Photosynthesis Research **24:** 1–13, 1990.
© 1990 *Kluwer Academic Publishers. Printed in the Netherlands.*

Minireview

The structure and function of CPa-1 and CPa-2 in Photosystem II*

Terry M. Bricker
Departments of Botany and Biochemistry, Louisiana State University, Baton Rouge, LA 70803, USA

Received 29 May 1989; accepted 14 September 1989

Key words: CPa-1, CPa-2, CP47, CP43, Photosystem II

Abstract

This review presents a summary of recent investigations examining the structure and function of the chlorophyll-proteins CPa-1 (CP47) and CPa-2 (CP43). Comparisons of the derived amino acid sequences of these proteins suggest sites for chlorophyll binding and for interactions between these chlorophyll-proteins and other Photosystem II components. Hydropathy plot analysis of these proteins allows the formulation of testable hypotheses concerning their topology and orientation within the photosynthetic membrane. The role of these chlorophyll-proteins as interior light-harvesting chlorophyll-*a* antennae for Photosystem II is examined and other possible additional roles for these important Photosystem II components are discussed.

1. Introduction

In the late 70's, a minor chlorophyll-protein which, during electrophoresis, migrated between CP I and CP II (LHCP) was identified by a number of workers (Hayden and Hopkins 1977, Anderson et al. 1978, Henriques and Park 1978, Wessels and Borchert 1978). This chlorophyll-protein was termed CPa and was found to contain no detectable chlorophyll-*b*. Delepelaire and Chua (1979), using low temperature LiDS–PAGE, were able to resolve CPa into two components, which they termed CP III and CP IV in the green alga *Chlamydomonas*; they had earlier established that the apoproteins of these chlorophyll-proteins were associated with Photosystem II (PS II) (Chua and Benoun 1975). Soon thereafter these chlorophyll-proteins were observed by numerous workers with a variety of plant and algal materials using a number of different electrophoretic systems. These chlorophyll-proteins are termed CPa-1 (or CP47) and CPa-2 (or CP43) in the current literature (see Green 1988 for a more complete discussion). CPa-1 and CPa-2 each consist of a single polypeptide chain. The apparent molecular masses of the apoproteins (from spinach) are 45–51 kDa and 40–45 kDa, respectively, as determined by polyacrylamide gel electrophoresis. The lower apparent molecular masses are observed after strongly denaturing electrophoretic conditions (i.e. heated samples or inclusion of urea). The primary functional roles which have been assigned to these proteins is that of interior chlorophyll-*a* antennae for PS II. In this capacity, these proteins appear to serve as intermediary transducers of excitation energy from the light-harvesting pigment bed (light-harvesting chlorophyll-*a*/*b* proteins in green algae and higher plants and phycobilosomes in cyanobacteria) to the reaction center of PS II. Other possible roles for these proteins in PS II cannot be excluded based on the experimental data which are currently available.

At least seven proteins appear to be required for photosynthetic oxygen evolution within the PS II membrane protein complex (Ghanotakis et al. 1987). These include CPa-1 (CP47), CPa-2 (CP43), D1, D2, the extrinsic manganese-stabilizing protein, and the α and β subunits of cytochrome b_{559}. Additionally, several low molecular mass polypeptides have recently been identified which appear to be associated with PS II (Ikeuchi et al. 1989). At least one of these, the 4.1 kDa product of the *psb* I

*This work was supported by NSF Grant DMB-8740292.

gene, appears to be associated with the PSII reaction center (Ikeuchi and Inoue 1988). The oxygen-evolving complex also contains a number of inorganic cofactors which are absolutely required for water oxidation. These cofactors include bound manganese, calcium and chloride (Amesz 1983, Dismukes 1986, Homann 1987). It is possible that, *in vivo*, the proteins D1, D2, cytochrome b_{559}, and the *psb* I gene product form the reaction center of PS II and that these proteins bear as prosthetic groups the primary and secondary electron acceptors and donors of the photosystem (Y_Z, P_{680}, Pheophytin, Q_A, and, probably, Q_B). PS II core preparations which contain these proteins have been isolated from both higher plants (Nanba and Satoh 1987, Akabori et al. 1988, Ghanotakis et al. 1989) and recently from cyanobacteria (Gounaris et al. 1989). These preparations, however, do not contain bound manganese and are non-functional with respect to oxygen evolution, Additionally, they do not contain Q_A or Q_B, and do not exhibit the characteristic EPR signals associated with Y_Z or Y_D (Frank et al. 1989). It is possible that other intrinsic components of PSII are required to stabilize these functional groups and EPR signals. Additionally, the proteins which bind the inorganic cofactors required for oxygen evolution and the manganese-stabilizing protein have not been unambiguously identified. Interestingly, all PS II submembrane complexes which have been isolated to date which have the ability to evolve oxygen contain the two chlorophyll-proteins CPa-1 and CPa-2. The object of this review is to summarize the structural and functional characteristics of these two proteins with particular emphasis on the identification of possible functionally important domains on these proteins and to suggest other possible roles which these important PS II components may perform.

2. Structural organization of CPa-1 and CPa-2

2.1. CPa-1 structure

CPa-1 is encoded by the *psb*B gene, which is located in the chloroplast genome in higher plants and green algae and in the genomic DNA of cyanobacteria. Comparisons of the derived amino acid sequences of CPa-1 (Fig. 1) from *Synechocystis* 6803 (Vermaas et al. 1987), *Marchantia* (Ohyama et al. 1986), tobacco (Shinozaki et al. 1986) and spinach (Morris and Herrmann, 1984) indicate that this protein is highly conserved. 72% of the amino acids are identical in these four species, a value which increases to 85% if conservative amino acid replacements are included. It is generally assumed that regions on functionally similar but structurally divergent proteins, which exhibit high degrees of similarity, represent functionally important domains. This type of analysis is difficult to apply to CPa-1 from different species since these proteins are functionally identical and exhibit only slight structural divergence. One interesting feature, however, is apparent upon examination of the derived primary structure of CPa-1. 40% of the non-conservative amino acid replacements are found in four small regions encompassing only 12% of the protein (residues 73–89, 182–200, 277–296, and 407–416). In these variable domains almost half of the residues are non-conservatively replaced. These regions all lie on predicted hydrophilic loop portions of CPa-1 which are probably lumenally exposed (see below). Additionally, these variable regions are either immediately adjacent to, or in the vicinity of, predicted membrane-spanning helixes of CPa-1. These variable regions would appear to represent domains in which a considerable amount of amino acid substitution is tolerated without apparent effects on protein functionality.

Numerous lines of biochemical evidence suggests that CPa-1 is an intrinsic membrane protein (Piccioni et al. 1981, Piccioni et al. 1982, Bricker et al. 1983). Hydropathy plots of CPa-1 have been interpreted to suggest the presence of either 7 (Morris and Herrmann 1984) or 6 (Vermaas et al. 1987) transmembrane alpha helixes. It appears that the second transmembrane helix predicted by Morris and Herrmann (1984) is probably too short and too hydrophilic to span the thylakoid membrane. In spinach, the six predicted membrane-spanning alpha helixes appear to be located in approximately the following positions: Helix I, Leu^{18}–Leu^{39}; Helix II, Val^{96}–Trp^{116}; Helix III, Ile^{138}–$1Leu^{161}$; Helix IV, Ile^{198}–Val^{219}; Helix V, Val^{237}–Trp^{257}; Helix VI, Trp^{450}–Gly^{470} (Fig. 2). The positions of these putative alpha helixes and their primary structure are strongly conserved in all species examined.

The distribution of histidyl residues within these helixes is intriguing. Histidyl residues are primary

```
             1                                                                                                           100
       6803  MGLPWYRVHT VVLNDPGRLI SVHLMHTALV AGWAGSMALY ELAIFDSSDA VLNPMWRQGM FVLPFMARLG VTSSWNGWSV TGETGLDPGF WSFEGVAAAH
MARCHANTIA  MGLPWYRVHT VVLNDPGRLI AVHLMHTALV SGWAGSMALY ELAVFDPSDP VLDPMWRQGM FVIPFMTRLG ITKSWGGWSI TGETVTNAGI WSYEGVAAVH
    SPINACH  MGLPWYRVHT VVLNDPGRLI SVHIMHTALV AGWAGSMALY ELAVFDPSDP VLDPMWRQGM FVIPFMTRLG ITNSWGGWSI TGGTITDPSI WSYEGVAGAH
    TOBACCO  MGLPWYRVHT VVLNDPGRLL SVHIMHTALV AGWAGSMALY ELAVFDPSDP VLDPMWRQGM FVIPFMTRLG ITNSWGGWSI TGGTVTNPGI WSYEGVAGAH
  CONSENSUS  MGLPWYRVHT VVLNDPGRL- -VH-MHTALV -GWAGSMALY ELA-FD-SD- VL-PMWRQGM FV-PFM-RLG -T-SW-GWS- TG-T------ WS-EGVA--H
                                   I

             101                                                                                                         200
       6803  IVLSGLLFLA AVWHWVFWDL ELFVDPRTGE SALDLPKMFG IHLFLSGLLC FGFGAFHLTG VWGPGMWVSD PYGLTGHVQP VAPEWGPAGF NPFNPGGVVA
MARCHANTIA  IVLSGLLFLA AIWHWVYWDL ELFRDERTGK PSLDLPKIFG IHLFLSGVLC FAFGAFHVTG LHGPGIWISD PYGLTGKVQP VAPAWGAEGF DPFVPGGIAS
    SPINACH  IMFSGLCFLA AIWHWVYWDL EIFSDERTGK PSLDLPKIFG IHLFLSGVAC FGFGAFHVTG LYGPGIWVSD PYGLTGKVQP VCSAWGVEGF DPFVPRGIAS
    TOBACCO  IVFSGLCFLA AIWHWVYWDL EIFCDERTGK PSLDLPKIFG IHLFLSGVAC FGFGAFHVTG LYGPGIWVSD PYGLTGKVQP VNPAWGVEGF DPFVPGGIAS
  CONSENSUS  I--SGL-FLA A-WHWV-WDL E-F-D-RTG- --LDLPK-FG IHLFLSG--C F-FGAFH-TG --GPG-W-SD PYGLTG-VQP V---WG--GF -PF-P-G---
              II                                        III

             201                                                                                                         300
       6803  HHIAAGIVGI IAGLFHLTVR PPERLYKALR MGNIETVLSS SIAAVFFAAF VVAGTMWYGN ATTPIELFGP TRYQWDKGYF QEEIQRRVDS QLAEGASLSE
MARCHANTIA  HHIAAGILGI LAGLFHLSVR PPQRLYKGLR MGNVETVLSS SIAAVFFAAF VVAGTMWYGS AATPIELFGP TRYQWDQGFF QQEIDRRIRS SKAENLSLSE
    SPINACH  HHIAAGTLGI LAGLFHLSVR SPQRLYKGLR MGNIETVLSS SIAAVFFAAF VVAGTMWYGS ATTPIELFGP TRYQWDQGYF QQEIYRRVSA GLAENQSFSE
    TOBACCO  HHIAAGTLGI LAGLFHLSVR PPQRLYKGLR MGNIETVLSS SIAAVFFAAF VVAGTMWYGS ATTPIELFGP TRYQWDQGYF QQEIYRRVSA GLAENQSLSE
  CONSENSUS  HHIAAG--GI -AGLFHL-VR -P-RLYK-LR MGN-ETVLSS SIAAVFFAAF VVAGTMWYG- A-TPIELFGP TRYQWD-G-F Q-EI-RR--- GLAE--S-SE
              IV                                    V

             301                                                                                                         400
       6803  AWSTIPEKLA FYDYVGNSPA KGGLFRTGAM NSGDGIAQEW IGHPIFKDKE GRELEVRRMP NFFETFPVIM TDADGVVRAD IPFRRSESKF SVEQTGVTVS
MARCHANTIA  AWSKIPEKLA FYDYIGNNPA KGGLFRAGAM DNGDGIAVGW LGHAVFKDKE GNELFVRRMP TFFETFPVVL VDEQGIVRAD VPFRRAESKY SVEQVGVTVE
    SPINACH  AWSKIPEKLA FYDYIGNNPA KGGLFRAGSM DNGDGIAVGW LGHPIFRDKE GRELFVRRMP TFFETFPVVL IDGDGIVRAD VPFRRAESKY SVEQVGVTVE
    TOBACCO  AWSKIPEKLA FYDYIGNNPA KGGLFRAGSM DNGDGIAVGW LGHPIFRDKE GRELFVRRMP TFFETFPVVL VDGDGIVRAD VPFRRAESKY SVEQVGVTVE
  CONSENSUS  AWS-IPEKLA FYDY-GN-PA KGGLFR-G-M --GDGIA--W -GH--F-DKE G-EL-VRRMP -FFETFPV-- -D--G-VRAD -PFRR-ESK- SVEQ-GVTV-

             401                                                                                                         500
       6803  FYGGALDGQT FSNPSDVKKF ARKAQLGEGF DFDTETFNSD GVFRTSPRGW FSFGHAVFAL LFFFGHIWHG SRSLFRDVFA GVDPGLEEQVE FGVFAKVGD
MARCHANTIA  FYGGALDGVS FSDPATVKKY ARRAQLGEIF EFDRATLKSD GVFRSSPRGW FTFGHATFAL LFFFGHIWHG ARTLFRDVFA GIDPDLEAQVE FGAFQKLGD
    SPINACH  FYGGELNEVS YSDPATVKKY ARRAQLGEIF ELDRASLKSD GVFRSSPRGW FSFGHASFAL LFFFGHIWHG SRSLFRDVFA GIDPDLDVQVE FGAFQKIGD
    TOBACCO  FYGGELNGVS YSDPATVKKY ARRAQLGEIF ELDRATLKSD GVFRSSPRGW FTFGHASFAL LFFFGHIWHG ARTLFRDVFA GIDPDLDAQVE FGAFQKLGD
  CONSENSUS  FYGG-L---- -S-P--VKK- AR-AQLGE-F --D-----SD GVFR-SPRGW F-FGHA-FAL LFFFGHIWHG -R-LFRDVFA G-DP-L--QVE FG-F-K-GD
                                                                      VI

             501  508
       6803  LSTRKEA.
MARCHANTIA  PTTKRQVI
    SPINACH  PSSRRQGV
    TOBACCO  PTTKRQAA
  CONSENSUS  --------
```

Fig. 1. Comparison of the derived amino acid sequences of CPa-1 from four species. A consensus sequence is shown; amino acid residues which are not conserved in all species are indicated by a dashed line (-). Predicted location of putative membrane-spanning helixes are shown by underlining and are indicated with Roman numerals (I–VI)

candidates for axial ligands for chlorophyll molecules. All of the bacteriochlorophyll molecules in the *Rhodopseudomonas viridis* reaction center (Deisenhofer et al. 1985), and five of seven chlorophylls in the *Prosthecochloris* light-harvesting chlorophyll-protein (Tronrud et al. 1986), the only chlorophyll-proteins for which crystal structures are available, are ligated to histidyl residues. The location of the twelve histidyl residues which are present in the predicted membrane-spanning helixes of CPa-1 are shown in Fig. 2. All of these histidyl residues are located near the membrane surface. A similar positioning of histidyl residues in membrane-spanning domains has also been predicted in other chlorophyll-proteins such as the alpha and beta polypeptides of the purple and green bacteria antenna complex and the light-harvesting chlorophyll *a*/*b* proteins of higher plants (Zuber et al. 1987). Additionally, Helixes I and VI of both CPa-1 and CPa-2 possess a pair of histidyl residues (. . . HxyH. . .) located near the predicted stromal membrane surface. There exists a striking similarity between the location of histidyl residues in CPa-1 and those located in the membrane-spanning regions of CPa-2 (Fig. 4) even though these two proteins, while obviously related, exhibit only limited overall primary sequence homology (< 20%). Even within the predicted membrane-spanning helixes the homology is only about 30%. This suggests that the position of these histidyl residues may be critical to the functioning of CPa-1 and CPa-2. de Vitry et al. (1984) demonstrated, using ^{14}C labeling techniques, that both CPa-1 and CPa-2 bound 20–25 chlorophylls/protein. Since the total number of conserved histidyl residues in CPa-1 is 14 and in CPa-2 is 12, one must conclude that other sites must act as chlorophyll ligands in these chlorophyll-proteins. Candidates for such alterna-

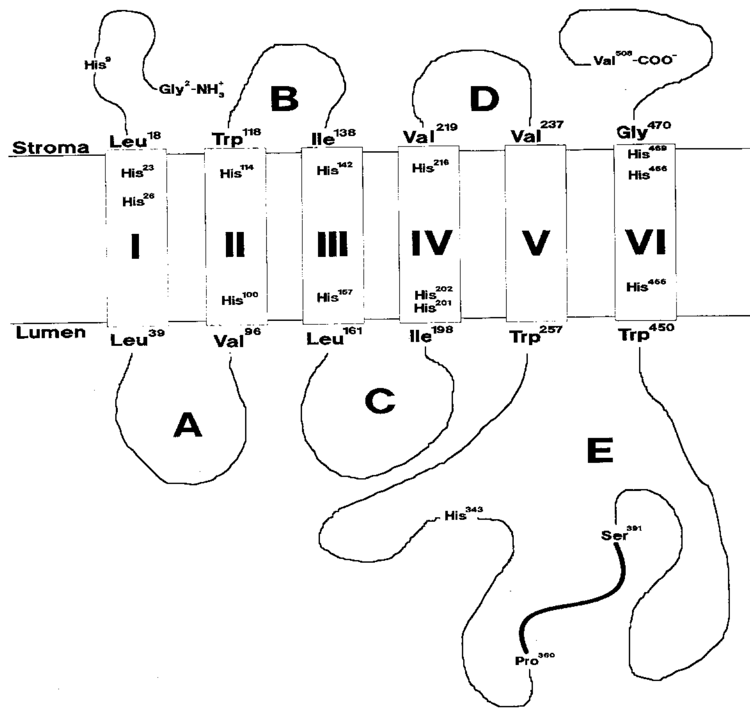

Fig. 2. Predicted membrane folding pattern of spinach CPa-1. Membrane-spanning helixes are indicated by Roman numerals I–VI. The extrinsic loop regions A–E are also indicated. Location of the epitope for the monoclonal antibody FAC2 is shown by a bold line. The approximate locations of the histidyl residues are shown

tive ligands include the amino acid residues Asn and Gln (Wechsler et al. 1985) as well as backbone carbonyl groups and water molecules (Tronrud et al. 1986). In addition to chlorophyll, CPa-1 and CPa-2 also contain beta-carotene (Akabori et al. 1988). The binding site(s) for this pigment have not been determined.

Five hydrophilic loops connect the predicted transmembrane alpha helixes in CPa-1. In spinach, these loops are approximately located in the following positions: Loop A, Tyr^{40}–Gly^{95}, Loop B Glu^{119}–Lys^{137}, Loop C, Tyr^{162}–Gly^{197}, Loop D, Arg^{220}–Glu^{236}, and Loop E, Tyr^{258}–Gly^{449}. Analysis of a variety of intrinsic membrane proteins of both procaryotic and eucaryotic origin has indicated that short hydrophilic segments (<100 amino acid residues) which are translocated across membranes have a much lower abundance of arginyl and lysyl residues than to hydrophilic segments which are not translocated (von Heijne and Gavel 1988). Analysis of CPa-1 by these methods shows that the short hydrophilic segments, Loops B and D, as well as the N-terminus (Met^{1}–Arg^{18}) and C-terminus (Ser^{471}–Val^{508}) of the protein' contain a relatively high abundance of lysyl and arginyl residues suggesting that these segments are not translocated across the thylakoid membrane and are consequently located on the stromal side of the membrane. Loops A and C have a much lower relative abundance of these positively charged amino acids and are probably translocated to the lumenal side of the thylakoid. It should be noted that Loops A and C are also significantly larger than Loops B and D. It is not known if this is of any functional significance. Large hydrophilic domains (>100 amino acid residues), such as Loop E of CPa-1, are apparently translocated across the membrane by different mechanisms than are short segments. No correlation between the amino acid composition of long segments and their translocation was found (von Heijne and Gavel 1988). Location of the other hydrophilic segments of CPa-1 as well as the location of the transmembrane regions of this protein lead to the predition that Loop E is located on the lumenal side of the thylakoid membrane. It should be noted that other models have been presented which have suggested that the N-terminus of CPa-1 is located to the

lumenal side of the thylakoid membrane (Murphy 1986). No direct evidence elucidating the absolute topology of CPa-1 is currently available. Studies, such as those performed on the D1 protein by Sayre et al. (1986), which would unequivocally determine the topology of this membrane-spanning protein, are obviously required.

A word of caution is in order here. The above analysis is based primarily on the examination of hydropathy plots from derived amino acid sequence data to identify transmembrane alpha helixes. It has been noted (Paul and Rosenbusch 1985, Ferenci 1989) that other hydrophobic structures, such as extended beta sheets, *could* span the biological membrane. If arranged in an extended beta sheet configuration, 5–6 hydrophobic residues could span the bilayer. Such structures would be difficult or impossible to identify using currently available computational techniques.

2.2. CPa-2 structure

CPa-2 is encoded by the *psb*C gene, which is also located on the chloroplast genome of higher plants and green algae and in the genomic DNA of cyanobacteria. The derived amino acid sequences for nine species, *Synechocystis* 6803 (Chisholm and Williams 1988), *Synechococcus* 7942 (Golden and Stearns 1988), *Chlamydomonas* (Rochaix et al. 1989), *Marchantia* (Ohayama et al. 1986), spinach (Alt et al. 1984), tobacco (Shinozaki et al. 1986), pea (Bookjans et al. 1986), wheat (Hird S, Dyer T and Gray J unpublished observations) and maize (Krebbers 1983), are shown in Fig. 2. In all species examined, the *psb*C gene overlaps a *psb*D gene (which encodes D2) and is translated in a different reading frame. This overlap accounts for the large number of *psb*C genes which have been sequenced since research on the structure of the *psb*D gene has been quite intensive. Additionally, a partial sequence for *Euglena* (Montandon et al. 1986) is included. The complete sequence for the *Euglena* gene is unavailable at this time since it contains large introns and has not been completely sequenced. CPa-2, like CPa-1, is highly conserved, with 67% of the amino acids being identical, a value which increases to 80% if conservative amino acid replacements are included. Five regions of particularly high variability (residues 1–37, 101–124, 133–150, 201–216 and 462–473) are evident upon examination of the predicted amino acid sequence of CPa-2. Of particular interest is the extremely variable region at the N-terminus of CPa-2. *Synechococcus* and *Chlamydomonas* exhibit an apparent twelve amino acid residue deletion with respect to the other predicted *psb*C sequences which have been determined. Previously, it had been speculated that the actual start codon in *Synechocystis* and *Synechococcus* was a GTG codon thirty-six nucleotides downstream from the putative AUG start codon identified in the other species (Chisholm and Williams 1988, Golden and Stearns 1988). This alternative possible translation start codon is also conserved in all species examined. Additionally, a strong shine–Dalgarno ribosome binding site is positioned upstream from the hypothesized GTG start codon. Thus, two possibilities exist, either *Synechococcus* and *Chlamydomonas* are exhibiting a true twelve amino acid residue N-terminal deletion, or, the actual start codon for all of the species is the GTG start codon. If the first possibility proves to be true, it would appear that these twelve N-terminal residues are neither required for the stable maintenance of this protein in the membrane nor for the formation of functional configuration of this molecule. Resolution of this problem will require S1 nuclease digestion analysis and/or N-terminal sequence information from unprocessed CPa-2 (see below). In this review, the amino acid sequences which were originally presented by the various investigators, are shown in Fig. 3.

Within the other four variable regions of CPa-2, over 40% of the amino acid residues exhibit nonconservative replacement. Two of these variable regions are adjacent to Helix II of CPa-2, a third is situated in the middle of the hydrophilic loop separating Helixes III and IV, while the fourth is located at the C-terminus of the protein. Again, as with the analogous regions in CPa-1, these domains represent portions of CPa-2 which can apparently tolerate a substantial amount of amino acid variability.

As with CPa-1, hydropathy plots of CPa-2 can be interpreted to suggest either seven (Alt et al. 1984) or six (Chisholm and Williams 1988) membrane-spanning helixes. It appears that the second helix predicted by Alt et al. (1984) is probably too short and too hydrophilic to span the thylakoid

```
              1                                                                                                            100
         6803 MKTLSSLRRF SPVVTLSN.T SMVGGRDLPS TGFAWWSGNA RLINLSGKLL GAHVRHAGLI VFWAGAMTLF EVAHFIPEKP MYEQGLILLP HIATLGWGVG
         7942                MVTLSSPS VIAGGRDIDS TGYAWWSGNA RLINLSGKLL GAHVAHAGLI VFWAGAMTLF EVAHFVPEKP MYEQGIILLS HLATLGWGVG
      EUGLENA                                          SGNA RLINVSGKLL GAHVAHAGLI VFWTGAMNLF EVAHFVPQKP MYEQGLILLP HLATLGFGVG
CHLAMYDOMONAS                METLFNGT LTVGGRDQET TGFAWWSGNA RLINLSGKLL GAHVAHAGLI VFWAGAMNLF EVSHFVPEKP MYEQGLILLP HIATLGYGVG
   MARCHANTIA MKILYSQRRF YPVETLFNGL LALGGRDQET TGFAWWAGNA RLINLSGKLL GAHVAHAGLI VFWAGAMNLF EVAHFVPEKP MYEQGLILLP HLATLGWGVG
      SPINACH MKTLYSLRRF YPVETLFNGT LTLAGRDQET TGFAWWAGNA RLINLSGKLL GAHVAHAGLI VFWAGAMNLF EVAHFVPEKP MYEQGLILLP HLATLGWGVG
      TOBACCO MKTLYSLRRF YHVETLFNGT LALAGRDQET TGFAWWAGNA RLINLSGKLL GAHVAHAGLI VFWAGAMNLF EVAHFVPEKP MYEQGLILLP HLATLGWGVG
          PEA MKTLYSLRRF YHVETLFNGT LALTGRDQET TGFAWWAGNA RLINLSGKLL GAHVAHAGLI VFWAGAMNLF EVAHFVPEKP MYEQGLILLP HLATLGWGVG
        WHEAT MKILYSLRRF YHVETLFNGT FVLAGRDQET TGFAWWAGNA RLINLSGKLL GAHVAHAGLI VFWAGAMNLF EVAHFVPEKP MYEQGLILLP HLATLGWGVG
        MAIZE MKILYSLRRF YHVETLFNGT FVLAGRDQET TGFAWWAGNA RLINLSGKLL GAHVAHAGLI VFWAGAMNLF EVAHFVPEKP MYEQGLILLP HLATLGWGVG
    CONSENSUS ---------- ----TL---- ----GRD--- TG-AWW-GNA RLIN-SGKLL GAHV-HAGLI VFW-GAM-LF EV-HF-P-KP MYEQG-ILL- H-ATLG-GVG
                                                                          I

              101                                                                                                          200
         6803 PAGEVTDIFP FFVVGVLHLI SSAVLGLGGI YHALRGPEVL EEYSSFFGYD WKDKNQMTNI IGYHLILLGC GALLLVFKAM FFGGVYDTWA PGGGDVRVIT
         7942 PGGEVVDTFP YFVVGVLHLI SSAVLGLGGI YHALRGPESL EEYSTFFSQD WKDKNQMTNI IGYHLILLGL GAFLLVFKAM FFGGVYDTWA PGGGDVRIIS
      EUGLENA FDGIVLDTYP FFVCGVLHLI SSAVLGFGGV
CHLAMYDOMONAS PGGEIIDTFP YFVSGVLHLI SSAVLGFGGV YHSLIGPETL EESYPFFGYV WKDKNKMTNI LGYHLIMLGL GAWLLVWKAM YFGGVYDTWA PGGGDVRVIT
   MARCHANTIA PGGEIVDTFP YFVSGVLHLI SSAVLGFGGI YHALIGPETL EESFPFFGYV WKDKNKMTTI LGIHLILLGA GAFLLVFKAL YFGGIYDTWA PGGGDVRKIT
      SPINACH PGGEVIDTFP YFVSGVLHLI SSAVLGFGGI YHALLGPETL EESFPFFGYV WKDRNKMTTI LGIHLILLGI GAFLLVFKAL YFGGVYDTWA PGGGDVRKIT
      TOBACCO PGGEVIDTFP YFVSGVLHLI SSAVLGFGGI YHALLGPETL EESFPFFGYV WKDRNKMTTI LGIHLILLGL GAFLLVFKAL YFGGVYDTWA PGGGDVRKIT
          PEA PGGEVIDTFP YFVSGVLHLI SSAVLGFGGI YHALLGPETL EESFPFFGYV WKDRNKMTTI LGIHLILLGI GSFLLVFKAF YFGGIYDTWA PGGGDVRKIT
        WHEAT PGGEVLDTFP YFVSGVLHLI SSAVLGFGGI YHALLGPETL EESFPFFGYV WKDRNKMTTI LGIHLILFGL GAFLLVLKAL YFGGVYDTWA PGGGDVRKIT
        MAIZE PGGEVLDTFP YFVSGVLHLI SSAVLGFGGI YHALLGPETL EESFPFFGYV WKDRNKMTTI LGIHLILLGL GAFLLVLKAL YFGGVYDTWA PGGGDVRKIT
    CONSENSUS --G---D--P -FV-GVLHLI SSAVLG-GG- YH-L-GPE-L EE---FF--- WKD-N-MT-I -G-HLI--G- G--LLV-KA- -FGG-YDTWA PGGGDVR-I-
                                  II                                                 III

              201                                                                                                          300
         6803 NPTLNPAIIF GYLLKAPFGG EGWIISVNNM EDIIGGHIWI GLICISGGIW HILTKPFGWA RRALIWSGEA YLSYSLGALS LMGFIASVFV WFNNTAYPSE
         7942 NPTLNPAVIF GYLLKSPFGG DGWIVSVDNL EDVIGGHIWI GLICISGGIW HILTKPFGWV GRAFIWNGEA YLSYSLGALS LMGFIASTMV WYNNTVYPSE
CHLAMYDOMONAS NPTTNAAVIF GYLVKSPFGG DGWICSVDNM EDIIGGHIWI GTLEILGGIW HIYTTPWPWA RRAFVWSGEA YLSYSLGAIG VMGFIACCMS WFNNTAYPSE
   MARCHANTIA NLTLSPGVIF GYLLKSPFGG EGWIVSVDNL EDIIGGHVWL GSICIFGGIW HILTKPFAWA RRALVWSGEA YLSYSLGAIA VFGFIACCFV WFNNTAYPSE
      SPINACH NVTLSPSIIF GCLLKSPFGG EGWIVSVDDL EDIIGGHVWI GVICILGGIW HILTKPFAWA RRALVWSGEA YLSYSLAALS VFGFIACCFV WFNNTAYPSE
      TOBACCO NLTLSPSIIF GYLLKSPFGG EGWIVSVDDL EDIIGGHVWL GSICILGGIW HILTKPFAWA RRALVWSGEA YLSYSLGALS VFGFIACCFV WFNNTAYPSE
          PEA NFTLSPSILF GYLLKSPFGG EGWIVSVDDL EDIIGGHVWL GSICILGGIW HILTKPFAWA RRALVWSGEA YLSYSLGALA VFGFIACCFV WFNNTAYPSE
        WHEAT NLTLSPSVIF GYLLKSPFGG EGWIVSVDDL EDIIGGHVWL GFICVFGGIW HILTKPFAWA RRAFVWSGEA YLSYSLAALS VFGFIACCFV WFNNRAYPSE
        MAIZE NLTLSPGVIF GYLLKSPFGG EGWIVSVDDL EDIIGGHVWL GSICVLGGIW HILTKPFAWA RRAFVWSGEA YLSYSLGALS VFGFIACCFV WFNNTAYPSE
    CONSENSUS N-T------F G-L-K-PFGG -GWI-SV--- ED-IGGH-W- G-----GGIW HI-T-P--W- -RA--W-GEA YLSYSL-A-- --GFIA---- W-NN--YPSE
                                                       IV                                          V

              301                                                                                                          400
         6803 FYGPTGMEAS QSQAFTFLVR DQRLGANIAS AQGPTGLGKY LMRSPSGEII FGGETMRFWD FRGPWLEPLR GPNGLDLDKL RNDIQPWQVR RAAEYMTHAP
         7942 FFGPTAAEAS QSQAFTFLVR DQRLGANIGS AQGPTGLGKY LMRSPTGEII FGGETMRFWD FRGPCVEPLR GPNGLDLDKL TNDIQPWQAR RAAEYMTHAP
CHLAMYDOMONAS FYGPTGPEAS QSQAFTFLVR DQRLGANVAS AQGPTGLGKY LMRSPTGEII FGGETMRFWD FRGPWLEPLR GPNGLDLNKL KNDIQPWQER RAAEYMTHAP
   MARCHANTIA FYGPTGPEAS QAQAFTFLVR DQRLGANVGS AQGPTGLGKY IMRSPTGEII FGGETMRFWD LRAPWLEPLR GPNGLDLSKL KKDIQPWQER RSAEYMTHAP
      SPINACH FYGPTGPEAS QAQAFTFLVR DQRLGANVGS AQGPTGLGKY LMRSPTGEVI FGGETMRFWD LRAPWLEPLR GPNGLDLSRL KKDIQPWQER RSAEYMTHAP
      TOBACCO FYGPTGPEAS QAQAFTFLVR DQRLGANVGS AQGPTGLGKY LMRSPTGEVI FGGETMRFWD LRAPWLEPLR GPNGLDLSRL KKDIQPWQER RSAEYMTHAP
          PEA FYGPTGPEAS QAQAFTFLVR DQRLGANVGS AQGPTGLGKY LMRSPTGEVI FGGETMRFWD LRAPWLEPLR GPNGLDLSRL KKDIQPWQER RSAEYMTHAP
        WHEAT FYGPTGPEAS QAQAFTFLVR DQRLGANVGS AQGPTGLGKY LMRSPTGEVI FGGETMRFWD LRAPWLEPLR GPNGLDLSRL KKDIQPWQER RSAEYMTHAP
        MAIZE FYGPTGPEAS QAQAFTFLVR DQRLGANVGS AQGPTGLGKY LMRSPTGEVI FGGETMRFWD LRAPWLEPLR GPNGLDLSRL KKDIQPWQER RSAEYMTHAP
    CONSENSUS F-GPT--EAS Q-QAFTFLVR DQRLGAN--S AQGPTGLGKY -MRSP-GE-I FGGETMRFWD -R-P--EPLR GPNGLDL--L --DIQPWQ-R R-AEYMTHAP

              401                                                                   473
         6803 LGSLNSVGGV ITDVNSFNYV SPRAWLATSH FVLGFFFLVG HLWHAGRARA AAAGFEKGID RETEPTLFMP DLD
         7942 LGSLNSVGGV ATEINSVNFV SPRAWLATSP FVLAFFFLVG HLWHAGRARA AAAGFEKGID RATEPVLAMR DLD
CHLAMYDOMONAS LGSLNSVGGV ATEINAVNFV SPRSWLACSH FCLGFFFFIG HLWHAGRARA AAAGFEKGID RFDEPVLSMR PLD
   MARCHANTIA LGSLNSVGGV ATEINAVNYV SPRSWLATSH FVLGFFFFVG HLWHAGRARA AAAGFEKGID RDFEPVLSMT PLN
      SPINACH LGSLNSVGGV ATEINAVNYV SPRSCLSTSH FVLGFFLFVG HLWHAGRARA AAAGFEKGID RDFEPVLSMT PLN
      TOBACCO LGSLNSVGGV ATEINAVNYV SPRSWLSTSH FVLGFFFFVG HLWHAGRARA AAAGFEKGID RDFEPVLSMT PLN
          PEA LGSLNSVGGV ATEINAVNYV SPRSWLATSH FVLGFFLFVG HLWHAGRARA AAAGFEKGID RDFEPVLSMT PLN
        WHEAT LGSLNSVGGV ATEINAVNYV SPRSWLSTSH FVLGFFFFVG HLWHAGRARA AAAGFEKGID RDLEPVLYMN PLN
        MAIZE LGSLNSVGGV ATEINAVNYV SPRSWLATSH FVLGFFFFVG HLWHAGRARA AAAGFEKGID RDLEPVLYMT PLN
    CONSENSUS LGSLNSVGGV -T--NAVN-V SPR--L--SH F-LGFF---G HLWHAGRARA AAAGFEKGID R--EP-L-M- -L-
                                             VI
```

Fig. 3. Comparison of the derived amino acid sequences of CPa-2 from nine species. A consensus sequence is shown; amino acid residues which are not conserved in all species are indicated by a dashed line (-). Predicted location of putative membrane-spanning helixes are shown by underlining and are indicated with Roman numerals (I–VI). Location of the alternative start codon with respect to these amino acid sequences is shown by an arrow

membrane. In spinach the six predicted membrane-spanning alpha helixes appear to be located in approximately the following positions: Helix I, Leu^{49}–Phe^{70}, Helix II, Phe^{112}–Leu^{135}, Helix III, Ile^{160}–Phe^{182}, Helix IV, Ile^{233}–Leu^{253}, Helix V, Leu^{272}–Phe^{292}, and Helix VI, Leu^{426}–Gly^{446} (Fig. 4). Five hydrophilic loops connect these predicted membrane-spanning domains. These loops are located in the following positions: Loop A, Glu^{71}–Tyr^{111}, Loop B, Gly^{136}–Thr^{159}, Loop C, Gly^{183}–Asp^{232}, Loop D, Lys^{254}–Tyr^{271}, and Loop E, Asn^{293}–Arg^{425}. Analysis of CPa-2 by the method of von Heijne and Gavel (1988) suggests that Loops A and C, which are depleted in lysyl and arginyl

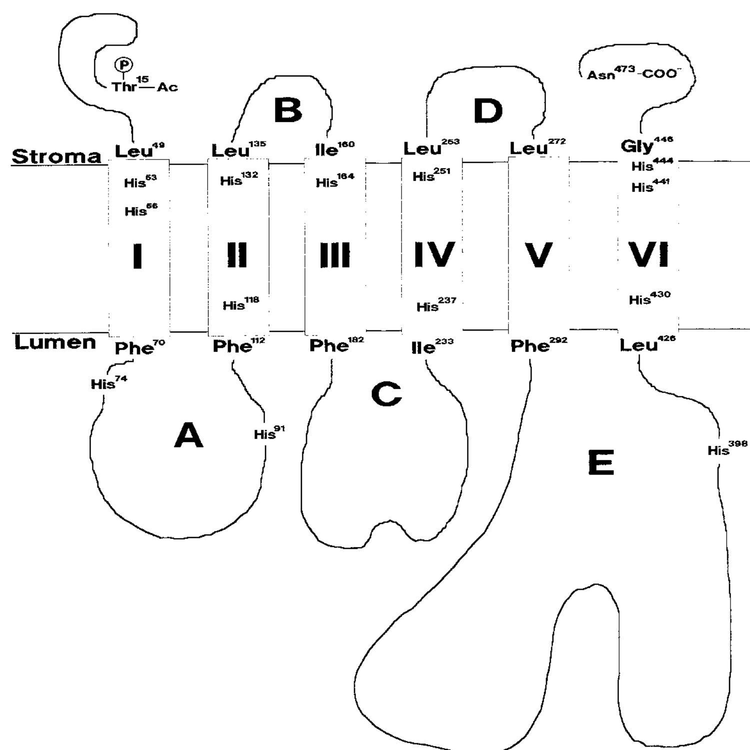

Fig. 4. Predicted membrane folding pattern of spinach CPa-2. Membrane spanning helixes are indicated by Roman numrals I–VI. The extrinsic loop regions A–E are also indicated. the approximate locations of the histidyl residues are shown. His^{430} is not conserved in *Synechococcus*

residues, are located on the lumenal side of the thylakoid membrane. Loops B and D, as well as the hydrophilic N- (Thr^{15}–Lys^{48}) and C-termini (Arg^{447}–Asn^{473}), are predicted to be exposed at the lumenal surface of the membrane. The large extrinsic loop of CPa-2 appears, based on the above analysis, to be located on the lumenal side of the thylakoid membrane.

In their role as internal chlorophyll-a antennae for PS II, CPa-1 and CPa-2 may interact with each other, the light-harvesting pigment-protein complex (the phycobilisomes in cyanobacteria and the light-harvesting chlorophyll-*a*/*b* protein in higher plants and green algae), other chlorophyll-proteins, such as CP29 (Camm and Green 1989), which appear to be associated with the PS II core, and the reaction center of the photosystem. Since excitation energy apparently flows from the light-harvesting pigment beds through CPa-1 and/or CPa-2, these proteins almost certainly interact structurally with the proteins of the light-harvesting pigment beds. In principal, comparison of the primary sequence of the *psb*B and *psb*C genes from the cyanobacterial species with those from higher plants and green algae could suggest possible sites for such structural interactions. This is due to the large structural differences between the cyanobacterial and higher plant (and green algal) light-harvesting pigment-protein complexes. Since there are so few *psb*B sequences available, such analysis is currently impossible for this protein. Comparison of the *psb*C sequences, however, suggests some possible sites for CPa-2-phycobilisome interaction. Specifically, when examined for amino acid residues which are conserved in the cyanobacterial species but which are non-conservatively replaced in higher plants and green algae, 14 such residues are identified. Eight of these residues are located on predicted stromally exposed hydrophilic regions. Interestingly, four of these residues (Arg^{135}, Tyr^{143}, Asp^{150}, Gln^{156} in the cyanobacteria; Ile^{135}, Ser^{143}, Val^{150}, Lys^{156} in green algae and higher plants) lie in Loop B, which is only 25 residues in length. I suggest that these, and possibly the other

stromally exposed residues, may be sites of interaction between CPa-2 and the phycobilisomes. Obviously, it will be interesting to modify these residues by use of site-directed mutagenesis techniques. It is interesting to note in this regard that Carpenter and Vermaas (1988) have replaced 75% of the 3′ end of the *Synechocystis psb*C gene with the homologous spinach gene sequence. The chimeric mutant generated completely lacks CPa-2 and can only grow photoheterotrophically. Transformation of this chimeric mutant with a plasmid containing the 3′ half of the cyanobacterial *psb*C gene restores photoautotrophic growth. These results imply that either the above sequence analysis must be substantially modified or that factors which complicate the interpretation of the experiments of Carpenter and Vermaas are occurring.

A number of investigators previously observed a phosphorylated protein with an apparent molecular mass of between 40 and 45 kDa (Millner et al. 1986). This protein was shown by Ikeuchi et al. (1987) to be CPa-2. In a very elegant study, Michel et al. (1988) demonstrated that CPa-2 undergoes a number of post-translational modifications including an endoproteolytic cleavage between Glu^{14} (or, if the alternative start codon is correct, Glu^{2}) and Thr^{15}, acetylation of Thr^{15} and phosphorylation of Thr^{15}. These authors also suggest that there may exist a second weakly phosphorylated site on this protein. The enzymology of these post-translational modifications as well as their physiological significance remain obscure. CPa-2 (as well as D1 and D2) appears to be phosphorylated by a different mechanism from that operating on the light-harvesting chlorophyll-*a*/*b* protein. Mutants, bearing alterations in the cytochrome b_6f complex, of maize and *Lemna* which cannot phosphorylate the light-harvesting chlorophyll-*a*/*b* protein are still capable of phosphorylating the proteins of the PS II core including CPa-2 (Gal et al. 1987, Coughlan, 1988, Bennett et al. 1988). Since Bennett (1980) had shown previously that the phosphorylation site on the 45 kDa phosphoprotein could be removed by trypsin treatment of thylakoid membranes, it is highly likely that the N-terminus of CPa-2 protrudes from the stromal side of the thylakoid membrane. It should be noted that this study, as well as analysis of the derived amino acid sequence of CPa-2 by the method of von Heijne and Gavel (1988) (see above), cast doubt on the accuracy of models which place the N-terminus of CPa-2 to the lumenal side of the thylakoid membrane (Murphy 1986).

3. The Function of CPa-1 and CPa-2

The primary function of both CPa-1 and CPa-2 appears to be that of interior chlorophyll-a light-harvesting antennae for PS II. Early work suggested that CPa-1 bound the primary donor of PS II, P_{680} (Nakatani et al. 1984), and associated cofactors. This hypothesis has fallen into almost universal disfavor. Excitation energy which is harvested by the light-harvesting pigment-protein complex is funneled through CPa-2 and CPa-1 and transferred to the putative reaction center complex, D1–D2-cytochrome b_{559}. In their apparent role of light-harvesting chlorophyll-proteins, CPa-1 and CPa-2 probably interact functionally with the light-harvesting pigment-protein complexes, components of the reaction center core complex, and with each other.

When excited with 420–440 nm light (which is primarily absorbed by chlorophyll) at 77 K, cyanobacteria, green plants and algae exhibit three major fluorescence emission bands which arise from the pigment beds of PS II and PS I. These emissions are generally observed at 685, 695 and 720–740 nm. The 685 and 695 nm peaks arise predominantly from PS II while the 720–740 nm peak arises predominantly from PS I (Krause and Weis 1984). Some contribution of PS I to the shorter wavelength emissions has been noted (Mullet et al. 1980) as has a contribution of PS II to the longer wavelength emission band (Stahl et al. 1989). Isolated PS II preparations which contain CPa-1, CPa-2, D1, D2, and cytochrome b_{559} (and possibly, the *psb* I gene product) exhibit the 685 and 695 nm fluorescence emissions (Satoh et al. 1983, Bricker et al. 1985). The 77 °K fluorescence emission spectra of CPa-1 and CPa-2 after their isolation by 'mildly-denaturing' electrophoresis (Nakatani et al. 1984, Pakrasi et al. 1985) or of protein isolated by chromatographic procedures (van Dorssen et al. 1987) indicate that CPa-2 and CPa-1 exhibit major emission bands at 685 nm and 695 nm, respectively. Additionally, cartridge mutagenesis of the *psb*B and *psb*C genes in *Synechocystis* induces the loss of

the 695 nm and 685 nm fluorescence emission peaks, respectively (Vermaas et al. 1986, Carpenter and Vermaas 1988). These studies suggest that CPa-1 and CPa-2 are responsible, at least in part, for the 685 nm and 695 nm fluorescence emission peaks associated with PS II. Experiments with oriented PS II membranes (Tapie et al. 1984) and oriented isolated CPa-1 (van Dorssen et al. 1987) indicate that the 695 nm emission originates from an emitting species whose dipoles are oriented perpendicular to the plane of the membrane. Breton (1982) had suggested that the emitting species was the pheophytin molecule which acts as the primary acceptor in PS II. Recent evidence, including the isolation of the D1–D2-cytochrome b_{559} complex, argues against this view. A recent study suggests that the 695 nm emission arises from a monomeric chlorophyll-a molecule (van Dorssen et al. 1987) associated with CPa-1. As mentioned previously, de Vitry et al. (1984) have determined that 20–25 chlorophyll molecules are associated with each CPa-1 and CPa-2 holoprotein. After electrophoresis (Delepelaire and Chua 1979) or isolation with octyl glucoside solubilization of a PS II core complex (Tang and Satoh 1984), 4–7 chlorophylls remain associated with these proteins. This suggests that there may be two pools of chlorophyll associated with these proteins, a weakly bound pool which appears to be released during isolation and a strongly bound pool of chlorophyll which is more difficult to disassociate. Since the 695 nm and 685 nm fluorescence emission peaks can be observed after electrophoresis of CPa-1 and CPa-2 (Nakatani et al. 1984, Pakrasi et al. 1985), it would appear that the emitting pigment molecules are associated with this strongly bound pool of chlorophyll. Interestingly, recent resonance Raman data suggest that CPa-1 (and the D1–D2-cytochrome b_{559} complex) binds chlorophylls in two different conformations: one in which the $C_9 = O$ group is relatively free and a second in which the $C_9 = O$ group interacts with protein or other chromophores (de Paula et al. 1990). It should be noted that the D1–D2-cytochrome b_{559} complex exhibits a 77 °K fluorescence emission at 684–685 nm (van Dorssen et al. 1987; Gounaris et al. 1989) which is oriented parallel to the plane of the membrane (van Dorssen et al. 1987). It has been suggested that this fluorescence emission may arise from the back reaction between P^{+}_{680} and Pheophytin^{-} (Barber et al. 1987), although other recombinational pathways leading to this fluorescence emission may be operating in this system (Eckert et al. 1987).

A number of investigators have observed that CPa-2 can be removed from the PS II core complex by treatment with either chaotropic agents such as potassium thiocyanate (Yamaguchi et al. 1988) or lithium perchlorate (Ghanotakis et al. 1989) or by additional detergent treatments (Yamagishi and Katoh 1985, Akabori et al. 1988). These treatments yield CPa-2 and a CPa-1–D1–D2-cytochrome b_{559} complex. Interestingly, removal of CPa-2 appears to destabilize the association of plastoquinone from the CPa-1–D1–D2-cytochrome b_{559} complex (Decker, Bowlby, and Yocum personal communication). The ability to remove CPa-2 from PS II preparations without the removal of CPa-1 and without affecting some partial chain PS II electron transport activities implies that CPa-2 is less closely associated with the D1–D2-cytochrome b_{559} reaction center than is CPa-1. Studies involving the selective solubilization of components of the PS II antennae seem to indicate that CPa-2 is required for the binding of LHCP to the CPa-1–D1–D2-cytochrome b_{559} complex and the transduction of excitation energy from the peripheral antennae to the core of PS II (Bassie et al. 1987). Additionally, these authors suggest that magnesium ions appear to stabilize the association of LHCP to CPa-2 and CPa-2 to CPa-1.

In addition to its role as a light-harvesting chlorophyll-protein, a number of lines of evidence suggest that CPa-1 may perform other functions in PS II. CPa-1 appears to be closely associated with the extrinsic 33 kDa manganese-stabilizing protein. The presence of the extrinsic 33 kDa protein shields CPa-1 (and CPa-2) from tryptic attack (Bricker and Frankel 1987) and prevents the labeling of amino groups on CPa-1 by the low molecular weight, amino group-specific reagent N-hydrosuccinimidobiotin (NHS-biotin) (Bricker et al. 1988). This reagent appears to primarily label lysyl residues located in the extrinsic loop region of CPa-1 (Frankel and Bricker, unpublished observations). Protein crosslinking data from three independent groups indicate the formation of crosslinked products between CPa-1 and the manganese stabilizing protein in both PS II membrane preparations (Bricker et al. 1988, Camm EL personal com-

munication, Enami et al. 1989) and in oxygen-evolving PS II core complex (Enami et al. 1987, Bricker et al. 1988). Three different crosslinking reagents have been used in these studies: Dithio*bis*(succinimidylpropionate) (DTSP) (Enami et al. 1987, Bricker et al. 1988), the water-soluble carbodiimide, 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide (EDC) (Bricker et al. 1988), and 2-iminothiolane (Camm EL personal communication). It should be noted that proteins must be essentially within van der Waals contact to be crosslinked with EDC and such proteins are usually thought to be interacting via a charge-pair mechanism (Hackett and Strittmatter 1984). Enami et al. (1989) have recently shown that oxygen evolution is stabilized after DTSP treatments which crosslink the 33 kDa extrinsic protein with CPa-1. Urea/NaCl washes which, in control experiments, remove the extrinsic 33 kDa protein and inhibit oxygen evolution, failed to remove the 33 kDa extrinsic protein which has been crosslinked to CPa-1. In these experiments, oxygen evolution was stabilized proportionally to the amount of 33 kDa protein which was crosslinked. Alkaline pH has been shown to effectively inhibit oxygen evolution and induces the release of the bound manganese associated with the oxygen-evolving site. Both oxygen evolution and manganese content were stabilized against the effects of alkaline pH proportionally to the amount of 33 kDa protein which was crosslinked. Interestingly, neither manganese nor oxygen evolution was stabilized against treatment with 0.8 M Tris although the crosslinked 33 kDa protein was stabilized to this treatment. These crosslinking data may suggest that CPa-1 provides *a binding site* for the extrinsic 33 kDa protein. Additional binding sites may be present on the D1-D2-cytochrome b_{559} complex (Gounaris et al. 1988).

In addition, my group has isolated a monoclonal antibody (designated FAC2) which recognizes CPa-1 both on 'Western blots' and on oxygen-evolving membranes (Bricker and Frankel 1987) when assayed by ELISA techniques. Interestingly, in the membrane environment, the antigenic determinant of the monoclonal antibody is accessible only upon removal of the strongly bound pool of manganese associated with the oxygen-evolving complex. The antigenic determinant for FAC2 has recently been mapped and is located between Pro^{360} and Ser^{391} in spinach (see Fig. 4), which is approximately in the center of the large extrinsic Loop E (Frankel and Bricker, in press). This data suggests that the removal of the strongly bound manganese from the oxygen-evolving site induces conformational changes in the proteins associated with the oxygen-evolving site which lead to the exposure of the antigenic determinant of the monoclonal antibody on CPa-1. The nature of this hypothesized conformational change has not yet been determined.

It is possible that CPa-1 is required for the functional assembly of PS II. In experiments examining alterations of the PS II structure and function after mutagenesis of CPa-1 in *Synechocystis*, Vermaas et al. (1988) concluded that CPa-1 was required for the assembly of a functional PS II complex. Interruption of the *psb*B gene with a kanamycin resistance cartridge leads to a PS II^- phenotype with no detectable amounts of CPa-1, D1 or D2 found in the photosynthetic membrane. Two chimeric mutants containing portions of the *psb*B gene from spinach, however, while exhibiting a PS II^- phenotype do accumulate measurable levels of D1 and D2. These authors hypothesize that the small amount of chimeric CPa-1 which may be synthesized in these mutants (but which rapidly turns over) is sufficient to stabilize small amounts of D1 and D2 in the membrane. It is unclear at this time how these results correlate to the findings of Callahan et al. (1990) which indicate that in higher plant thylakoid membranes CPa-1 and CPa-2 are located solely in the granal lamellae while D1 and D2 are also found in the stromal lamellae. These authors suggest that D1 and D2 are inserted into the stromal lamellae, processed, and imported into the granal lamellae (Mattoo and Edelman 1987, Callahan et al. 1987) where they become associated with CPa-1 and CPa-2 (Callahan et al. 1990).

While CPa-2 does not appear to be as closely associated with the PS II reaction center as is CPa-1, interruption mutagenesis of the *psb*C gene in *Synechocystis* does diminish the amounts of the other intrinsic PS II proteins to low levels. Interestingly, in these mutants electron transport from diphenylcarbazide to artificial electron acceptors still occurs, indicating that the PS II core is still functional. The water oxidation site and the reducing side of the photosystem are damaged, however, since these mutants apparently no longer

evolve oxygen nor do they exhibit binding of ^{14}C-diuron (Carpenter and Vermaas 1988).

4. Future directions

It is obvious from the above discussion that our understanding of both the structural and functional properties of CPa-1 and CPa-2 is in a state of flux. While we have primary structure information for these proteins, their organization within the membrane and relationship to the other PS II components remain unclear. While it is apparent that these proteins are interior chlorophyll-a antennae for PS II, other functional roles for these proteins have only recently been suggested. At least three lines of experimentation are required to clarify the properties of these proteins. First, the absolute topology of CPa-1 and CPa-2 must be determined. This information would allow the formation of testable hypotheses concerning the identification of domains on these proteins which interact with other PS II components. Second, *in vitro* mutagenesis of these proteins in cyanobacterial systems will provide important information concerning functionally important regions on these proteins. Finally, studies reconstituting purified CPa-1 and CPa-2 with the reaction center of PS II and with the light-harvesting pigment-protein complexes will clarify the roles for these proteins in the stabilization of PS II cofactors as well as allowing a more detailed examination of the mechanics of excitation energy transfer within the PS II antenna. It is hoped that in the next few years a more detailed understanding of the roles of these important PS II components will be forthcoming.

References

Akabori K, Tsukamoto H, Tsukihara J, Nagatsuka T, Motokawa O and Toyoshima Y (1988) Disintegration and reconstitution of Photosystem II reaction center core complex. I. Preparation and characterization of three different types of subcomplex. Biochim Biophys Acta 932: 345–357

Alt J, Morris J, Westhoff P and Herrmann RG (1984) Nucleotide sequence of the clustered genes for the 44 kd chlorophyll a apoprotein and the '32 kd'-like protein of the Photosystem II reaction center in the spinach plastid chromosome. Curr Genetics 8: 597–606

Amesz J (1983) The role of manganese in photosynthetic oxygen evolution. Biochim Biophys Acta 726: 1–12

Anderson JM, Waldron JC and Thorne SW (1978) chlorophyll-protein complexes of spinach and barley thylakoids. Spectral characteristics of six complexes resolved by an improved electrophoretic procedure. FEBS Lett 92: 227–233

Barber J, Chapman DJ and Tefler A (1987) Characterization of a PS II reaction center isolated from the chloroplasts of *Pisum sativum*. FEBS Lett 220: 67–73

Bassi R, Hoyer-Hansen G, Barbato R, Giacometti GM and Simpson, DJ (1987) Chlorophyll proteins of the Photosystem II antenna system. J Biol Chem 262: 13333–13341

Bennett J (1980) Chloroplast phosphoproteins. Evidence for a thylakoid-bound phosphoprotein phosphatase. Eur J Biochem 104: 85–89

Bennett J, Shaw EK and Michel H (1988) Cytochrome b_6f is required for phosphorylation of light-harvesting chlorophyll *a*/*b* complex II in chloroplast photosynthetic membranes. Eur J Biochem 171: 95–100

Bookjahns G, Stummann BM, Rasmussen OF and Henningsen KW (1986) Structure of a 3.2 kb region of the pea chloroplast DNA containing the gene for the 44 kD Photosystem II polypeptide. Plant Mol Biol 6: 359–366

Breton J (1982) The 695 nm fluorescence (F695) of chloroplasts at low temperature is emitted from the primary acceptor of photosystem 2. FEBS Lett 147: 16–20

Bricker TM, Metz JG, Miles D and Sherman LA (1983) Biochemical characterization of a highly active O_2-evolving Photosystem II preparation from maize. Biochim Biophys Acta 724: 447–455

Bricker TM, Pakrasi HB and Sherman LA (1985) Characterization of a spinach Photosystem II preparation isolated by a simplified method. Arch Biochem Biophys 237: 170–176

Bricker TM and Frankel LK (1987) Use of a monoclonal antibody in structural investigations of the 49 kDa polypeptide of Photosystem II. Arch Biochem Biophys 256: 295–301

Bricker TM, Odom WR and Queirolo CB (1988) Close association of the 33 kDa extrinsic protein with the apoprotein of CPa-1 in Photosystem II. FEBS Lett 231: 111–117

Camm EL and Green BR (1989) The chlorophyll *ab* complex, CP29, is associated with the Photosystem II reaction centre core. Biochim Biophys Acta 974: 180–184

Callahan FE, Edelman M and Mattoo AK (1987) Posttranslational acylation and intra-thylakoid translocation of specific chloroplast proteins. In: Biggins J (ed) Progress in Photosynthesis Research, pp. 799–802. Martinus Nijhoff Publishers, Dordrecht, The Netherlands

Callahan FE, Wergin WP, Nelson N, Edelman M and Mattoo AK (1989) Distribution of thylakoid proteins between stromal and granal lamellae in *Spirodela*: dual location of Photosystem II components. 91: 629–635 Plant Physiology

Carpenter SD and Vermaas WFJ (1988) *Synechocystis* 6803 mutants with a CP43 protein from spinach. In: Stevens SE and Bryant DA (eds) Light-energy transduction in photosynthesis: higher plant and bacterial models, pp 327–331. Rockvill, Maryland: American Society of Plant Physiologists

Chisholm D and Williams JGK (1988) Nucleotide sequence of *psb*C, the gene encoding the CP-43 chlorophyll a-binding protein of Photosystem II, in the cyanobacterium *Synechocystis* 6803. Plant Mol Biol 10: 293–301

Chua N-H and Benoun P (1975) Thylakoid membrane polypeptides of *Chlamydomonas reinhardtii*: wild type and mutant

strains deficient in photosystem II reaction center. Proc Natl Acad Sci (USA) 72: 2175–2179
Coughlan SJ (1988) Chloroplast thylakoid membrane phosphorylation is influenced by mutations in the cytochrome *bf* complex. Biochim Biophys Acta 933: 413–422
Deisenhofer J, Epp O, Miki K, Huber R and Michel H (1985) Structure of the protein subunits in the photosynthetic reaction center of *Rhodopseudomonas viridis* at 3Å resolution. Nature 318: 618–624
Delepelaire P and Chua N-H (1979) Lithium dodecyl sulfate/polyacrylamide gel electrophoresis of thylakoid membranes at 4°C: Characterization of two additional chlorophyll-*a* protein complexes. Proc Natl Acad Sci (USA) 76: 111–115
Dismukes GC (1986) The metal centers of the photosynthetic oxygen-evolving complex. Photochem Photobiol 43: 99–115
van Dorssen RJ, Breton J, Plitjter JJ, Satoh K, van Gorkom HJ and Amez J (1987) Spectroscopic properties of the reaction center and of the 47 kDa chlorophyll protein of Photosystem II. Biochim Biophys Acta 893: 267–274
Eckert H-J, Renger G, Bernarding J, Faust P, Eichler H-J and Salk J (1987) Examination of fluorescence lifetimes and radical-pair decay in Photosystem II membrane fragments from spinach. Biochim Biophys Acta 893: 208–218
Enami I, Satoh K and Katoh S (1987) Crosslinking between the 33 kDa extrinsic protein and the 47 kDa chlorophyll-carrying protein of the PS II reaction center core complex. FEBS Lett 226: 161–165
Enami I, Miyaoka T, Mochizuki Y, Shen J-R, Satoh K and Katoh S (1989) Nearest neighbor relationships among the constituent proteins of oxygen-evolving photosystem II membranes: binding and function of the extrinsic 33 kDa protein. Biochim Biophys Acta 973: 35–40
Ferenci T (1989) Is "hydrophobicity analysis" sufficient to predict topology of membrane proteins' Trends Biol Sci 14: 96
Frank HA, Hansson O and Mathis P (1989) EPR and optical changes of the Photosystem II reaction center produced by low temperature illumination. Photosyn Res 20: 279–289
Gal A, Shahak Y, Schuster G and Ohad I (1987) Specific loss of LHC II phosphorylation in the *Lemna* mutant 1073 lacking the cytochrome b_6/f complex. FEBS Lett 221: 205–210
Ghanotakis DF, Demetriou DM and Yocum CF (1987) Isolation and characterization of an oxygen-evolving Photosystem II reaction center core preparation and a 28 kDa Chl-*a*-binding protein. Biochim Biophys Acta 891: 15–21
Ghanotakis DF, de Paula JC, Demetriou DM, Bowlby NR, Petersen J, Babcock GT and Yocum CF (1989) Isolation and characterization of the 47 kDa protein and the D1–D2-cytochrome *b*-559 complex. Biochim biophys Acta 974: 44—53
Golden SS and Stearns GW (1988) Nucleotide sequence and transcript analysis of three Photosystem II genes from the cyanobacterium *Synechococcus* sp. PCC7942. Gene 67: 85–96
Gounaris K, Chapman DJ and Barber J (1988) The interaction between the 33 kDa manganese-stabilizing protein and the D1/D2 cytochrome *b*-559 complex. FEBS Lett 234: 374–378
Gounaris K, Chapman DJ and Barber J (1989) Isolation and characterization of a D1/D2/cytochrome *b*-559 complex from *Synechocystis* 6803. Biochim Biophys Acta 973: 296–301.
Green BR (1988) the chlorophyll-protein complexes of higher plant photosynthetic membranes or Just what green band is that? Photosyn Res 15: 3–32
Hackett CS and Strittmatter P (1984) Covalent crosslinking of the active sites of vesicle-bound cytochrome b_5 and NADH cytochrome b_5 reductase. J. Biol Chem 259: 3275–3282
Hayden DB and Hopkins WG (1977) A second chlorophyll-a protein complex in mesophyll chloroplasts. Can J Bot 55: 2525–2529
von Heijne G and Gavel Y (1988) Topogenic signals in integral membrane proteins. Eur J Biochem 174: 671–678
Henriques F and Park RB (1978) Characterization of three new chlorophyll-protein complexes. Biochem Biophys Res Commun 81: 1113–1118
Homann PH (1987) the relations between chloride, calcium, and polypeptide requirements of photosynthetic water oxidation. J Bioenerg Biomem 19: 105–123
Ikeuchi M, Plumley FG, Inoue Y and Schmidt GW (1987) Phosphorylation of Photosystem II components, CP43 apoprotein, D1, D2, and 10 to 11 kilodalton protein in chloroplast thylakoids of higher plants. Plant Physiol. 85: 638–642
Ikeuchi M and Inoue Y (1988) A new photosystem II reaction center component (4.8 kDa protein) encoded by the chloroplast genome. FEBS Lett 241: 99–104
Ikeuchi M, Tako K and Inoue Y (1989) N-terminal sequencing of Photosystem II low molecular mass proteins. FEBS Lett 242: 263–269
Krause GH and Weis E (1984) Chlorophyll fluorescence as a tool in plant physiology. II. Interpretation of fluorescence signals. Photosyn Res 5: 139–157
Krebbers ET (1983) Thesis, Harvard University
Mattoo AK and Edelman M (1987) Intramembrane translocation and post-translational palmitoylation of the chloroplast 32 kDa herbicide-binding protein. Proc Natl Acad Sci (USA) 84: 1497–1501
Michel H, Hunt DF, Shabanowitz J and Bennett J (1988) Tandem mass spectrometry reveals that three Photosystem II proteins of spinach chloroplasts contain N-acetyl-O-phosphothreonine at their NH_2 termini. J Biol Chem 263: 1123–1130
Millner PA, Marder JB, Gounaris K and Barber J (1986) Localization and identification of phosphoproteins within the PS II core of higher plant thylakoid membranes. Biochim Biophys Acta 852: 30–37
Montandon P-E and Stutz E (1986) *Euglena gracilis* chloroplast DNA: analysis of a 1.6 kb intron of the *psb* C gene containing an open reading frame of 458 codons. Curr Genet 11: 35–39
Morris J and Herrmann RG (1984) Nucleotide sequence of the gene for the P680 chlorophyll a apoprotein of the Photosystem II reaction center from spinach. Nucleic Acid Res 12: 2837–2850
Mullet JE and Arntzen CJ (1980) Simulation of grana stacking in a model membrane system: mediation by a purified light-harvesting pigment protein complex from cholorplasts. Biochim Biophys Acta 589: 100–117
Murphy DJ (1986) The molecular organization of the photosynthetic membranes of higher plants. Biochim Biophys Acta 864: 33–94
Nakatani HY, Ke B, Dolan E and Arntzen CJ (1984) Identity of the Photosystem II reaction center polypeptide. Biochim

Biophys Acta 765: 347–352
Nanba O and Satoh K (1987) Isolation of a Photosystem II reaction center consisting of D1 and D2 polypeptides and cytochrome *b*-559. Proc Natl Acad Sci USA 84: 109–112
Ohyama K, Kohchi T, Sano T, Sano S, Umesono K, Shiki Y, Takeuchi MM, Chang Z, Aota S, Inokuchi H and Ozeki H (1986) Chloroplast gene organization deduced from complete sequence of liverwort *Marchantia polymorpha* chloroplast DNA. Nature 322: 572–574
Pakrasi HB, Riethman HC and Sherman LA (1985) Organization of the pigment proteins of the Photosystem II complex of the cyanobacterium *Anacystis nidulans* R2. Proc Natl Acad Sci (USA) 82: 6903–6907
Paul C and Rosenbusch JP (1985) Folding patterns of porin and bacteriorhodopsin. EMBO J 4: 1593–1597
de Paula JC, Ghanotakis DF, Bowlby NR, Dekker JP, Yocum CF and Babcock GT (1990) Chlorophyll-protein interactions in Photosystem II. Resonance Raman spectroscopy of the D1–D2-cytochrome b_{559} complex and the 47 kDa protein. To appear in the proceedings of the VIIIth International Congress of Photosynthesis.
Piccioni R, Benoun P and Chua NH (1981) A nuclear mutant of *Chlamydomonas reinhardtii* defective in photosynthetic phosphorylation. Eur J Biochem 117: 93–102
Piccioni R, Bellemare R and Chua N-H (1982) Methods of polyacrylamide gel electrophoresis in the analysis and preparation of plant polypeptides. In: Edelman M, Hallik RB and Chua N-H (eds) Methods of Chloroplast Molecular Biology, pp 985–1014. Elsevier Biomedical Press, Amsterdam
Rochaix J-D, Kuchka M, Mayfield S, Schirmer-Rahire M, Girard-Bascou J and Bennoun P (1989) Nuclear and chloroplast mutations affect the synthesis or stability of the chloroplast *psbC* gene product in *Chlamydomonas reinhardtii* EMBO J 8: 1013–1021
Sayre RT, Andersson B and Bogorad L (1986) The topology of a membrane protein: the orientation of the 32 kD Qb-binding chloroplast thylakoid membrane protein. Cell 47: 601–608
Satoh K, Nakatani HY, Steinback KE, Watson J and Arntzen CJ (1983) Biochim Biophys Acta 724: 142–150
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, Shimada H and Sugiura M (1986) The complete nucleotide sequence of the tobacco chloroplast genome: its gene organization and expression. EMBO J 5: 2043–2049
Stahl U, Tusov VB, Paschenko VZ and Voigt J (1989) Spectroscopic investigations of fluorescence behavior, role and function of the long-wavelength pigments of Photosystem I. Biochim Biophys Acta 973: 198–204
Tang XS and Satoh K (1984) Characterization of a 47-kilodalton chlorophyll-binding polypeptide (CP-47) isolated from a Photosystem II core complex. Plant Cell Physiol 25: 935–945
Tapie P, Acker S, Arntzen CJ, Choquet Y, Delepelaire P, Diner BA, Wollman FA and Breton J (1984) Orientation of absorbing and fluorescencing dipoles in PS II particles. In Sybsma C (ed) Advances in Photosynthesis Research, Vol 1, pp 693–696. Nijhoff/Junk Publishers, Amsterdam
Tronrud DE, Schmidt MF and Matthews BW (1986) Structure and X-ray amino acid sequence of a bacteriochlorophyll-*a* protein from *Prosthecochloris aestuarii* refined at 1.9 Å resolution. J Mol Biol 188: 443–454
Vermaas WFJ, Williams JGK, Rutherford AW, Mathis P and Arntzen CJ (1986) Genetically engineered mutant of the cyanobacterium *Synechocystis* 6803 lacks the Photosystem II chlorophyll-binding protein CP-47. Proc Natl Acad Sci (USA) 83: 9474–9477
Vermaas WFJ, Williams, JGK and Arntzen CJ (1987) Sequencing and modification of *psb* B, the gene encoding the CP-47 protein of Photosystem II, in the cyanobacterium *Synechocystis* 6803. Plant Mol Biol 8: 317–326
Vermaas WFJ, Ikeuchi M and Inoue Y (1988) Protein composition of the Photosystem II core complex in genetically engineered mutants of the cyanobacterium *Synechocystis* sp. PCC 6803. Photosyn Res 17: 97–113
de Vitry C, Wollmann F-A and Delepelaire P (1984) Function of the polypeptides of the Photosystem II reaction center in *Chlamydomonas reinhardtii*. Biochim Biophys Acta 767: 415–422
Wechsler T, Brunisholz R, Suter F, Fuller RC and Zuber H (1985) Primary structures of the bacteriochlorophyll *c* and *a* binding polypeptides of the green thermophilic bacterium *Chloroflexus aurantiacus*. Abstract Vth International Symposium on Photosynthetic Procaryotes, p 296
Wessels JSC and Borchert MT (1978) Polypeptide profiles of chlorophyll-protein complexes and thylakoid membranes of spinach chloroplasts. Biochim Biophys Acta 503: 78–93
Yamagishi A and Katoh S (1985) Further characterization of the two Photosystem II reaction center preparations from the thermophilic cyanobacterium *Synechococcus* sp. Biochim Biophys Acta 807: 74–80
Yamaguchi N, Takahashi Y and Satoh K (1988) Isolation and characterization of a Photosystem II core complex depleted in the 43 kDa chlorophyll-binding subunit. Plant Cell Physiol 29: 123–129
Zuber H, Brunisholz R and Sidler W (1987) Structure and function of light-harvesting pigment-protein complexes. In Amesz J (ed) Photosynthesis pp 233–271. Amsterdam: Elsevier Press