897]
- Gonzalez-Halphen D, Lindorfer MA, Capaldi RA. *J Biol Chem.* 1988; 27:7021–7031.
- Graham LA, Brandt U, Sargent JS, Trumpower BL. *J Bioenerg Biomembr.* 1993; 25:245–257. [PubMed: 8394320]
- Grasberger B, Minton AP, DeLisi C, Metzger H. *Proc Natl Acad Sci USA.* 1986; 83:6258–6262. [PubMed: 3018721]
- Gray JC. *Eur J Biochem.* 1978; 82:133–141. [PubMed: 620667]
- Gray JC. *Photosynth Res.* 1992; 34:359–374. [PubMed: 24408832]
- Gray KA, Daldal F. *Biophys J.* 1993; 64:105a.
- Greet KL, Golden SS. *Plant Mol Biol.* 1992; 19:355–365. [PubMed: 1623188]
- Gross EL. *Photosyn Res.* 1993; 37:103–116. [PubMed: 24317707]

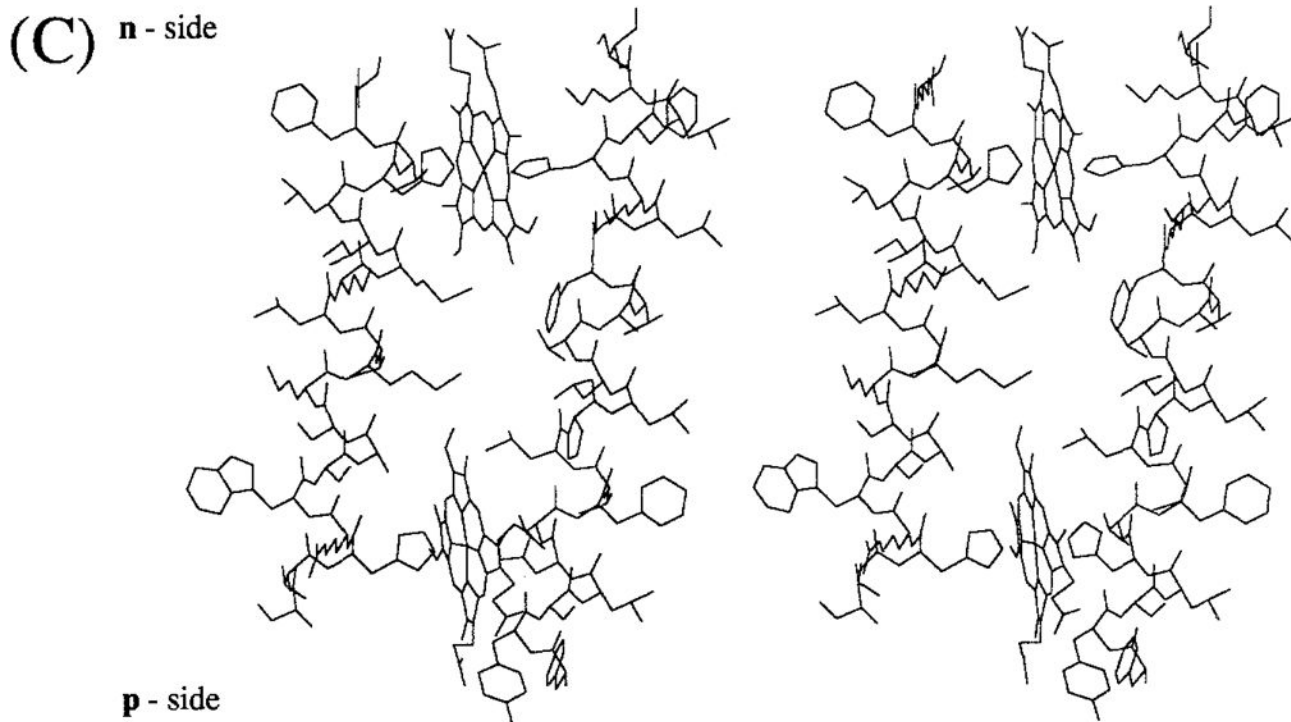
- Hacker B, Barquera B, Crofts AR, Gennis RB. *Biophys J*. 1993; 64:105a.
- Haley J, Bogorad L. *Proc Natl Acad Sci USA*. 1989; 86:1534–1538. [PubMed: 2922397]
- Hauska G, Hurt E, Gabellini N, Lockau W. *Biochim Biophys Acta*. 1983; 726:97–133. [PubMed: 6307358]
- Hauska G, Nitschke W, Herrmann RG. *J Bioenerg Biomem*. 1988; 20:211–228.
- Heinemeyer W, Alt J, Herrmann RG. *Curr Genet*. 1984; 8:543–549. [PubMed: 24177956]
- Hiratsuka J, Shimada H, Whittier R, Ishibashi T, Sakamoto M, Mori M, Kondo C, Honji Y, Sun CR, Meng BY, Li YQ, Kanno A, Nishizawa Y, Hirai A, Shinozaki K, Sugiura M. *Mol Gen Genet*. 1989; 217:185–194. [PubMed: 2770692]
- Hird SM, Webber AN, Wilson RJ, Dyer TA, Gray JC. *Plant Mol Biol*. 1991; 16:745–747. [PubMed: 1868207]
- Ho KK, Krogmann DW. *J Biol Chem*. 1980; 255:3855–3861. [PubMed: 6768742]
- Hope AB. *Biochim Biophys Acta*. 1993; 1143:1–22. [PubMed: 8388722]
- Howell N, Gilbert K. *J Mol Biol*. 1988; 203:607–618. [PubMed: 3210228]
- Huang DH, Everly M, Cheng RH, Heymann JB, Schägger H, Baker TS, Cramer WA. 1993 Submitted.
- Hurt EC, Hauska G. *FEBS Lett*. 1983; 153:413–419.
- Jacobs RE, White SH. *Biochemistry*. 1989; 8:3421–3437. [PubMed: 2742845]
- Johnson EM, Schabelrauch LS, Sears BB. *Mol Gen Genet*. 1991; 225:106–112. [PubMed: 2000083]
- Joliot P, Joliot A. *Biochim Biophys Acta*. 1988; 933:319–333.
- Kallas T, Spiller S, Malkin R. *J Biol Chem*. 1988; 263:14334–14342. [PubMed: 2844767]
- Kirwin PM, Elderfield PD, Williams RS, Robinson C. *J Biol Chem*. 1991; 263:18128–18132. [PubMed: 3056939]
- Kramer, DM.; Crofts, AR. *Curr Res Photosyn. Baltscheffksy, M., editor. Vol. III. Kluwer Academic; Dordrecht: 1990. p. 283-286.*
- Kramer, DM.; Crofts, AR. *Res Photosyn. Murata, N., editor. Vol. II. Kluwer; Dordrecht: 1992. p. 491-494.*
- Kück U. *Mol Gen Genet*. 1989; 218:257–265. [PubMed: 2476655]
- Kutoh E, Sone N. *J Biol Chem*. 1988; 263:9020–9026. [PubMed: 2837471]
- Lemmon MA, Flanagan JM, Hunt JF, Adair BD, Bormann BJ, Engelman DE. *J Biol Chem*. 1992; 267:7683–7689. [PubMed: 1560003]
- Leonard K, Wingfield P, Arod T, Weiss H. *J Mol Biol*. 1981; 149:259–274. [PubMed: 6273583]
- Li LB, Zou YP, Yu L, Yu CA. *Biochim Biophys Acta*. 1991; 1057:215–222. [PubMed: 2015249]
- Link TA, Haase U, Brandt U, von Jagow G. *J Bioenerg Biomembr*. 1993; 25:221–232. [PubMed: 8394318]
- Martinez SE, Szezepaniak A, Smith JL, Cramer WA. *Biophys J*. 1991; 59:524a.
- Martinez, SE.; Smith, JL.; Huang, D.; Szezepaniak, A.; Cramer, WA. *Research in Photosynthesis. Murata, N., editor. Vol. II. Kluwer Academic; Dordrecht: 1992. p. 495-498.*
- Martinez SE, Huang D, Szezepaniak A, Cramer WA, Smith JL. 1993 Manuscript in preparation.
- Milburn MV, Prive GG, Mittigan DL, Scott WG, Yeh J, Jancarik J, Koshland DE Jr, Kim SH. *Science*. 1991; 254:1342–1347. [PubMed: 1660187]
- Morand L, Frame MK, Colvert KK, Johnson DA, Krogmann DW, Davis DJ. *Biochemistry*. 1989; 28:8039–8047. [PubMed: 2605172]
- Mörschel E, Staehelin A. *J Cell Biol*. 1983; 97:301–310. [PubMed: 6193124]
- Nalecz MJ, Azzi A. *Arch Biochem Biophys*. 1985; 240:921–931. [PubMed: 2992386]
- Nieboer P, Berden JA. *Biochim Biophys Acta*. 1992; 1101:90–96. [PubMed: 1321664]
- Nitschke W, Hauska G, Crofts AR. *FEBS Lett*. 1988; 232:204–208.
- Ohnishi T, Schägger H, Meinhardt SW, LoBrutto R, Link TA, von Jagow G. *J Biol Chem*. 1989; 264:735–744. [PubMed: 2536023]
- Osiewicz HD. *Arch Microbiol*. 1992; 157:336–342. [PubMed: 1590707]
- Pakula A, Simon M. *Nature*. 1992; 355:496–497. [PubMed: 1741027]
- Pfanner N, Neupert W. *Annu Rev Biochem*. 1990; 59:331–353. [PubMed: 2197978]

- Phillips AC, Gray JC. *Mol Gen Genet.* 1984; 194:477–484.
- Piccioni, R.; Bellemare, G.; Chua, N-H. *Methods in Chloroplast Molecular Biology.* Edelman, M., et al., editors. Vol. Chap 80. 1982. p. 985-1015.
- Popot JL, Engelman DE. *Biochemistry.* 1990; 29:4031–4039. [PubMed: 1694455]
- Qin L, Kostic NM. *Biochemistry.* 1993; 32:6073–6080. [PubMed: 8507642]
- Redinbo MR, Yeates TO, Merchant S. *J Bioenerg Biomem.* 1994 in press.
- Reimann A, Kueck U. *Plant Mol Biol.* 1989; 13:255–256. [PubMed: 2519117]
- Reverdatto SV, Andreeva AV, Buryakova AA, Chakhmakhcheva OG, Efimov VA. *Nucleic Acids Res.* 1989; 17:2859–2860. [PubMed: 2654887]
- Rich PR, Madgwick SA, Moss DA. *Biochim Biophys Acta.* 1991; 1058:312–328.
- Rich P, Madgwick SA, Brown S, von Jagow G, Brandt U. *Photosynth Res.* 1992; 34:465–477. [PubMed: 24408840]
- Riedel A, Rutherford W, Hauska G, Müller A, Nitschke W. *J Biol Chem.* 1991; 266:17838–17844. [PubMed: 1655728]
- Rigby SEJ, Moore GR, Gray JC, Godsby PMA, George SJ, Thomson AJ. *Biochem J.* 1988; 256:571–577. [PubMed: 3223931]
- Robertson DE, Dutton PL. *Biochim Biophys Acta.* 1988; 935:273–291. [PubMed: 2844257]
- Robertson DE, Daldal F, Dutton PL. *Biochemistry.* 1990; 29:11249–11260. [PubMed: 2176897]
- Rock CD, Barkan A, Taylor WC. *Curt Genet.* 1987; 12:69–77.
- Saraste M. *FEBS Lett.* 1984; 166:367–372. [PubMed: 6363134]
- Schägger H, von Jagow G. *Anal Biochem.* 1991; 199:223–231. [PubMed: 1812789]
- Schlunegger B, Stutz E. *Curr Genet.* 1984; 8:629–634. [PubMed: 24178003]
- Schmitt ME, Trumpower BL. *J Biol Chem.* 1990; 265:17005–17011. [PubMed: 2170363]
- Selak MA, Whitmarsh J. *FEBS Lett.* 1982; 150:286–292.
- Shinozaki K, Ohme M, Tanaka M, Wakasugi T, Hayashida N, Matsubayashi T, Zaita N, Chunwongse J, Obokata J, Yamaguchi-Shinozaki K, Ohto C, Torazawa K, Meng BY, Sugita M, Deno H, Kamogashira T, Yamada K, Kusuda J, Takaiwa F, Kato A, Tohdoh N, Shimda H, Sugiura M. *EMBO J.* 1986; 5:2043–2049. [PubMed: 16453699]
- Shiver JW, Peterson AA, Widger WR, Cramer WA. *Methods Enzymol.* 1989; 172:439–461. [PubMed: 2747538]
- Siedow JN, Vickery LE, Palmer G. *Arch Biochem Biophys.* 1980; 203:101–107. [PubMed: 6250479]
- Simpkin D, Palmer G, Devlin FJ, McKenna MC, Jensen GM, Stephens PJ. *Biochemistry.* 1989; 28:8033–8039. [PubMed: 2557894]
- Szczepaniak A, Cramer WA. *J Biol Chem.* 1990; 265:17720–17726. [PubMed: 1698778]
- Szczepaniak AD, Huang D, Keenan TW, Cramer WA. *EMBO J.* 1991; 10:2757–2764. [PubMed: 1915260]
- Thony-Meyer L, James P, Hennecke H. *Proc Natl Acad Sci USA.* 1991; 88:5001–5005. [PubMed: 1647023]
- Trumpower BL. *J Biol Chem.* 1990; 265:11409–11412. [PubMed: 2164001]
- Turmel M, Boulanger J, Bergeron A. *Nucleic Acids Res.* 1989; 17:3593–3593. [PubMed: 2726496]
- Verner K, Schatz G. *Science.* 1988; 241:1307–1313. [PubMed: 2842866]
- Von Jagow G, Sebald W. *Ann Rev Biochem.* 1980; 49:281–314. [PubMed: 6250444]
- Von Heijne G. *J Mol Biol.* 1992; 225:487–494. [PubMed: 1593632]
- Weiss H, Kolb HJ. *Eur J Biochem.* 1979; 99:139–149. [PubMed: 226365]
- Widger WR, Cramer WA, Herrmann R, Trebst A. *Proc Natl Acad Sci, USA.* 1984; 81:674–678. [PubMed: 6322162]
- Widger, WR.; Cramer, WA. *Cell Culture and Somatic Cell Genetics of Plants: The Molecular Biology of Plastids and the Photosynthetic Apparatus.* Vasil, IK.; Bogorad, L., editors. Academic Press; Orlando: 1991. p. 149-176.
- Willey DL, Auffret AD, Gray JC. *Cell.* 1984; 36:555–562. [PubMed: 6319031]
- Yang M, Trumpower BL. *J Biol Chem.* 1993 in press.

Yun CH, Van Doren SR, Crofts AR, Gennis RB. J Biol Chem. 1991; 266:10967–10973. [PubMed: 1645718]



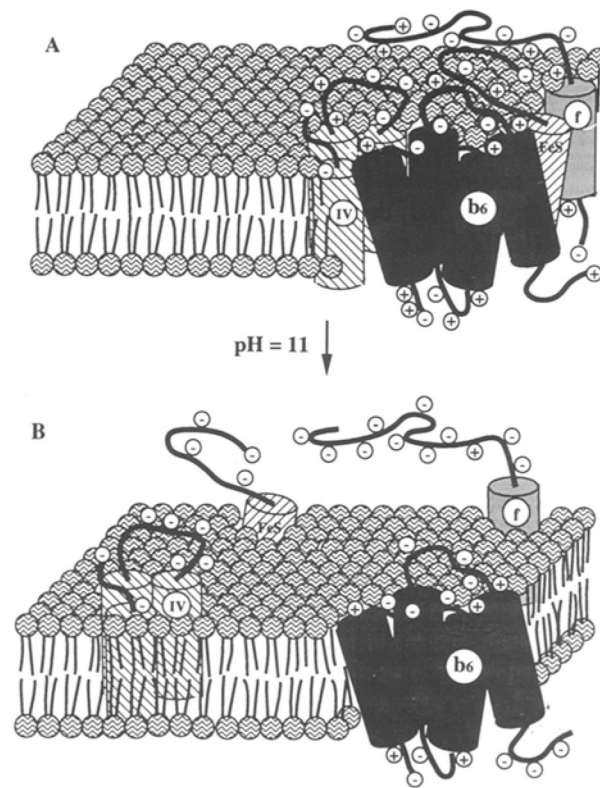




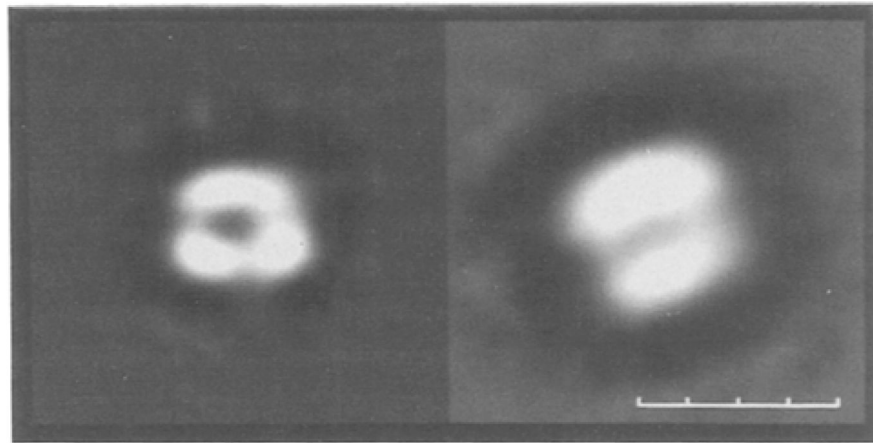
**Fig. 1.**

Folding pattern across the thylakoid membrane of (A) cytochrome  $b_6$  and (B) subunit IV predicted by the distribution and conservation in the 13 and 16 sequences shown in Table IA, B of long (ca. 20 residue) hydrophobic segments. The hydrophobic membrane bilayer is delineated by the rectangular boxes. Hatched circles represent residues that are not invariant or pseudo-invariant. (C) Molecular model (stereo view) of heme-bridged helices B and D of spinach cyt  $b_6$ . The amino acid sequences of these helices are: B helix, NH<sub>2</sub> ( $p$ -side)-SVHRWSASMMVLMMLHVF-COOH; D helix, NH<sub>2</sub> ( $p$ -side)-FYSLHTFVLPLLTAVFMLMHFLMI-COOH. *Nomenclature:* (1) The two sides of the membrane are designated  $n$ - and  $p$ -, the electrochemically negative and positive sides, which correspond to stroma and lumen, respectively, in the case of chloroplast thylakoid membranes. This notation is chosen instead of  $i$  and  $o$  proposed in the recent review of Degli Esposti *et al.* (1993) because it is conceptually simple, which is important for teaching purposes, allows comparative discussion of bacteria, chloroplasts, and mitochondria without reference to the definition of all of the respective extramembrane compartments (lumen, matrix, cytoplasmic, stroma, intermembrane, periplasmic), and is unambiguous, unlike  $i$  and  $o$  which can readily be mistaken by nonspecialists for “inside” and “outside.” (2) Putative *trans*-membrane helices and connecting peripheral domains are labeled A – G and ab...fg, respectively, according to the proposal of Crofts *et al.* (1990). (3) Cytochrome  $b$  hemes on the  $n$ - and  $p$ -sides of membrane coordinated by His-99 and His-201 and His-85 and His-186, respectively, bridging helices B and D, are designated hemes  $b_n$  and  $b_p$ , respectively. The notation is preferred over that of  $b_h$  and  $b_l$ , for high- and low-potential hemes, proposed by Degli Esposti *et al.* (1993) because (a) there is experimental disagreement for cyt  $b_6$  *in situ* concerning the existence of experimentally resolvable high- and low-potential  $b$  hemes, and

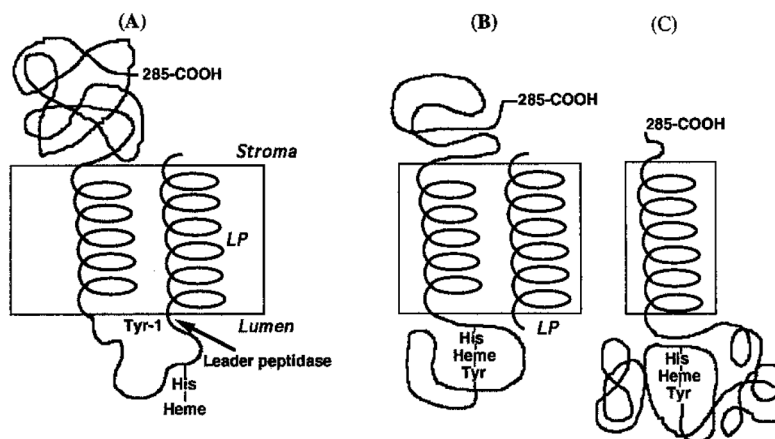
(b) the redox titration of cyt *b* of the  $bc_1$  complex from the PS3 thermophilic bacterium, which shows some similarity to the  $b_6f$  complex, does not show a discernible difference in  $E_m$  of the two hemes (Kutoh and Sone, 1988). Regarding (a), two reports (Furbacher *et al.*, 1989; Rich *et al.*, 1991) show that the  $E_m$  values of the two hemes are not separately resolved ( $E_m < 50\text{mV}$ ) in thylakoid membranes in the presence of  $\text{Mg}^{2+}$ ; one (Kramer and Crofts, 1990) showed potentiometric resolution ( $b_h = -15\text{ mV}$ ,  $b_l = -110\text{ mV}$ ,  $E_m = 95\text{ mV}$ ) of the two hemes. The latter result was obtained in the absence of  $\text{Mg}^{2+}$  but is in close agreement with the heme potentials measured in the isolated  $b_6f$  complex (Hurt and Hauska, 1983; Nitschke *et al.*, 1988). It was proposed (Kramer and Crofts, 1990) that the use of a continuous high-measuring light intensity in the redox titrations of Girvin and Cramer (1984) and Furbacher *et al.* (1989) might account for the discrepancy, but the intensity ( $0.6\ \mu\text{E} \cdot \text{m}^{-2} \cdot \text{sec}^{-1}$ ) found by Kramer and Crofts (1990) to cause cytochrome oxidation was at least 30 times that used ( $10\text{--}20\ \text{nE} \cdot \text{m}^{-2} \cdot \text{sec}^{-1}$ ) by the latter authors.



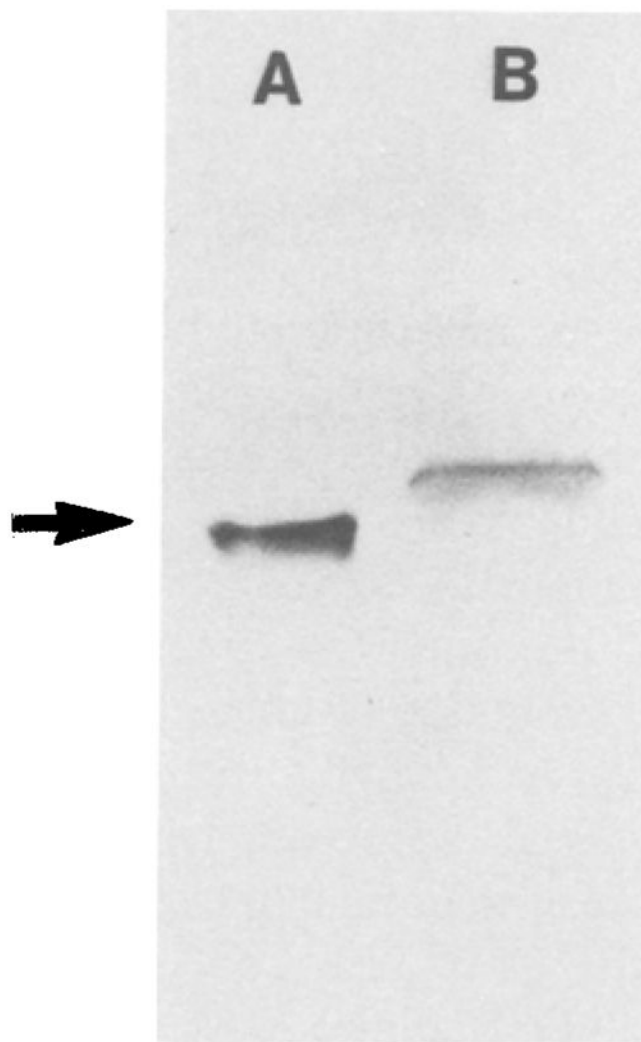
**Fig. 2.** Schematic diagram of the major subunits of the cytochrome  $b_6f$  complex (A) as an organized cluster at neutral ambient pH, and (B) after lateral separation of the subunits in the membrane at alkaline pH (Cramer *et al.*, 1992; figure reproduced with permission of The American Society for Biological Chemistry and Molecular Biology).



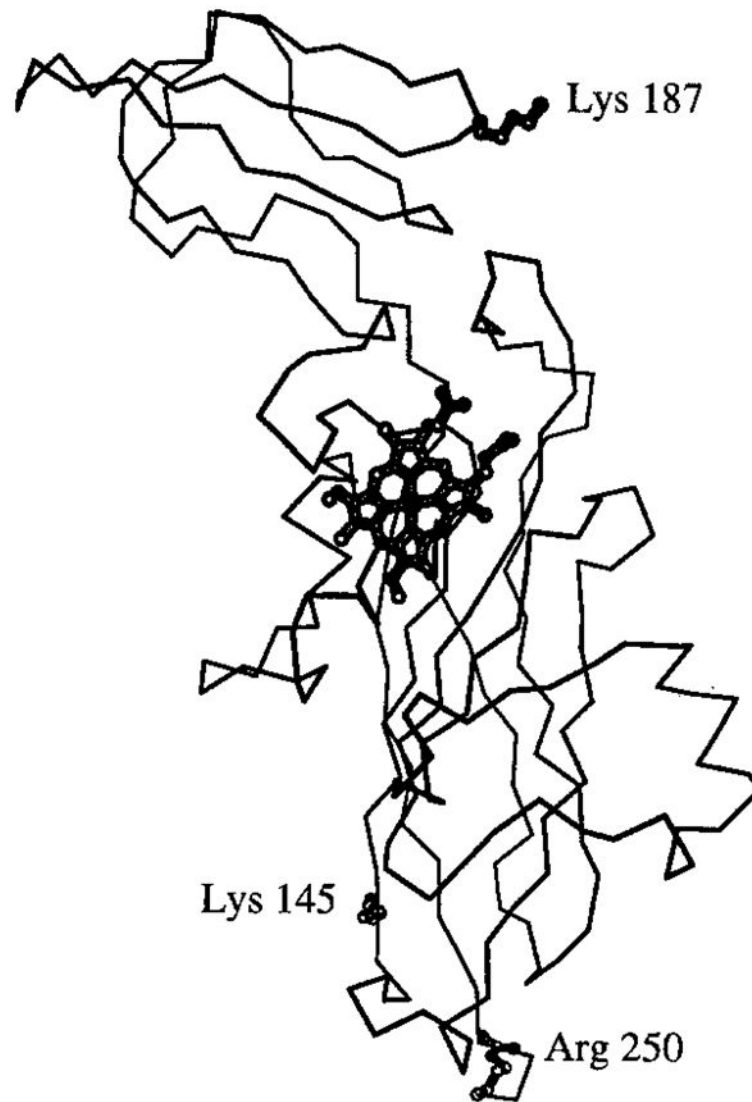
**Fig. 3.** Electron micrographs of negatively stained (left) monomer and (right) dimer forms of the cytochrome *b<sub>6</sub>f* complex. Monomer and dimer images were computed as the average of 58 and 59 single-particle images, respectively (Huang *et al.*, 1993). Size bar = 10 nm.



**Fig. 4.** (A, B) Model for processing and assembly of intermediate forms of *cyt f* and (C) membrane topography of mature form. The schematic shows that complete heme ligation involving the  $\alpha$ -amino group of Tyr-1 cannot occur until the 35-residue leader peptide (LP) has been removed by the processing peptidase that is believed to reside on the lumen side of the membrane (Johnson *et al.*, 1991; Kirwin *et al.*, 1991).



**Fig. 5.** SDS-PAGE of cytochrome *f* purified from turnip thylakoid membranes (lane A), compared to that of the cytochrome in membranes that were dissolved in SDS and run on the gel without prior extraction (lane B).  $M_r = 3,000$  between the two bands that correspond to polypeptides with (A) the 252 residue turnip *cyt f* fragment (arrow) and (B) the complete turnip *cyt f* polypeptide with 285 residues. The molecular weight of the 252-residue polypeptide is 27,500 including the covalently bound heme. The purification procedure was essentially as described by Gray (1978), except for the use of PMSF (0.5 mM), benzamidine (2 mM), and  $\epsilon$ -aminocaproic acid (2 mM) (Martinez *et al.*, 1992, 1993). For electrophoresis, the cytochrome was solubilized in 50 mM Tris-HCl, pH 8.6, 4 M urea, 4% SDS, 10% glycerol, and 5%  $\beta$ -mercaptoethanol, and run with a gel system containing 15% acrylamide (Piccioni *et al.*, 1982).



**Fig. 6.**  $\alpha$ -Carbon model of the lumen-side extrinsic polypeptide of turnip (*French*, “le navet”; *German*, “die Steckrübe” or “der Kohlrabi”) cytochrome *f* (SDS-PAGE in Fig. 5) showing the large and small domains of the cytochrome, the heme, and residues Lys-145 and Lys-187 in the large and small domains, and Arg-250 at the C-terminus that are separated by 33, 28, and 45 Å, respectively, from the heme iron.

Table I

Compilation of Aligned Amino Acid Sequences of (A) Cytochrome *b<sub>6</sub>* and (B) Subunit IV<sup>a</sup>

(A) Amino Acid Sequence Alignment of Cytochrome <i>b<sub>6</sub></i>							
	1						50
AGMEN	MFTKEV	TS.KLYKWF	NER LEIQ	AISDDI	SSKYV	PPHVN	IFYCLGGITL TCFIIQFATG
HOLLANDICA	MFTKQV	QES.GVYKWF	NDR LEIE	AISDDI	SSKYV	PPHVN	IFYCLGGITL VCFIIQFATG
NOSTOC	.....MA	.NVYDWF	EER LEIQ	AIAEDV	TSKYV	PPHVN	IFYCLGGITL TCFIIQFATG
SPINACH	.....MS	.VYDWF	EER LEIQ	AIAADDI	TSKYV	PPHVN	IFYCLGGITL TCFIIQFATG
BARLEY	.....	KVYDWF	EER LEIQ	AIAADDI	TSKYV	PPHVN	IFYCLGGITL TCFIIQFATG
MAIZE	.....	KVYDWF	EER LEIQ	AIAADDI	TSKYV	PPHVN	IFYCLGGITL TCFIIQFATG
RICE	.....MS	.KVYDWF	EER LEIQ	AIAADDI	TSKYV	PPHVN	IFYCLGGITL TCFIIQFATG
TOBACCO	.....MSVK	VYDWF	EER LEI	PAIADDI	TSKYV	PPHVN	IFYCLGGITL TCFIIQFATG
WHEAT	.....MS	.KVYDWF	EER LEIQ	AIAADDI	TSKYV	PPHVN	IFYCLGGITL TCFIIQFATG
LVWORT	.....M.V	.VYDWF	EER LEIQ	AIAADDI	TSKYV	PPHVN	IFYCLGGITL TCFIIQFATG
GRACILIS	.....MS	.RVYDWF	EER LEIQ	AIAADDV	SSKYV	PPHVN	IFYCLGGITL TCFIIQFATG
REINHARDTII	.....MS	.KVYDWF	EER LEIQ	AIAADDI	TSKYV	PPHVN	IFYCLGGITL TCFIIQFATG
PROTOTHCOI	.....MS	.KIYDWF	EER LEIQ	SIADDI	SSKYV	PPHVN	IFYCLGGITL TCFIIQFATG
							HELIX A
	51						110
AGMEN	FAMTFY	YKPT	VAEAF	TSVQY	IMNEV	NFGWL	IRSIHRWSAS MMVLMMLIHI FRVYLTGGFK
HOLLANDICA	FAMTFY	YKPS	VTEAF	TSVQY	LMNEV	SFGWL	IRSIHRWSAS MMVLMMLIHI FRVYLTGGFK
NOSTOC	FAMTFY	YKPT	VAEAF	SSVQY	IMNEV	NFGWL	IRSIHRWSAS MMVLMMLIHI FRVYLTGGFK
SPINACH	FAMTFY	YRPT	VTEAF	ASVQY	IMTEA	NFGWL	IRSVHRWSAS MMVLMMLIHI FRVYLTGGFK
BARLEY	FAMTFY	YRPT	VTEAF	SSVQY	IMTEA	NFGWL	IRSVHRWSAS MMVLMMLIHI FRVYLTGGFK
MAIZE	FAMTFY	YRPT	VTEAF	SSVQY	IMTEA	NFGWL	IRSVHRWSAS MMVLMMLIHI FRVYLTGGFK
RICE	FAMTFY	YRPT	VTEAF	SSVQY	IMTEA	NFGWL	IRSVHRWSAS MMVLMMLIHI FRVYLTGGFK
TOBACCO	FAMTFY	YRPT	VTEAF	ASVPY	IMTEA	NFGWL	IRSVHRWSAS MMVLMMLIHI FRVYLTGGFK
WHEAT	FAMTFY	YRPT	VTEAF	SSVQY	IMTEA	NFGWL	IRSVHRWSAS MMVLMMLIHI FRVYLTGGFK
LVWORT	FAMTFY	YRPT	VTEAF	SSVQY	IMTEA	NFGWL	IRSVHRWSAS MMVLMMLIHI FRVYLTGGFK
GRACILIS	FAMTFY	YRPT	VTEAF	LSVKY	IMNEV	NFGWL	IRSIHRWSAS MMVLMMLIHI FRVYLTGGFK
REINHARDTII	FAMTFY	YRPT	VAEAF	ASVQY	IMTDV	NFGWL	IRSIHRWSAS MMVLMMLIHI FRVYLTGGFK
PROTOTHCOI	FAMTFY	YRPT	VAEAF	TSVQY	LMTQV	NFGWL	IRSIHRWSAS MMVLMMLIHI FRVYLTGGFK
							HELIX B
	111						170
AGMEN	KPRELT	WVTG	VIMATI	TVSF	GVTGY	SLPWD	QVGYWAVKIV SGVPAIIPV GDQMVELLRG
HOLLANDICA	KPRELT	WVTG	VILAVI	TVSF	GVTGY	SLPWD	QVGYWAVKIV SGVPEAIPLV GPLMVELLRG
NOSTOC	KPRELT	WVSG	VILAVI	TVSF	GVTGY	SLPWD	QVGYWAVKIV SGVPEAIPV GVLIIDLLRG
SPINACH	KPRELT	WVTG	VVLGVL	TASF	GVTGY	SLPWD	QIGYWAVKIV TGVDAIPVI GSPLVELLRG
BARLEY	KPRELT	WVTG	VVLAVL	TASF	GVTGY	SLPWD	QIGYWAVKIV TGVDAIPVI GSPLVELLRG
MAIZE	KPRELT	WVTG	VVLAVL	TASF	GVTGY	SLPWD	QIGYWAVKIV TGVDAIPVI GSPLVELLRG
TOBACCO	KPRELT	WVTG	VVLAVL	TASF	GVTGY	SLPWD	QIGYWAVKIV TGVDAIPVI GSPLVELLRG
WHEAT	KPRELT	WVTG	VVLAVL	TASF	GVTGY	SLPWD	QIGYWAVKIV TGVDAIPVI GSPLVELLRG
LVWORT	KPRELT	WVTG	VILAVL	TVSF	GVTGY	SLPWD	QIGYWAVKIV TGVDAIPVI GSPLVELLRG
GRACILIS	KPRELT	WVTG	IILAIL	TVSF	GVTGY	SLPWD	QVGYWAVKIV TGVPEAIPVI GNFIVELLRG
REINHARDTII	KPRELT	WVTG	VIMAVC	TVSF	GVTGY	SLPWD	QVGYWAVKIV TGVDAIPVI GGFIVELLRG
PROTOTHCOI	KPRELT	WVTG	VIMAVC	TVSF	GVTGY	SLPWD	QIGYWAVKIV TGVDAIPVI GQVLELLRG
							HELIX C



	171				214
AGMEN	GASVGQATLT	RFYSLHTFVL	PWLIAVFMLA	HFLMIRKQGI	SGPL*
HOLLANDICA	SASVGQATLT	RFYSLHTFVL	PWFIAVFMLM	HFLMIRKQGI	SGPL*
NOSTOC	GSSVGQATLT	RYSAHTFVL	PWLIAVFMLF	HFLMIRKQGI	SGPL*
SPINACH	SASVGQSTLT	RFYSLHTFVL	PLLTAVFMLM	HFLMIRKQGI	SGPL*
BARLEY	SASVGQSTLT	RFYSLHTFVL	PLLTAVFMLM	HFLMIRKQGI	SGPL*
MAIZE	SASVGQSTLT	RFYSLHTFVL	PLLTAVFMLM	HFLMIRKQGI	SGPL*
RICE	SASVGQSTLT	RFYSLHTFVL	PLLTAVFMLM	HFLMIRKQGI	SGPL*
WHEAT	SASVGQSTLT	RFYSLHTFVL	PLLTAVFMLM	HFLMIRKQGI	SGPL*
TOBACCO	SASVGPSTLT	RFYSLHTFVL	PLLTAVFMLM	HFLMIRKQGI	SGPL*
LVWORT	SVSVGQSTLT	RFYSLHTFVL	PLLTAVFMLM	HFLMIRKQGI	SGPL*
GRACILIS	SVSVGQSTLT	RFYSLHTFVL	PLLTATFMLG	HFLMIRKQGI	SGPL*
REINHARDTII	GVGVGQATLT	RFYSLHTFVL	PLLTAVFMLM	HFLMIRKQGI	SGPL*
PROTOTHCOI	GVAVGQSTLT	RFYSLHTFVL	PLTAVFMLM	HFLMIRKQGI	SGPL*

## HELIX D

## (B) Amino Acid Sequence Alignment of Subunit IV

	215				264
SPINACH	MGVTKKPDLN	DPVLRACLAK	GMGHNYGEP	AWPNDLLYIF	PVVILGTIAC
BARLEY	MGVTKKPDLN	DPVLRACLAK	GMGHNYGEP	AWPNDLLYIF	PVVILGTIAC
MAIZE	MGVTKKPDLN	DPVLRACLAK	GMGHNYGEP	AWPNDLLYIF	PVVILGTIAC
RICE	MGVTKKPDLN	DPVLRACLAK	GMGHNYGEP	AWPNDLLYIF	PVVILGTIAC
TOBACCO	MGVTKKPDLN	DPVLRACLAK	GMGHNYGEP	AWPNDLLYIF	PVVILGTIAC
PEA	MG-----	-----	---HNYGEP	AWPNDLLYIF	PVVILGTIAC
WHEAT	MGVTKKPDLN	DPVLRACLAK	GMGHNYGEP	AWPNDLLYIF	PVVILGTIAC
LVWORT	MGVTKKPDLN	DPVLRACLAK	GMGHNYGEP	AWPNDLLYIF	PVVILGTIAC
REINHARDTII	MSVTKKPDLN	DPVLRACLAK	GMGHNYGEP	AWPNDLLYIF	PVVILGTIAC
EUGAMETOS	MSVTKKPDLN	DPVLRACLAK	GMGHNYGEP	AWPNDLLYIF	PVVILGTIAC
PROTOTHCOI	MAVTKKPDLN	DPVLRACLAK	GMGHNYGEP	AWPNDLLYIF	PVVILGTIAC
AGMEN	MSIMKKPDLN	DPVLRACLAK	GMGHNYGEP	AWPNDLLYIF	PVVILGTIAC
S. 6803	MSIIMKKPDLN	DPVLRACLAK	GMGHNYGEP	AWPNDLLYIF	PVVILGTIAC
NOSTOC	MATQKKPDLN	DPVLRACLAK	GMGHNYGEP	AWPNDLLYIF	PVVILGTIAC
HOLLANDICA	MSVLKKPDLN	DPVLRACLAK	GMGHNYGEP	AWPNDLLYIF	PVVILGTIAC
S. obliquus	MSVTKKPDLN	DPVLRACLAK	GMGHNYGEP	AWPNDLLYIF	PVVILGTIAC

## HELIX E

	265				314
SPINACH	NVGLAVLEPS	MIGEPADPFA	TPLEILPEWY	FFPVFQILRT	VPNKLLGVLL
BARLEY	NVGLAVLEPS	MIGEPADPFA	TPLEILPEWY	FFPVFQILRT	V-----
MAIZE	NVGLAVLEPS	MIGEPAD---	-----	-----	-----
RICE	NVGLAVLEPS	MIGEPADPFA	TPLEILPEWY	FFPVFQILRT	VPNKLLGVLL
TOBACCO	NVGLAVLEPS	MIGEPADPFA	TPLEILPEWY	FFPVFQILRT	VPNKLLGVLL
PEA	NVGLAVLEPS	MIGEPADPFA	TPLEILPEWY	FFPVFQILRT	VPNKLLGVLL
WHEAT	NVGLAVLEPS	MIGEPADPFA	TPLEILPEWY	FFPVFQILRT	VPNKLLGVLL
LVWORT	TVGLAVLEPS	MIGEPANPFA	TPLEILPEWY	FFPVFQILRT	VPNKLLGVLL
REINHARDTII	VIGLSVLDPA	AMGEPANPFA	TPLEILPEWY	FYPVFQILRV	VPNKLLGVLL
EUGAMETOS	CIGLAVLDPA	AMGEPANPFA	TPLEILPEWY	FYPVFQILRV	VPNKLLGVLL
PROTOTHCOI	VIGLAVLDPA	AIGEPANPFA	TPLEILPEWY	FYPVFQILRV	VPNKLLGVLL
AGMEN	ITGLAILDPA	MIGEPGNPFA	TPLEILPEWY	LYPVFQILRV	VPNKLLGIAC
S. 6803	IAGLAILDPA	MIGEPADPFA	TPLEILPEWY	LYPTFQILRI	LPNKLLGIAC
NOSTOC	IVALAVLDPA	MTGEPANPFA	TPLEILPEWY	LYPVFQILRS	VPNKLLGVLL
HOLLANDICA	VVGLAVLDPA	MVGEANPFA	TPLEILPEWY	LYPVFQILRV	VPNKLLGILL
S. obliquus	VIGLSVLDPA	AIGEPANPFA	TPLEILPEWY	FYPVFQILRV	VPNKLLGVLL

	315				364
SPINACH	MASVPAGLLT	VPFLENVNKF	QNPFRPVAT	TVFLVGTVVA	LWLGIGATLP
BARLEY	-----	-----	-----	-----	-----
MAIZE	-----	-----	-----	-----	-----
RICE	MVSVPAGLLT	VPFLENVNKF	QNPFRPVAT	TVFLIGTAVA	LWLGIGATLP
TOBACCO	MVSVPAGLLT	VPFLENVNKF	QNPFRPVAT	TVFLIGTAVA	LWLGIGATLP
PEA	MVSVPAGLLT	VPFLENVNKF	QNPFRPVAT	TVFLIGTAVA	LWLGIGATLP
WHEAT	MVSVPAGLLT	VPFLENVNKF	QNPFRPVAT	TVFLIGTAVA	LWLGIGATLP
LVWORT	MAAVPAGLLT	VPFLENVNKF	QNPFRPVAT	TVFLIGTAVA	LWLGIGAALP
REINHARDTII	MAAVPAGLLT	VPFLENVNKF	QNPFRPVAT	TVFLIGTAVA	LWLGIGAALP
EUGAMETOS	MAAVPAGLLT	VPFLENVNKF	QNPFRPVAT	TVFLIGTAVA	LWLGIGAALP
PROTOTHCOI	MAAVPAGLLT	VPFLENVNKF	QNPFRPVAT	TVFLIGTAVA	LWLGIGAALP
AGMEN	QGAIPGLMM	VPFIESVNKF	QNPFRPVAM	AVFLFGTAVT	LWLGAGACFP
S. 6803	MAAIPGLML	VPFIESVNKF	QNPFRPIAM	TVFLFGTAAA	LWLGAGATFP
NOSTOC	MASVPLGLL	VPFIENVNKF	QNPFRPVAT	TVFLFGTLVT	LWLGIGAALP
HOLLANDICA	QTAIPGLML	VPFIENVNKF	QNPFRPIAM	AVFLFGTLVT	LWLGIGAALP
S. obliquus	MAAVPAGLLT	VPFIENVNKF	QNPFRPIAM	TVFLFGTLVA	LWLGIGATLP
	HELIX F			HELIX G	

	365	374
SPINACH	IDKSLTLGLF*	
BARLEY	-----*	
MAIZE	-----*	
RICE	IEKSLTLGLF*	
TOBACCO	IDKSLTLGLF*	
PEA	IEKSLTLGLF*	
WHEAT	IDKSLTLGLF*	
LVWORT	IDKSLTLGLF*	
REINHARDTII	IDISLTLGLF*	
EUGAMETOS	IDISLTLGLF*	
PROTOTHCOI	IDISLTLGLF*	
AGMEN	IDESLTLGLF*	
S. 6803	IDKSLTLGLF*	
NOSTOC	LDKSLTLGLF*	
HOLLANDICA	IDKFTLGLF*	
OOCSP	IEISLTFGLF*	

<sup>a</sup> AGMEN (*Agmenellum quadruplicatum*. Brand *et al.*, 1992); NOSTOC (Cyanobacterium Nostoc strain PCC 7906. Kallas *et al.*, 1988); HOLLANDICA (*Prochlorothrix hollandica*. Greer and Golden, 1992); BARLEY (*Hordeum vulgare*. Reverdatto *et al.*, 1989); SPINACH (*Spinacia oleracea*. Heinemeyer *et al.*, 1984); WHEAT (*Triticum aestivum*. Hird *et al.*, 1991); MAIZE (*Zea mays*. Rock *et al.*, 1987, PROTOTHCOI (*Chlorella protothecoides*. Reimann and Kueck, 1989); RICE (*Oryza sativa*. Cote *et al.*, 1988; Hiratsuka *et al.*, 1989); TOBACCO (*Nicotiana tabacum*. Shinozaki *et al.*, 1986); GRACILIS (*Euglena gracilis*. Schlunegger and Stutz, 1984); S. 6803 (*Synechocystis* sp. PCC 6803. Osiewacz, 1992); EUGAMETOS (*Chlamydomonas eugametos*. Tunnel *et al.*, 1989); REINHARDTII (*Chlamydomonas reinhardtii*. Büschlen *et al.*, 1991); LVWORT (*Marchantia polymorpha*. Fukuzawa *et al.*, 1987); PEA (*Pisum sativum*. Phillips and Gray, 1984); S. obliquus (*Scenedesmus obliquus*. Kück, 1989).

**Table II**

Summary of Residues of Cyt  $b_6$  and SU IV in Sequence Compilations of Tables IA, B and Figs. 2A, B That Are Most Highly Conserved (Degli Esposti *et al.*, 1993)

Residue	Invariance	Location (Figs. 2A, B)	Function
<b>Cytochrome <math>b_6</math></b>			
Gly 36	Complete <sup>b</sup>	<i>n</i> -side, A helix	Heme pocket <sup>a</sup>
Gly 50	Essentially complete <sup>b</sup>	A helix	—
Gly 78	Essentially complete <sup>b</sup>	“ab” region	Turn between “ab” and B helices
Arg 82	(+) Residue essentially complete <sup>b</sup>	<i>p</i> -side interface, B helix	Stop transfer, B helix
His 85	Complete <sup>b</sup>	<i>p</i> -side, B helix	Ligand of <i>p</i> -side heme <sup>c,d</sup>
Ser 90	Incomplete <sup>b</sup>	B helix	—
His 99	Complete <sup>b</sup>	<i>n</i> -side, B helix	Ligand of <i>n</i> -side heme <sup>c,d</sup>
Arg 102	(+) Residue complete <sup>b</sup>	<i>n</i> -side interface, B helix	Stop transfer, B helix
Trp 117	Complete <sup>b</sup>	<i>n</i> -side, C helix	—
Gly 120	Essentially complete <sup>b</sup>	<i>n</i> -side, C helix	—
Gly 134	Essentially complete <sup>b</sup>	<i>p</i> -side, C helix	—
Tyr 135	Incomplete <sup>b</sup>	<i>p</i> -side, C helix	—
Trp 145	Incomplete <sup>b</sup>	<i>p</i> -side, “cd” helix	—
Thr 178	Essentially complete <sup>b</sup>	<i>p</i> -side interface—“cd” helix	—
His 186	Complete <sup>b</sup>	<i>p</i> -side, D helix	Ligand of <i>p</i> -side heme <sup>c,d</sup>
His 201	Complete <sup>b</sup>	<i>n</i> -side, D helix	Ligand of <i>n</i> -side heme <sup>c,d</sup> One extra residue between two His ligands on D helix <sup>c</sup>
<b>SU IV<sup>e</sup></b>			
Asp 249	Essentially complete <sup>b</sup>	<i>n</i> -side interface, E helix	—
Ile 289	Incomplete <sup>b</sup>	<i>p</i> -side, “ef” loop	—
Pro 291	Complete <sup>b</sup>	<i>p</i> -side, “ef” loop	—
Glu 292	Essentially complete <sup>b</sup>	<i>p</i> -side, “ef” loop	—
Trp 293	Essentially complete <sup>b</sup>	<i>p</i> -side, “ef” loop	—
Leu 302	Essentially complete <sup>b</sup>	<i>p</i> -side, “ef” loop	—
Lys 308	Incomplete <sup>b</sup>	<i>p</i> -side interface, F helix	Stop transfer, F helix
Gly 311	Essentially complete <sup>b</sup>	<i>p</i> -side interface, F helix	Turn between “ef” and F helix

<sup>a</sup>Tron *et al.*, 1991; Yun *et al.*, 1991;

<sup>b</sup>Degli Esposti *et al.*, 1993;

<sup>c</sup>Widger *et al.*, 1984;

<sup>d</sup>Saraste, 1984;

<sup>e</sup> note that the numbering of SU IV is different from that used by Degli Esposti *et al.* (1993) who omitted the first 30 residues of the SU IV sequence. Use of the numbering system of the latter authors would cause overlap of numbering with the cyt *b*<sub>6</sub> sequence from residue 195–214. The basis of the numbering system used in the present work is based on the concept that SU IV is a functional extension of cyt *b*<sub>6</sub>.

**Table III**Amino Acid sequence of Turnip Cytochrome *f* (Gray, 1992).

1	YPIFAQQNYE	NPREATGRIV	CANHLASKP	VDIEVPQAVL	PDTVFEAVVK	50
51	IPYDMQLKQV	LANGKKGALN	VGAVLILPEG	FELAPPDRIS	PEMKEKIGNL	100
101	SFQNYRPNKK	NILVIGPVPV	QKYSEITFPI	LAPDPATNKD	VHFLKYPIYV	150
151	GGNRGRGQIY	PDGSKSNNTV	YNATAGGIIS	KILRKEKGGY	EITIVDASNE	200
201	RQVIDIIPRG	LELLVSEGES	IKLDQPLTSN	PNVGGFGQGD	AEIVLQDPLR	250
251	<u>VQGLLFFLGS</u>	<u>VVLAQIFLVL</u>	KKKQFEKVQL	SEMNF*		285
	↑					