Functionally Important Domains of the Large Hydrophilic Loop of CP47 As Probed by Oligonucleotide-Directed Mutagenesis in *Synechocystis* sp. PCC 6803[†]

Elisabeth Haag,*.1.8 Julian J. Eaton-Rye, 8.11 Gernot Renger, and Wim F. J. Vermaas 8

Department of Botany and the Center for the Study of Early Events in Photosynthesis, Arizona State University, Tempe, Arizona 85287, and Max Volmer Institute for Physical and Biophysical Chemistry, Technical University Berlin, 1000 Berlin 12, Germany

Received November 30, 1992; Revised Manuscript Received February 9, 1993

ABSTRACT: The chlorophyll a-binding protein CP47 serves as core antenna to photosystem II (PS II). The predicted topology of CP47 exhibits six membrane-spanning regions and a large hydrophilic loop (loop E) which roughly includes 200 residues (255-455) and is presumably exposed to the lumenal side of the thylakoid membrane. Several lines of experimental evidence suggest that loop E might be involved in binding or stabilizing functional manganese in the catalytic site of water oxidation or in interacting with the extrinsic PS II-O protein (the 33-kDa manganese-stabilizing protein). To scan loop E for functionally important domains, oligonucleotide-directed mutagenesis has been used to introduce deletions of 3-8 residues in conserved and charged regions of loop E. In addition, one single-site mutation of the only histidine present in loop E was created (H343L). Domains deleted in $\Delta 1$ (I265-F268), $\Delta 2$ (T271-K277), $\Delta 4$ (T304-L309). $\Delta 5$ (F311-N317), and $\Delta 12$ (D440-P447) are required for stable assembly of functional PS II complexes. Deletion of domains $\Delta 3$ (K277-E283) and $\Delta 11$ (R422-E428) significantly reduces the level of assembled PS II and impairs photoautotrophic growth and oxygen evolution. Deletion of domain $\Delta 8$ (A373–D380) enhances the susceptibility to photoinhibition. In contrast, deletion of domains Δ6 (G333-I336), Δ7 (K347-R352), $\Delta 9$ (V392–Q394), and $\Delta 10$ (D416–F420) and mutation of H343 to leucine do not seem to severly interrupt PS II structure and function, although all mutants exhibit a slightly decreased stability of PS II as compared to the wild type. Thus, selected domains of the large hydrophilic loop of CP47 are important for PS II structure and function. With respect to possible sites of interaction between loop E of CP47 and the extrinsic PS II-O protein, our results indicate that none of the deletions in the region from residue 330 to 420 ($\Delta 6$, $\Delta 7$, $\Delta 8$, $\Delta 9$, $\Delta 10$) completely interrupts a functional association of the manganese-stabilizing protein to PS II, although the binding characteristics might be changed in some cases.

In all oxygen-evolving photoautotrophic organisms, from cyanobacteria to higher plants, the oxidation of water to dioxygen is tightly coupled to the light-driven electron transport within photosystem II (PS II), a multimeric pigment-protein complex in the thylakoid membrane [for reviews, see Babcock (1987), Renger (1987), and Rutherford et al. (1992)]. In total, PS II is composed of more than a dozen different polypeptides [for a review, see Ikeuchi (1992)], only some of which can be related to specific functional subdivisions, e.g., reaction center, proximal core antenna, and regulatory extrinsic cap (for reviews, see Hansson and Wydrzynski (1990), and Vermaas and Ikeuchi (1991)]. In the current model of PS II, the chlorophyll a-binding proteins CP47 (the psbB gene product) and CP43 (the psbC gene product) are viewed to

constitute the core antenna engaged in transfer of excitation energy to the reaction center, which comprises two polypeptides, designated D1 and D2 (Trebst, 1986; Michel & Deisenhofer, 1988). However, strong evidence has accumulated over the past years that CP47 and CP43 contribute additional structural and functional roles to the operation of PS II (e.g., Bricker, 1990; Bricker et al., 1987, 1988; Vermaas et al., 1988; Rögner et al., 1991).

In terms of the overall organization and assembly of PS II, CP47 and CP43 are closely associated with the D1/D2 heterodimer. Any isolated PS II subcomplex that is capable of oxygen evolution contains both CP47 and CP43. A similarity in selected functions for CP47 and CP43 is further implied by a number of corresponding structural features. Both proteins are predicted to contain six transmembrane α -helices and to exhibit a large hydrophilic loop (loop E) between the fifth and sixth putative helices which is thought to be lumenally exposed (Vermaas et al., 1987; Bricker, 1990). However, in spite of the structural similarity, evidence for significant functional differences has accumulated. Biochemical studies reveal that CP47 is bound more tightly to the PS II core than is CP43 (Akabori et al., 1988; Ghanotakis et al., 1989), and deletion mutagenesis implies a more significant role for CP47 than CP43 for a stable assembly of PS II (Vermaas et al., 1986, 1988).

Several lines of evidence indicate an interaction between CP47 and the extrinsic PS II-O polypeptide [for reviews, see Bricker (1990) and Enami et al. (1992)]. This is generally interpreted to indicate a possible interaction between CP47 and the water-oxidizing system. Despite some recent findings

[†] This work has been supported by the U.S. Department of Energy (Grant DE-FG02-89ER14031). This is publication 134 from the Center for the Study of Early Events in Photosynthesis at Arizona State University. The Center is funded by the U.S. Department of Energy (Grant DE-FG02-88ER13969) as part of the Plant Science Centers Program of the U.S. Department of Agriculture/Department of Energy/National Science Foundation. E.H. acknowledges the support of the Deutsche Akademische Austauschdienst (DAAD). Financial support by the Deutsche Forschungsgemeinschaft (Re 354/11-2) is gratefully acknowledged.

^{*} Correspondence address: Technical University Berlin.

[‡] Technical University Berlin.

[§] Arizona State University.

Present address: Brookhaven National Laboratory, Upton, NY 11973.

¹ Abbreviations: Chl, chlorophyll; DMBQ, 2,5-dimethylbenzoquinone; HEPES, 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid; PMSF, phenylmethylsulfonyl fluoride; PS I, photosystem I; PS II, photosystem II; SDS, sodium dodecyl sulfate; PQ, plastoquinone.

implying a possible role for the reaction center polypeptides D1 and D2 for binding of manganese involved in water oxidation (Nixon & Diner, 1992; Vermaas et al., 1990b), the correct coordination sphere and the nature of the protein matrix of the manganese cluster remain largely obscure (Renger & Wydrzynski, 1991; Debus, 1992). Likewise, the precise role of the PS II-O protein for water oxidation is still a matter of debate, but it is known to be closely associated with the manganese-containing catalytic site of water oxidation. In PS II subcomplexes isolated from higher plants and under physiological Cl- concentrations, the PS II-O protein has proved to be neccessary for stabilization of at least two of the four manganese atoms involved in water oxidation (Miyao & Murata, 1984a,b; Mavankal et al., 1986). In contrast, recent genetic engineering in cyanobacteria has produced mutants which lack PS II-O but still are able to maintain photoautotrophic growth; however, these mutants exhibit a reduced oxygen evolution rate and an increased susceptibility toward photoinhibition (Burnap & Sherman, 1991; Philbrick et al., 1991; Mayes et al., 1991). On the other hand, genetic studies performed with the eukaryotic green alga Chlamydomonas reinhardtii favor the notion that the PS II-O protein is required for stability of the PS II complex in eukaryotes (Mayfield et al., 1987), and biochemical studies in PS II membrane fragments from higher plants reveal that depletion of PS II-O raises the Cl- requirement to nonphysiologically high concentrations above 100 mM (Miyao & Murata, 1984a; Bricker, 1992). Part of the apparent discrepancy between eukaryotic and prokaryotic systems might be explained by considering the slightly different organization of the donor side of PS II. In most eukaryotes two additional regulatory extrinsic proteins with apparent molecular weights of 24 kDa (PS II-P) and 18 kDa (PS II-Q) are present in addition to the PS II-O protein, whereas in cyanobacteria two other proteins might be additionally associated with the donor side of PS II (Shen et al., 1992).

Regardless of the detailed function of PS II-O, its close association with CP47 would imply a possible role of CP47 in the water-oxidizing system, i.e., binding or stabilization of a functional manganese cluster. In the past, evidence favoring a close interaction between CP47 and PS II-O has been provided by biochemical studies with PS II membrane fragments of higher plants. Extraction of PS II-O increases the accessibility of proteases, labeling reagents, and/or antibodies to CP47 (Bricker & Frankel, 1987; Bricker et al., 1988; Frankel & Bricker, 1990). Furthermore, a number of cross-linking experiments indicate close proximity between the PS II-O protein and CP47 (Enami et al., 1987, 1989, 1991; Bricker et al., 1988; Odom & Bricker, 1992). In summary, the available data support the view that the N-terminus of the extrinsic PS II-O protein interacts with a domain located on the C-terminal half of the large hydrophilic loop of CP47, i.e., residues 364-440 (Eaton-Rye & Murata, 1989; Frankel & Bricker, 1990; Odom & Bricker, 1992).

In the present study, the large lumen-exposed loop of CP47 has been scanned for functionally important domains by employing deletion mutagenesis in the cyanobacterium Synechocystis sp. PCC (Pasteur Culture Collection) 6803. The system used to manipulate the psbB gene of Synechocystis sp. PCC 6803 has been previously described, and results of its first application have demonstrated that the region between G351 and T365 is essential for the stable assembly of PS II, while the region from R384 to V392 is not (Eaton-Rye & Vermaas, 1991). In the current study, twelve short deletions (3-8 amino acids) containing conserved and charged residues and, in addition, one single-site mutation of the only histidine residue of loop E are presented. Together with the previous work (Eaton-Rye & Vermaas, 1991), the present paper provides a useful genetic approach to elucidate the functional role of domains in the extrinsic loop of CP47 under in vivo conditions.

MATERIALS AND METHODS

Mutant Construction and Verification. Oligonucleotidedirected deletion mutagenesis from the wild-type psbB cloned in M13mp19 was performed as described by Eaton-Rye & Vermaas (1991). Oligonucleotides 30 bases long were synthesized with each oligonucleotide hybridizing to bases up- and downstream of a particular region that was to be deleted. The oligonucleotide sequences are shown in Table I. Each oligonucleotide was designed such that a restriction site was created at the site of the deletion. In some cases silent base-pair changes had to be introduced into the codons flanking the deletion in order to obtain the desired restriction site.

After hybridization of the mutagenic oligonucleotide to a psbB template cloned into bacteriophage M13mp19, synthesis of the mutant psbB gene was completed in vitro (Eaton-Rye & Vermaas, 1991). Selective degradation of wild-type DNA and propagation of bacteriophage M13mp19 carrying the mutated psbB in Escherichia coli was performed according to Vermaas et al. (1990a). Subsequently, M13mp19 singlestranded DNA from single clones were sequenced, mutants were identified, and the appropriate mutated psbB fragment was excised and cloned into a plasmid providing complementary parts of psbB. Proper ligation into this plasmid also led to restoration of a complete kanamycin-resistance cartridge, thus providing convenient selection (Eaton-Rye & Vermaas, 1991). The resulting construct isolated from single E. coli clones was tested for the proper psbB sequence by restriction enzyme digestion prior to transformation of the photoheterotrophic psbB-strain of Synechocystis sp. PCC 6803 (Eaton-Rye & Vermaas, 1991). As a final check of the introduced mutation, genomic DNA of the transformants was isolated (Williams, 1988) and then amplified and sequenced, employing a double-stranded cycle sequencing system (BRL Life Technologies Inc.).

Cell Culture and Growth. Unless indicated otherwise, liquid cultures of wild-type and mutant strains were grown in regular BG-11 medium (Rippka et al., 1979) supplemented with 5 mM glucose. When cultures were kept on solid agar plates (1.5% w/v), 20 μM atrazine was added in order to suppress PS II activity and thus to suppress selection for secondary. spontaneous mutations leading to better photoautotrophic growth. In addition, all growth media for mutant strains contained 20 µg/mL kanamycin. Phototrophic growth was maintained at continuous illumination with 60 $\mu E/(m^2 s)$ at 30 °C. During the entire growth period the liquid cultures were ventilated with a low but constant stream of air.

Preparation of Thylakoids. Thylakoid membranes were prepared from cultures with an optical density of 0.7-0.9 at 730 nm according to Yu & Vermaas (1990). Cells of a 200mL culture were collected, washed, and subsequently disrupted at 4 °C using fine glass beads (100-150 μ m). After removal of glass beads and cell debris by centrifugation, the thylakoid membranes were pelleted and washed once. Thylakoid membranes were either used directly or stored in the presence of an additional 10% glycerol at -80 °C. Protease inhibitors were only added for thylakoids subjected to gel electrophoresis.

Chlorophyll Analysis. The chlorophyll (Chl) of cells was extracted in 100% methanol; the Chl of isolated thylakoids, in 80% acetone. The Chl a concentration was determined according to MacKinney (1941).

Table I: Oligonucleotide Sequences Used to Generate Site-Directed Deletions and Mutation

mutant			oligonucleotide sequence ^a	restriction endonuclease
Δ1	5' AAT	GCC ACC ACC	C CCC ▼ GGG CCC ACC CGT TAT 3'	SmaI
Δ2	5' GAA	CTC TTT GG	C CCC ▼ GGG TAT TTT CAA GAA 3'	SmaI
Δ3	5' CGT	TAT CAG TGO	G GAT ▼ ATC CAA CGC CGG GTT 3'	EcoRV
Δ4	5' TCT	GAA GCC TGC	G AGC ▼ GCT TTC TAC GAT TAC 3'	Eco47III
Δ 5	5' CCT	GAA AAG TTO	G GCT ▼ AGC CCC GCT AAA GGT 3'	NheI
Δ6	5' GGT	GCT ATG AAC	AGC ▼ GCT CAG GAA TGG ATT 3'	Eco47III
Δ 7	5' GGT	CAC CCC ATA	A TTC ▼ GAA CTG GAG GTA CGG 3'	BstBI
Δ8	51 GTC	ATC ATG ACC	GAT ▼ ATC CCC TTC CGT CGT 3'	EcoRV
Δ9	5' GAG	TCT AAA TTO	AGT ▼ ACT GGT GTT ACC GTC 3'	Scal
Δ 10	5' TTC	AGC AAT CCC	C AGC ▼ GCT CGG AAA GCT CAG 3'	Eco47III
Δ11	5' GTG	AAG AAG TT	I GCC ▼ GGC TTC GAC TTC GAT 3'	NaeI
Δ12	5' GAA	ACC TTC AAC	C TCG ▼ CGA GGT TGG TTT ACC 3'	NruI
H343L	5' GAA	TGG ATT GG	CTC CCC ATA TTC AAA GAC 3'	

^a The site of deletion is indicated by (▼); the resulting restriction site is underlined; single-base changes are indicated by italic letters.

Gel Electrophoresis and Immunoblot Analysis. SDS-polyacrylamide gel electrophoresis was performed according to Chua (1980) in the presence of 5 M urea. Samples were incubated in Laemmli buffer (Laemmli, 1970) containing 5% $(v/v)\beta$ -mercaptoethanol for 30 min at room temperature prior to application. In each case 7 μ g of Chl was applied to the gel. Immunoblot analysis with antibodies raised against PS II polypeptides D1, D2, CP43, and CP47 was carried out as described previously (Vermaas et al., 1988).

Herbicide-Binding Assay. The relative concentration of stable assembled PS II centers in wild-type and mutant cells was determined by measuring the binding of 14 C-labeled diuron on a Chl basis. Cells were collected at an optical density of 0.7–0.9 at 730 nm, washed in 25 mM HEPES/NaOH, pH 7.0, and resuspended to 50 μ M Chl. Aliquots were then incubated with different [14 C] diuron concentrations in dim light for 20–30 min at room temperature. To eliminate contributions of nonspecific diuron binding, the difference between samples assayed in the absence and presence of 20 μ M atrazine was evaluated. For details, see Vermaas et al. (1990b).

Oxygen Evolution. Comparative measurements of the oxygen evolution rate of wild-type and mutant strains were performed with cell cultures grown to an optical density of 0.7–0.9 at 730 nm. Oxygen evolution was detected on a Clark-type electrode. Measurements were performed in 25 mM HEPES/NaOH, pH 7.0, in the presence of 0.5 mM K_3 [Fe(CN)₆] and 0.1 mM DMBQ. Samples (10 μ g/mL Chl) were illuminated with orange light (570-nm cutoff filter) at a saturating intensity of 5000 μ E/(m²·s) at 30 °C.

77 K Fluorescence Emission Spectra. 77 K fluorescence emission spectra were recorded on a SPEX Fluorolog 2 spectrometer equipped with a glass dewar for liquid nitrogen. The samples (whole cells or thylakoid membranes) were frozen in 60% glycerol and 25 mM HEPES/NaOH, pH 7.0, and excited at 440 or 590 nm. In each case the Chl concentration of the sample was $5 \mu g/mL$. Spectra were corrected for the wavelength dependence of the sensitivity of the photodetector.

Prediction of Secondary Structure. The secondary structure of the large extrinsic loop of CP47 was predicted for

wild-type and mutant sequences using the Genetics Computer Group (GCG) software programs, version 7.0 (Devereux et al., 1984; Jameson & Wolf, 1988). The prediction evaluated in this study is based on the Garnier-Osguthorpe-Robson method (Garnier et al., 1978).

RESULTS

To determine regions of possible structural or functional importance in loop E of CP47, we evaluated sequence conservation within this loop. Figure 1 compiles available sequence data for the predicted large loop between transmembrane helices V and VI on the lumenal side of CP47 from different organisms, including cyanobacteria, algae, and higher plants, and shows the position of the deletions in the twelve deletion mutants which were created and characterized in this study. None of these deletions comprises more than eight amino acids. The locations are specifically chosen in order to cover predominantly conserved and negatively charged amino acid residues. The goal of this study is to determine if specific regions of this loop are of special relevance for the structural and functional integrity of PS II. To test this, the mutants were characterized by measuring the following properties: (a) photoautotrophic growth, (b) abundance of the main polypeptides and stability of the PS II complex, (c) oxygen evolution capacity, and (d) fluorescence emission spectra. As a control, all characterization procedures were performed with a wild-type and a kanamycin-resistant psbBreconstituted strain which does not carry a mutation in the psbB gene. As described previously, the phenotype of the psbB-reconstituted strain is indistinguishable from wild type except for the kanamycin resistance (Eaton-Rye & Vermaas, 1991). In the present study the data obtained for the wildtype and the psbB-reconstituted strain are referred to as wildtype control. In addition to the functional characterization, the predicted secondary structure of loop E was analyzed in all mutants in order to compare mutationally induced functional changes with possible effects on the protein

Photoautotrophic Growth. The capability for photoautotrophic growth was tested for all mutant strains in BG-11

Totacco		251	261	271	281	291	301	311
Spinach	Tobacco					•	AWSKIPEKLA	FYDYIGNNPA
NONITHING		VVAGTMWYGS	ATTPIELFGP	TRYQWDQGYF	QQEIYRRVSA	GLAENQSFSE	AWSKIPEKLA	FYDYIGNNPA
No.	-	VVAGTMWYGS	ATTPIELFGP	TRYQWDQGYF	QQEIYRRVGA	GLAKNQSLSE	AWSKIPEKLA	FYDYIGNNPA
Note	Maize	VVAGTMWYGS	ATTPIELFGP	TRYQWDQGYF	QQEIYRRVSD	GLAENLSLSE	AWSKIPEKLA	FYDYIGNNPA
Barley VVAGTMYGS ATFIELFGP TRYGWDGOFF QGEIYRRUSH GLARNSLSE AMSKIPEKLA FYDYIGNNPA Liverwort VVAGTMYGS AATFIELFGP TRYGWDGOFF QGEIGRRUGA SLAEGALSLS AMSKIPEKLA FYDYIGNNPA Chlamydon. VVAGTMYGS AATFIELFGP TRYGWDGOFF QGEIGRRUGA SLAEGALSLS AMSKIPEKLA FYDYIGNNPA FYDYIGNNPA FYDROCHIOPO. VVAGTMYGK AATFIELFGP TRYGWDGOFF QGEIGRRUGA SLAEGALSLS AMSKIPEKLA FYDYIGNNPA ANABABABA VVAGTMYGH VTTPIELFGP TRYGWDGOFF TGEIGRRUGS GLAEGASLER AMSKIPEKLA FYDYIGNNPA ANABABABA VVAGTMYGH ATTFIELFGP TRYGWDGOFF TGEIGRRUGS GLAEGASLER AMSKIPEKLA FYDYIGNSPA TRYGWDGOFF TGEIGRRUGS GLAEGASLER AMSKIPEKLA FYDYIGNSPA AMILE KGGLFRAGSM DNGGGIAVGH LGHPIFRDKE GRELFVRRMP TFFETFPVUL VDGGGIVAAD VPFRAESKY TRYGWDGOFF TO THE TO THE TO THE TO THE THE TOTAL THE TO	Rice	VVAGTMWYGS	ATTPIELFGP	TRYQWDQGYF	QQEIYRRVSD	GLAENLSLSE	AWSKIPEKLA	FYDYIGNNPA
	Rye/Wheat	VVAGTMWYGS	ATTPIELFGP	TRYQWDQGYF	QQEIYRRVSN	GLAENLSLSE	AWSKIPEKLA	FYDYIGNNPA
	•	VVAGTMWYGS	ATTPIELFGP	TRYQWDQGYF	QQEIYRRVSN	GLAENLSLSE	AWSKIPEKLA	FYDYIGNNPA
Euglana	•	VVAGTMWYGS	AATPIELFGP	TRYQWDQGFF	QQEIDRRIRS	SKAENLSLSE	AWSKIPEKLA	FYDYIGNNPA
Prochloro. VVAGTMYNYCH VITTPIELFOP TRYOMOGNYF TOELQRRUDS QLAEGASLSE AWSSIPEKLA FYDYVGNSPA Anabaena VVAGTMYGN ATTPIELFOP TRYQMDQGYF HQEICRRUDS SVAQASLSE AWSQIPEKLA FYDYVGNSPA Syn.8803 VVAGTMYGN ATTPIELFOP TRYQMDQFF GEECRRUDS QLAEGASLSE AWSTTPEKLA FYDYVGNSPA Consensus VV-GTMYGN -TEPELFOP TRYQMDQFF -EE-R-R S-S- AW-LIKLA FYDYVGNSPA London KGGLFRAGSM DNGDGIAVGW LGHPIFRDKE GRELFVRRMP TFFETFPVVL VDGDGIVRAD VPFRAESKY Maize KGGLFRAGSM DNGDGIAVGW LGHPUFRDKE GRELFVRRMP TFFETFPVVL VDEGGIVRAD VPFRAESKY Rice KGGLFRAGSM DNGDGIAVGW LGHPUFRDKE GRELFVRRMP TFFETFPVVL VDEGGIVRAD VPFRAESKY Livervort KGGLFRAGSM DNGDGIAVGW LGHPUFRDKE GRELFVRRMP TFFETFPVVL VDEGGIVRAD VPFRAESKY Livervort KGGLFRAGSM NNGDGIAVGW LGHAVFIDGE G	Chlamydom.	VVAGTMWYGS	AATPIELFGP	TRYQWDQGFF	QQEIQKRVQA	SLAEGASLSD	AWSRIPEKLA	FYDYIGNNPA
Anabaena VVAGTMYGM ATTFIELEGP TRYGWOGYF HGEIERRVGS SVAQGASLSE AWSGIPEKLA FYDYVGNSPA Syn.6803 VVAGTMYGM ATTFIELEGP TRYGWOGYF GEELGRAVDS QLAEGASLSE AWSGIPEKLA FYDYVGNSPA AWSTIPEKLA FYDYVGNSPA FYDYGNSPA FYDYGNSPA FYDYGNSPA FYDYYGNSPA FYDYGNSPA FYD	Euglena	VVSGTMWYGS	ASTPIELFGP	TRYQWDKGYF	QEEIERRVQA	SLSDGCSLSE	AWGAISPKLA	FYDYIGNNPA
Syn.6801 VVACTHWYGN ATTETELED TRYQWENGYF QEECGREVOS QLAEGASLE AWSTIPERLA FYDYVONSPA CONSENSUS VV-GTMWYGTPIELED TRYQWING-FEFRS-S- AW-IRIA FYDY-GN-PA 321 331 341 351 361 371 381 TODAGCO KGGLFRAGSM DNGDGIAVGW LGHPIFRDKE GRELFVRRMP TFFETFFVVL VDGDGIVRAD VPFRRAESKY Spinach KGGLFRAGSM DNGDGIAVGW LGHPIFRDKE GRELFVRRMP TFFETFFVVL VDGDGIVRAD VPFRRAESKY Maize KGGLFRAGSM DNGDGIAVGW LGHPIFRDKE GRELFVRRMP TFFETFFVVL VDEGGIVRAD VPFRRAESKY MAIZE KGGLFRAGSM DNGDGIAVGW LGHPYFRDKE GRELFVRRMP TFFETFFVVL VDEGGIVRAD VPFRRAESKY RYP-Wheat KGGLFRAGSM DNGDGIAVGW LGHPYFRDKE GRELFVRRMP TFFETFFVVL VDEGGIVRAD VPFRRAESKY LIVETWOTK KGGLFRAGSM DNGDGIAVGW LGHPYFRDKE GRELFVRRMP TFFETFFVVL VDEGGIVRAD VPFRRAESKY LIVETWOTK KGGLFRAGSM DNGDGIAVGW LGHPYFRDKE GRELFVRRMP TFFETFFVVL VDEGGIVRAD VPFRRAESKY EUGlena KGGLFRAGSM DNGDGIAVGW LGHAVFRDKE GRELFVRRMP TFFETFFVVL VDEGGIVRAD VPFRRAESKY EUGlena KGGLFRAGSM DNGDGIAVGW LGHAVFRDKE GRELFVRRMP TFFETFFVVL VDEGGIVRAD VPFRRAESKY EUGlena KGGLFRAGSM DNGDGIAVGW LGHAVFRDKE GRELFVRRMP TFFETFFVVL VDEGGIVRAD VPFRRAESKY EUGlena KGGLFRAGSM DNGDGIAVGW LGHAVFROKE GRELFVRRMP TFFETFFVVL VDEGGIVRAD VPFRRAESKY EUGlena KGGLFRAGSM DNGDGIAVGW LGHAVFRDKE GRELFVRRMP TFFETFFVVL VDEGGIVRAD VPFRRAESKY EUGlena KGGLFRAGSM NSGDGIAVGW GHAVFROKE GRELFVRRMP TFFETFFVVL VDEGGIVRAD VPFRRAESKY EUGlena KGGLFRAGSM NSGDGIAVGW GHAVFROKE GRELFVRRMP TFFETFFVVL TDADGVVRAD IPFRRAESKY EUGlena KGGLFRAGSM NSGDGIAVGW GHAVFROKE GRELFVRRMP NFFETFFVVL TDADGVVRAD IPFRRAESKY EUGlena SVEQVGVTVE FYGGELNGVS YSDPATVKKY ARRAQLGEIF ELDRATLKSD GVFRSSPRGW FTFGHASFAL ENGLES SVEQVGVTVE FYGGELNGVS YSDPATVKKY ARRAQLGEIF ELDRATLKSD GVFRSSPRGW FTFGHAFFAL ENGlena SVEQVGVTVE FYGGELNGVS YSDPATVKKY ARRAQLGEIF ELDRATLKSD GVFRSSPRGW FTFGHAFFAL ENGlena SVEQVGVTVE FYGGELNGVS YSDPATVKKY ARRAQLGEIF ELDRATLKSD GVFRSSPRGW FTFGHAFFAL ENGlena SVEQVGVTVE FYGGELNGVS YSDPATVKKY ARRAQLGEIF ELDRATLKSD GVFRSSPRGW FTFGHAFFAL ENGLANDER SVEQVGVTVE FYGGELNGVS YSDPATVKKY ARRAQLGEIF ELDRATLKSD GVFRSSPRGW FTFGHAFFAL ENGLANDER SVEQVGVTVE FYGGELNGVS FSDPATVKKY ARRAQLGEIF ELDRATLKSD	-	VVAGTMWYGH	VTTPIELFGP	TRYQWDQGYF	TQEIQRRVDS	QLAEGASLSE	AWSSIPEKLA	FYDYVGNSPA
	Anabaena	VVAGTMWYGN	ATTPIELFGP	TRYQWDQGYF	HQEIERRVQS	SVAQGASLSE	AWSQIPEKLA	FYDYVGNSPA
	Syn.6803	VVAGTMWYGN	ATTPLELFGP	TRYQWDKGYF	QEEIQRRVDS	QLAEGASLSE	AWSTIPEKLA	FYDYVGNSPA
TobaccoKGGLFRAGSMDMGDGIAVGWLGHPIFRDKEGRELFVRRMPTFFETFPVVLVDGDGIAVADVPFRRAESKYSpinachKGGLFRAGSMDMGDGIAVGWLGHPIFRDKEGRELFVRRMPTFFETFPVVLVDGDGIAVADVPFRRAESKYCenotheraKGGLFRAGSMDMGDGIAVGWLGHPVFRDKEGRELFVRRMPTFFETFPVVLVDGDGIAVADVPFRRAESKYMaizeKGGLFRAGSMDMGDGIAVGWLGHPVFRDKEGRELFVRRMPTFFETFPVVLVDEEGIVRADVPFRRAESKYRye/WheatKGGLFRAGSMDMGDGIAVGWLGHPVFRDKEGRELFVRRMPTFFETFPVVLVDEEGIVRADVPFRRAESKYLiverwortKGGLFRAGSMDMGDGIAVGWLGHAVFKDKEGRELFVRRMPTFFETFPVVLVDEEGIVRADVPFRRAESKYLiverwortKGGLFRAGAMDMGDGIAVGWLGHAVFKDKEGRELFVRRMPTFFETFPVVLVDEGGIVRADVPFRRAESKYLiverwortKGGLFRAGAMMNGDGIAVGWLGHAVFKDKEGRELFVRRMPTFFETFPVVLVDEGGIVRADVPFRRAESKYEuglenaKGGLFRAGAMNMGDGIAVGWLGHAVFKDKEGRELFVRRMPTFFETFPVVLLDKNGIVRADVPFRRAESKYProchloroKGGLFRAGAMMNGDGIATAWLGHAVFIDKEGRELFVRRMPTFFETFPVVLLDQMGVVRADIPFRAESKYSyn. 6803KGGLFRAGAMNMGDGIATAWLGHAVFIDKEGRELFVRRMPTFFETFPVVLLDQMGVVRADIPFRAESKYSyn. 6803KGGLFRAGAMNMGDGIATAWLGHAVFIDKEGRELFVRRMPTFFETFPVVLLDQMGVVRADIPFRAESKYTODAGOVARANSGDGIACWLGHPYFDKEGRELFVRRMPTFFETFPVVLLDQMGVVRADIPFRAESKY </td <td>Consensus</td> <td>VV-GTMWYG-</td> <td>TPIELFGP</td> <td>1 11</td> <td>EIR</td> <td>s-s-</td> <td>AW-IKLA</td> <td>FYDY-GN-PA</td>	Consensus	VV-GTMWYG-	TPIELFGP	1 11	EIR	s-s-	AW-IKLA	FYDY-GN-PA
TobaccoKGGLFRAGSMDMGDGIAVGWLGHPIFRDKEGRELFVRRMPTFFETFPVVLVDGDGIAVADVPFRRAESKYSpinachKGGLFRAGSMDMGDGIAVGWLGHPIFRDKEGRELFVRRMPTFFETFPVVLVDGDGIAVADVPFRRAESKYCenotheraKGGLFRAGSMDMGDGIAVGWLGHPVFRDKEGRELFVRRMPTFFETFPVVLVDEGIVRADVPFRRAESKYMaizeKGGLFRAGSMDMGDGIAVGWLGHPVFRDKEGRELFVRRMPTFFETFPVVLVDEGIVRADVPFRRAESKYRye/WheatKGGLFRAGSMDMGDGIAVGWLGHPVFRDKEGRELFVRRMPTFFETFPVVLVDEGIVRADVPFRRAESKYLiverwortKGGLFRAGSMDMGDGIAVGWLGHAVFKDKEGRELFVRRMPTFFETFPVVLVDEGIVRADVPFRRAESKYLiverwortKGGLFRAGAMDMGDGIAVGWLGHAVFKDKEGRELFVRRMPTFFETFPVVLVDEGIVRADVPFRRAESKYLiverwortKGGLFRAGAMMNGDGIAVGWLGHAVFKDKEGRELFVRRMPTFFETFPVVLVDEGIVRADVPFRRAESKYLiverwortKGGLFRAGAMNNGDGIAVGWLGHAVFKDKEGRELFVRRMPTFFETFPVVLVDEGIVRADVPFRRAESKYEuglenaKGGLFRAGAMNNGDGIAVGWLGHAVFKDKEGRELFVRRMPTFFETFPVVLLDQMGVVRADLPFRAESKYProchloroKGGLFRAGAMNNGDGIAVGWLGHAVFKDKEGRELFVRRMPTFFETFPVVLLDQMGVVRADLPFRRAESKYSyn. 6803KGGLFRAGAMNNGDGIAVGWLGHAVFKDKEGRELFVRRMPTFFETFPVVLLDQMGVVRADLPFRRAESKYTODAGOVRADNSGDGIAVGWLGHPFFRAESKYNRFALFERPVLNFFETFPVVLLDQMGVVRADLPFRRAESKY<							<u> </u>	
Spinach KGGLFRAGSM DNGOGIAVGW LGHPIFROKE GRELFVRRMP TFFETFPVVL VDGDGIVRAD VPFRRAESKY Maize KGGLFRAGSM DNGDGIAVGW LGHPIFROKE GRELFVRRMP TFFETFPVVL VDGDGIVRAD VPFRRAESKY Maize KGGLFRAGSM DNGDGIAVGW LGHPYRDKE GRELFVRRMP TFFETFPVVL VDEGIVRAD VPFRRAESKY Rye/Wheat KGGLFRAGSM DNGDGIAVGW LGHPYRDKE GRELFVRRMP TFFETFPVVL VDEGIVRAD VPFRRAESKY DNGDGIAVGW LGHPYRDKE GRELFVRRMP TFFETFPVVL VDEGIVRAD VPFRRAESKY LIVERVORT KGGLFRAGSM DNGDGIAVGW LGHPYRDKE GRELFVRRMP TFFETFPVVL VDEGIVRAD VPFRRAESKY LIVERVORT KGGLFRAGSM DNGDGIAVGW LGHAVFKOKE GRELFVRRMP TFFETFPVVL VDEGIVRAD VPFRRAESKY LIVERVORT KGGLFRAGAM DNGDGIAVGW LGHAVFKOKE GRELFVRRMP TFFETFPVVL VDEGIVRAD VPFRRAESKY LDAMBOARD KGGLFRAGAM DNGDGIAVGW LGHAVFKOKE GRELFVRRMP TFFETFPVVL VDEGIVRAD VPFRRAESKY LGHAVFKOKE GRELFVRRMP TFFETFPVL LDAMGIVRAD VPFRRAESKY LGHAVFKOKE GRELFVRRMP TFFETFPVL LDAMGIVRAD VPFRRAESKY PTOCHOLFO. ***AGGLFRAGAM NNGDGIAVGW LGHAVFKOKE GRELFVRRMP TFFETFPVL LDAMGIVRAD IPFRRAESKY PTOCHOLFO.** ***AGGLFRAGAM NNGDGIAVGW LGHAVFKOKE GRELFVRRMP TFFETFPVL LDAMGIVRAD IPFRRAESKY PTOCHOLFO.** ***AGGLFRAGAM NNGDGIAVGW LGHAVFKOKE GRELFVRRMP TFFETFPVL LDAMGIVRAD IPFRRAESKY PTOCHOLFO.** ***AGGLFRAGAM NNGDGIAVGW LGHAVFKOKE GRELFVRRMP TFFETFPVL TDAMGIVRAD IPFRRAESKY PTOCHOLFO.** ***AGGLFRAGAM NNGDGIAVGW LGHAVFKOKE GRELFVRRMP NFFETFPVL LDAMGIVRAD IPFRRAESKY PTOCHOLFO.** ***AGGLFRAGAM NNGDGIAVGW LGHAVFKOKE GRELFVRRMP NFFETFPVL TDAMGIVRAD IPFRRAESKY PTOCHOLFO.** ***AGGLFRAGAM NNGDGIAVGW LGHAVFKOKE GRELFVRRMP NFFETFPVL TDAMGIVRAD IPFRRAESKY PTFGHATFAL PTALFAL PTALF		321	331	341	351	361	371	381
Oenothera KGGLFRAGSM DSGGIAVGW LGHPIFROKE GRELFVRRMP TFFETFPVVL VDGDGIVRAD VPFRRAESKY Maize KGGLFRAGSM DNGDGIAVGW LGHPVFRDKE GRELFVRRMP TFFETFPVVL VDEEGIVRAD VPFRRAESKY Rye/Wheat KGGLFRAGSM DNGDGIAVGW LGHPVFRDKE GRELFVRRMP TFFETFPVVL VDEEGIVRAD VPFRRAESKY RYe/Wheat KGGLFRAGSM DNGDGIAVGW LGHPVFRDKE GRELFVRRMP TFFETFPVVL VDEEGIVRAD VPFRRAESKY LIVERWORT KGGLFRAGSM DNGDGIAVGW LGHAVFKDKE GRELFVRRMP TFFETFPVVL VDEEGIVRAD VPFRRAESKY LGHAVFKDKE GRELFVRRMP TFFETFPVVL LDXDGIVVRAD IPFRRAESKY SYN. 6803 KGGLFRTGAM NSGDGIAVGW LGHAVFKDKE GRELFVRRMP TFFETFPVVL LDXDGIVVRAD IPFRRAESKY SYN. 6803 KGGLFRTGAM NSGDGIAQEW LGHPVFRDKE GRELFVRRMP NFFETFPVVL TDADGVVRAD IPFRRAESKY SYN. 6803 KGGLFRTGAM NSGDGIAQEW LGHPVFRDKE GRELFVRRMP NFFETFPVIL TDADGVVRAD IPFRRAESKY SYN. 6803 KGGLFR-G-MGDGIA-W -GH-F-D- GL-VRR-P -FFE-FPVDG-VRAD IPFRRAESKY SYN. 6803 KGGLFR-G-MGDGIA-W -GH-F-D- GL-VRR-P -FFE-FPVDG-VRAD IPFRRAESKY SYDATVKKY ARRAQLGEIF ELDRATLKSD GVFRSSPRGW FTFGHASFAL SYEQVGVTVE FYGGELNGVS YSDPATVKKY ARRAQLGEIF ELDRATLKSD GVFRSSPRGW FTFGHASFAL SVEQVGVTVE FYGGELNGVS YSDPATVKKY ARRAQLGEIF ELDRATLKSD GVFRSSPRGW FTFGHAFFAL LIVERWORT SYEQVGVTVE FYGGELNGVS YSDPATVKKY ARRAQLGEIF ELDRATLKSD GVFRSSPRGW FTFGHAFFAL LIVERWORT SYEQVGVTVE FYGGELNGVS YSDPATVKKY ARRAQLGEIF ELDRATLKSD GVFRSSPRGW FTFGHATFAL LIVERWORT SYEQVGVTVE FYGGELNGVS YSDPATVKKY ARRAQLGEIF ELDRATLKSD GVFRSSPRGW FTFGHATFAL LIVERWORT SYEQVGVTVE FYGGELNGVS YSDPATVKKY ARRAQLGEIF ELDRATLKSD GVFRSSPRGW FTFGHATFAL LIVERWORT SYEQVGVTVE FYGGELNGVS TSDPATVKKY ARRAQLGEIF ELDRATLKSD GVFRSSPRGW FTFGHATFAL LIVERWORT SYEQVGVTVE FYGGELNGVS FSDPATVKKY ARRAQLGEIF ELDRATLKSD GVFRSSPRGW FTFGHATFAL LIVERWORT SYEQVGVTVE FYGGELNGVS FSDPATVKKY ARRAQLGEIF EFDRATLKSD GVFRSSPRGW FTFGHATFAL LIVERWORT SYEQVGVTVE FYGGELNGVS FSDPATVK	Tobacco	KGGLFRAGSM	DNGDGIAVGW	LGHPIFRDKE	GRELFVRRMP	TFFETFPVVL	VDGDGIVRAD	VPFRRAESKY
Maize KGGLFRAGSM DNDDGIAVGW LGHPVFRDKE GRELFVRRMP TFFETFPVVL VDEEGIVRAD VPFRRAESKY Rye/Wheat KGGLFRAGSM DNDDGIAVGW LGHPVFRDKE GRELFVRRMP TFFETFPVVL VDEEGIVRAD VPFRRAESKY LIVERWOTT KGGLFRAGSM DNDDGIAVGW LGHPVFRDKE GRELFVRRMP TFFETFPVVL VDEEGIVRAD VPFRRAESKY LIVERWOTT KGGLFRAGSM DNDDGIAVGW LGHPVFRDKE GRELFVRRMP TFFETFPVVL VDEEGIVRAD VPFRRAESKY LIVERWOTT KGGLFRAGSM DNDDGIAVGW LGHAVFKDKE GRELFVRRMP TFFETFPVVL VDEEGIVRAD VPFRRAESKY CHIAMYDOM. KGGLFRAGSM NSGDGIAVGW LGHAVFKDKE GRELFVRRMP TFFETFPVVL LDKDGIVRAD VPFRRAESKY LGHAVFLOKE GRELFVRRMP TFFETFPVL LDKDGIVRAD VPFRRAESKY LGHAVFLOKE GRELFVRRMP TFFETFPVL LDKDGIVRAD VPFRRAESKY LGHAVFLOKE GRELFVRRMP TFFETFPVL LDKDGIVRAD VPFRRAESKY LGHAVFLOKE GRELFVRRMP NFFETFPVL LDKDGIVRAD IPFRRAESKY GRELFVRRMP NFFETFPVL LDKDGIVRAD IPFRRAESKY GRELFVRRMP NFFETFPVL LDKDGIVRAD IPFRRAESKY GRELFVRRMP NFFETFPVL TDADGVVRAD IPFRRAESKY SVEQVGVTVE FYGGELNGVS YSDPATVKKY ARRAQLGEIF ELDRATLKSD GVFRSSPRGW FTFGHASFAL SVEQVGVTVE FYGGELNGVS YSDPATVKKY ARRAQLGEIF ELDRATLKSD GVFRSSPRGW FTFGHAFFAL RYE/Wheat SVEQVGVTVE FYGGELNGVS YSDPATVKKY ARRAQLGEIF ELDRATLKSD GVFRSSPRGW FTFGHATFAL LIVERWOTT SVEQVGVTVE FYGGELNGVS YSDPATVKKY ARRAQLGEIF ELDRATLKSD GVFRSSPRGW FTFGHATFAL LIVERWOTT SVEQVGVTVE FYGGELNGVS YSDPATVKKY ARRAQLGEIF ELDRATLKSD GVFRSSPRGW FTFGHATFAL CLIVERWOTT SVEQVGVTVE FYGGELNGVS YSDPATVKKY ARRAQLGEIF ELDRATLKSD GVFRSSPRGW FTFGHATFAL CLIVERWOTT SVEQVGVTVE FYGGELNGVS FSDPATVKKY ARRAQLGEIF ELDRATLKSD GVFRSSPRGW FTFGHATFAL CLIVERWOTT SVEQVGVTVE FYGGELNGVS FSDPATVKKY ARRAQLGEIF ELDRATLKSD GVFRSSPRGW FTFGHATFAL CLIVERWOTT SVEQVGVTVE FYGGELNGVS FSDPATVKKY ARRAQLGEIF EL	Spinach	KGGLFRAGSM	DNGDGIAVGW	LGHPIFRDKE	GRELFVRRMP	TFFETFPVVL	IDGDGIVRAD	VPFRRAESKY
Rice KGGLFRAGSM DNGDGIAVGW LGHPIFRDKE GRELFVRRMP TFFETFVVL VDEEGIVRAD VPFRRAESKY Rye/wheat KGGLFRAGSM DNGDGIAVGW LGHPVFRDKE GRELFVRRMP TFFETFVVL VDEEGIVRAD VPFRRAESKY Liverwort KGGLFRAGSM DNGDGIAVGW LGHAVFKDKE GRELFVRRMP TFFETFVVL VDEEGIVRAD VPFRRAESKY Liverwort KGGLFRAGSM DNGDGIAVGW LGHAVFKDKE GRELFVRRMP TFFETFVVL VDEEGIVRAD VPFRRAESKY Chlamydom. KGGLFRTGAM NSGDGIAVGW LGHAVFKDKE GRELFVRRMP TFFETFVVL VDEQGIVRAD VPFRRAESKY LDWGGIAVGW LGHAVFKDKE GRELFVRRMP TFFETFVVL LDKDGIVRAD VPFRRAESKY LDWGGIAVGW LGHAVFLDKE GRELFVRRMP TFFETFVVL LDKDGIVRAD VPFRRAESKY PROCHOECT TO THE TRETTY TO THE TRE	Oenothera	KGGLFRAGSM	DSGDGIAVGW	LGHPIFRDKE	GRELFVRRMP	TFFETFPVVL	VDGDGIVRAD	VPFRRAESKY
Rye/Wheat KGGLFRAGSM DNGDGIAVGW LGHPVFRDKE GRELFVRRMP TFFETFVVL VDEEGIVRAD VPFRRAESKY Liverwort KGGLFRAGSM DNGDGIAVGW LGHPVFRDKE GRELFVRRMP TFFETFVVL VDEEGIVRAD VPFRRAESKY Chlamydom. KGGLFRAGAM DNGDGIAVGW LGHAVFKDKE GNELFVRRMP TFFETFVVL VDEQGIVRAD VPFRRAESKY Chlamydom. KGGLFRAGAM NSGDGIAVGW LGHAVFKDKE GNELFVRRMP TFFETFVVL LDKDGIVRAD VPFRRAESKY LGHAVFLOKE GNELFVRRMP TFFETFVVL LDKDGIVRAD VPFRRAESKY LGHAVFLOKE GNELFVRRMP TFFETFVVL LDKDGIVRAD VPFRRAESKY LGHAVFLOKE GNELFVRRMP TFFETFVVL LDKDGIVRAD VPFRRAESKY LDKDGIVRAD VPFRRAESKY LDROWVRAD IPFRRAESKY NFFETFVVL LDKDGIVRAD IPFRRAESKY NFFETFVVL LDKDGIVRAD IPFRRAESKY SYN.6803 KGGLFRTGAM NSGDGIAEW LGHPVFQDGA GRALSVRRLP NFFETFVVL TDGDGVVRAD IPFRRAESKY SYN.6803 KGGLFRTGAM NSGDGIAQEW IGHPIFKDKE GRELTVRRLP NFFETFVVL TDADGVVRAD IPFRRAESKY CONSENSUS KGGLFRTGAM NSGDGIAQEW IGHPIFKDKE GRELTVRRLP NFFETFVVL TDADGVVRAD IPFRRAESKY SYN.6803 KGGLFR-G-MGDGIAW -GHF-D G-L-VRR-P -FFE-FFVDG-VRAD -PFRES TODAGCO SVEQVGVTVE FYGGELNGVS YSDPATVKKY ARRAQLGEIF ELDRATLKSD GVFRSSPRGW FTFGHASFAL SVEQVGVTVE FYGGELNGVS YSDPATVKKY ARRAQLGEIF ELDRATLKSD GVFRSSPRGW FTFGHASFAL SVEQVGVTVE FYGGELNGVS YSDPATVKKY ARRAQLGEIF ELDRATLKSD GVFRSSPRGW FTFGHATFAL LIVERWORT SVEQVGVTVE FYGGELNGVS YSDPATVKKY ARRAQLGEIF ELDRATLKSD GVFRSSPRGW FTFGHATFAL LIVERWORT SVEQVGVTVE FYGGELNGVS YSDPATVKKY ARRAQLGEIF ELDRATLKSD GVFRSSPRGW FTFGHATFAL CHLAMPDOM. SIEQVGVSVT FYGGELDGUS FSDPATVKKY ARRAQLGEIF EDRATLKSD GVFRSSPRGW FTFGHATFAL EUGHAM SIEQVGVSVT FYGGELDGUT FTDPATVKKY ARRAQLGEIF EDRATLKSD GVFRSSPRGW FTFGHATFAL DPROCHORO. SFEQGSTVS FYGGLDGKT FTNPSDVKKF ARRAQLGEIF GDRATLKSD GVFRSSPRGW FTFGHATFAL SIEQVGVTVF FYGGELDGUT FTNPSDVKKF ARRAQLGEIF GDRATLKSD GVFRSSPRGW FTGHATFAL STEAGH FTGHAVFAL SFEGGSTVS FYGGLDGKT FTNPSDVKKF ARRAQLGEIF GDRATLKSD GVFRSSPRGW FTGHATFAL SFEGGSTVS FYGGLDGKT FTNPSD	Maize	KGGLFRAGSM	DNGDGIAVGW	LGHPVFRDKE	GRELFVRRMP	TFFETFPVVL	VDEEGIVRAD	VPFRRAESKY
Barley KGGLFRAGSM DNGDGIAVGW LGHPVFRDKE GRELFVRRMP TFFETFPVUL VDEEGIVRAD VPFRRAESKY Liverwort KGGLFRAGAM DNGDGIAVGW LGHAVFKDKE GNELFVRRMP TFFETFPVUL VDEQGIVRAD VPFRRAESKY Chlamydom. KGGLFRAGAM NSGDGIAVGW LGHASFKDQE GRELFVRRMP TFFETFPVUL LDKDGIVRAD VPFRRAESKY Euglena KGGLFRAGAM NSGDGIAVGW LGHASFKDQE GRELFVRRMP TFFETFPVUL LDKDGIVRAD VPFRRAESKY Prochloro. KGGLFRAGAM DSGDGIAEW LGHAVFIDKE GNSLFVRRMP TFFETFPVUL LDQNGVVRAD IPFRRAESKY Anabaena KGGLFRAGAM NSGDGIAQEW LGHAVFIDKE GRELFVRRMP NFFETFPVUL TDADGVVRAD IPFRRAESKY Syn.6803 KGGLFRAGAM NSGDGIAQEW IGHPIFKDKE GRELFVRRMP NFFETFPVUL TDADGVVRAD IPFRRAESKY GCNSANSUS KGGLFRAGAM NSGDGIAQEW IGHPIFKDKE GRELFVRRMP NFFETFPVUL TDADGVVRAD IPFRRAESKY Syn.6803 KGGLFRAGAM NSGDGIAQEW IGHPIFKDKE GRELFVRRMP NFFETFPVUL TDADGVVRAD IPFRRAESKY GCNSANSUS KGGLFRAGAM NSGDGIAQEW IGHPIFKDKE GRELFVRRMP NFFETFPVUL TDADGVVRAD IPFRRAESKY Syn.6803 KGGLFRAGAM NSGDGIAQEW IGHPIFKDKE GRELFVRRMP NFFETFPVUL TDADGVVRAD IPFRRAESKY GRELFVRRMP NFFETFPVUL TDADGVVRAD IPFRRAESKY ITPRAESKY NFFETFPVUL LDQNGVVRAD IPFRRAESKY ITPRAESKY NFFETFPVUL LDQNGVVRAD IPFRRAESKY ITPRAESKY NFFETFPVUL LDQNGVVRAD IPFRRAESKY ITPRAESKY NFFETFPVUL LDQNGVVRAD IPFRRAESKY ITPRAESKY NFFETFPVUL TDADGVVRAD IPFRRAESKY ITPRAESKY NFFETFPVUL LDQNGVVRAD IPFRRAESKY ITPRAESKY NFFETFPVUL TDADGVVRAD IPFRRAESKY ITPRAESKY NFFETFPVUL LDQNGVVRAD IPFRRAESKY ITPRAESKY ITPRAESK	Rice	KGGLFRAGSM	DNGDGIAVGW	LGHPIFRDKE	GRELFVRRMP	TFFETFPVVL	VDEEGIVRAD	VPFRRAESKY
Liverwort KGGLFRAGAM DNGDGIAVGW LGHAVFKDKE GNELFVRRMP TFFETFPVUL VDEQGIVRAD VPFRRAESKY Chlamydom. KGGLFRTGAM NSGDGIAVGW LGHASFKDQE GRELFVRRMP TFFETFPVUL LDKDGIVRAD VPFRRAESKY Euglena KGGLFRSGPM NNGDGIATAW LGHAVFIDKE GNSLFVRRMP TFFETFPVUL LDQNGVVRAD IPFRRAESKY Prochloro. KGGLFRVGAM DSGDGIAEEW LGHPVFQDGA GRALSVRILP NFFETFPVUL TDQDGVVRAD IPFRRAESKY Syn.6803 KGGLFRTGAM NSGDGIAQSW QGHGVFKDAE GRELTVRRLP NFFETFPVUL TDADGVVRAD IPFRRAESKY Syn.6803 KGGLFRTGAM NSGDGIAQEW IGHPIFKDKE GRELEVRRMP NFFETFPVUL TDADGVVRAD IPFRRAESKY Syn.6803 KGGLFR-G-M -GD-IA-W -GH-F-D- G-L-VRR-P -FFE-FPVDG-VRAD -PFR-ES TODACCO SVEQVGVTVE FYGGELNGVS YSDPATVKKY ARRAQLGEIF ELDRATLKSD GVFRSSPRGW FTFGHASFAL SYEQVGVTVE FYGGELNGVS YSDPATVKKY ARRAQLGEIF ELDRATLKSD GVFRSSPRGW FTFGHASFAL Rice SVEQVGVTVE FYGGELNGVS YSDPATVKKY ARRAQLGEIF ELDRATLKSD GVFRSSPRGW FTFGHAFFAL Rye/Wheat SVEQVGVTVE FYGGELNGVS YSDPATVKKY ARRAQLGEIF ELDRATLKSD GVFRSSPRGW FTFGHAFFAL Chlamydom. SIEQVGVTVE FYGGELNGVS FSDPATVKKY ARRAQLGEIF ELDRATLKSD GVFRSSPRGW FTFGHAFFAL Chlamydom. SIEQVGVTVE FYGGELDGUT FTDPATVKKY ARRAQLGEIF EFDRATLKSD GVFRSSPRGW FTFGHAFFAL Chlamydom. SIEQVGVTVE FYGGELDGUT FTDPATVKKY ARRAQLGEIF EFDRSTLQSD GVFRSSPRGW FTFGHAFFAL Chlamydom. SIEQVGVTVE FYGGELDGUT FTDPATVKKY ARRAQLGEIF EFDRSTLQSD GVFRSSPRGW FTFGHAFFAL Chlamydom. SIEQVGVTVE FYGGELDGUT FTDPATVKKY ARRAQLGEIF EFDRSTLQSD GVFRSSPRGW FTFGHAFFAL Chlamydom. SFEQGCVTVS FYGGALDGQT FTNPSDVKKF ARRAQLGEIF EFDRSTLQSD GVFRSSPRGW FTFGHAFFAL SYR.6803 SVEQTGVTVS FYGGALDGQT FTNPSDVKKF ARRAQLGEIF EFDRSTLNSD GVFRTSFRGW FTFGHAFFAL SYR.6803 SVEGTGVTVS FYGGALDGQT FTNPSDVKKF ARRAQLGEIF EFDRSTLNSD GV	Rye/Wheat	KGGLFRAGSM	DNGDGIAVGW	LGHPVFRDKE	GRELFVRRMP	TFFETFPVVL	VDEEGIVRAD	VPFRRAESKY
Euglena KGGLFRTGAM NSGDGIAVGW LGHASFKDQE GRELFVRRMP TFFETFPVLL LDKDGIVRAD VFFRKAESKY Prochloro. KGGLFRVGAM DSGDGIAEW LGHAVFIDKE GNSLFVRRMP TFFETFPVLL LDQNGVVRAD IPFRRAESKY Anabaena KGGLFRTGAM NSGDGIAQEW LGHPVFQDGA GRALSVRLP NFFETFPVLL TDADGVVRAD IPFRRAESKY Syn.6803 KGGLFRTGAM NSGDGIAQEW IGHPIFKDKE GRELFVRRP NFFETFPVLL TDADGVVRAD IPFRRAESKY COBSENSUS KGGLFRTGAM NSGDGIAQEW IGHPIFKDKE GRELFVRRP NFFETFPVLM TDADGVVRAD IPFRRAESKY COBSENSUS KGGLFRTGAM NSGDGIAQEW IGHPIFKDKE GRELFVRRP NFFETFPVLM TDADGVVRAD IPFRRAESKY FTGHASFAL TOBACCO SVEQUGVTVE FYGGELNGVS YSDPATVKKY ARRAQLGEIF ELDRATLKSD GVFRSSPRGW FTFGHASFAL Oenothera SVEQUGVTVE FYGGELNGVS YSDPATVKKY ARRAQLGEIF ELDRATLKSD GVFRSSPRGW FTFGHASFAL Rice SVEQUGVTVE FYGGELNGVS YSDPATVKKY ARRAQLGEIF ELDRATLKSD GVFRSSPRGW FTFGHAFFAL Rye/Wheat SVEQUGVTVE FYGGELNGVS YSDPATVKKY ARRAQLGEIF ELDRATLKSD GVFRSSPRGW FTFGHAFFAL Rye/Wheat SVEQUGVTVE FYGGELNGVS YSDPATVKKY ARRAQLGEIF ELDRATLKSD GVFRSSPRGW FTFGHAFFAL Barley SVEQUGVTVE FYGGELNGVS YSDPATVKKY ARRAQLGEIF ELDRATLKSD GVFRSSPRGW FTFGHAFFAL LIVERWORT SVEQUGVTVE FYGGELNGVS YSDPATVKKY ARRAQLGEIF ELDRATLKSD GVFRSSPRGW FTFGHATFAL Chlamydom. SIEQUGVTVE FYGGELNGVS FSDPATVKKY ARRAQLGEIF ELDRATLKSD GVFRSSPRGW FTFGHATFAL Chlamydom. SIEQUGVTVF FYGGELNGVS FSDPATVKKY ARRAQLGEIF EFDRATLKSD GVFRSSPRGW FTFGHATFAL Euglena SIEQUGVTVF FYGGELDGLT FTDPATVKKY ARRAQLGEIF EFDRATLSD GVFRSSPRGW FTFGHATFAL Prochloro. SFEGTGVTVS FYGGALDGCT FTDPATVKKY ARRAQLGEIF EFDRATLSD GVFRSSPRGW FTFGHATFAL FTGHATFAL FTGHAT	Barley	KGGLFRAGSM	DNGDGIAVGW	LGHPVFRDKE	GRELFVRRMP	TFFETFPVVL	VDEEGIVRAD	VPFRRAESKY
Euglena KGGLFRSGPM NNGDGIATAW LGHAVFIDKE GNSLFVRRMP TFFETFPVIL LDQNGVVRAD IPFRRAESKY Prochloro. KGGLFRVGAM DSGDGIAGEW LGHPVFQDGA GRALSVRRLP NFFENFPVIL TDGDGVVRAD IPFRRAESKY Syn.6803 KGGLFRTGAM NSGDGIAQEW IGHPIFKDKE GRELTVRRLP NFFETFPVIL TDADGVVRAD IPFRRAESKY IGHPIFKDKE GRELTVRRLP NFFETFPVIL TDADGVVRAD IPFRRAESKY CONSENSUS KGGLFR-G-MGDGIA-W -GHF-D G-L-VRR-P -FFE-FPVDG-VRAD -PFRES TODAGCCO SVEQVGVTVE FYGGELNGVS YSDPATVKKY ARRAQLGEIF ELDRATLKSD GVFRSSPRGW FTFGHASFAL SVEQVGVTVE FYGGELNGVS YSDPATVKKY ARRAQLGEIF ELDRATLKSD GVFRSSPRGW FTFGHASFAL Rice SVEQVGVTVE FYGGELNGVS YSDPATVKKY ARRAQLGEIF ELDRATLKSD GVFRSSPRGW FTFGHATFAL SVEQVGVTVE FYGGELNGVS YSDPATVKKY ARRAQLGEIF ELDRATLKSD GVFRSSPRGW FTFGHATFAL Riverwort SVEQVGVTVE FYGGELNGVS YSDPATVKKY ARRAQLGEIF ELDRATLKSD GVFRSSPRGW FTFGHATFAL SVEQVGVTVE FYGGELNGVS YSDPATVKKY ARRAQLGEIF ELDRATLKSD GVFRSSPRGW FTFGHATFAL RIVErwort SVEQVGVTVE FYGGELNGVS YSDPATVKKY ARRAQLGEIF ELDRATLKSD GVFRSSPRGW FTFGHATFAL SVEQVGVTVE FYGGELNGVS FSDPATVKKY ARRAQLGEIF ELDRATLKSD GVFRSSPRGW FTFGHATFAL SVEQVGVTVE FYGGELNGVS FSDPATVKKY ARRAQLGEIF ELDRATLKSD GVFRSSPRGW FTFGHATFAL Chlamydom. SIEQVGVTVF FYGGELDGVS FSDPATVKKY ARRAQLGEIF EFDRATLKSD GVFRSSPRGW FTFGHATFAL SVEQVGVTVF FYGGELDGVS FNDPATVKKY ARRAQLGEIF EFDRATLKSD GVFRSS	Liverwort	KGGLFRAGAM	DNGDGIAVGW	LGHAVFKDKE	GNELFVRRMP	TFFETFPVVL	VDEQGIVRAD	VPFRRAESKY
Prochloro. KGGLFRVGAM DSGDGIAEEW LGHPVFQDGA GRALSVRRLP NFFETFPVIL TDADGVVRAD IPFRRESKY TDADGVVRAD IPFRESKY TDADGVVRAD IPFRRESK	Chlamydom.	KGGLFRTGAM	NSGDGIAVGW	LGHASFKDQE	GRELFVRRMP	TFFETFPVLL	LDKDGIVRAD	VPFRKAESKY
Anabaena KGGLFRTGPM VKGDGIAQSW QGHGVFKDAE GRELTVRRLP NFFETFPVIL TDADGVVRAD IPFRAESKY Syn.6803 KGGLFRTGAM NSCDGIAQEW IGHPIFKDKE GRELEVRRMP NFFETFPVIM TDADGVVRAD IPFRAESKY Consensus KGGLFR-G-M -GDGIAW -GHF-D G-L-VRR-P -FFE-FPVDG-VRAD -PFRES 391 401 411 421 431 441 451 Tobacco Sveqvgvtve Fyggelngvs ysdpatvkky ARRAQLGEIF ELDRATLKSD GVFRSSPRGW FTFGHASFAL Spinach Sveqvgvtve Fyggelngvs ysdpatvkky ARRAQLGEIF ELDRATLKSD GVFRSSPRGW FTFGHASFAL Maize Sveqvgvtve Fyggelngvs ysdpatvkky ARRAQLGEIF ELDRATLKSD GVFRSSPRGW FTFGHASFAL Rice Sveqvgvtve Fyggelngvs ysdpatvkky ARRAQLGEIF ELDRATLKSD GVFRSSPRGW FTFGHATFAL Rye/Wheat Sveqvgvtve Fyggelngvs ysdpatvkky ARRAQLGEIF ELDRATLKSD GVFRSSPRGW FTFGHATFAL Barley Sveqvgvtve Fyggelngvs ysdpatvkky ARRAQLGEIF ELDRATLKSD GVFRSSPRGW FTFGHATFAL Liverwort Sveqvgvtve Fyggelngvs ysdpatvkky ARRAQLGEIF ELDRATLKSD GVFRSSPRGW FTFGHATFAL Liverwort Sveqvgvtve Fyggelngvs FSDPATVKKY ARRAQLGEIF ELDRATLKSD GVFRSSPRGW FTFGHATFAL Chlamydom. SIEQvgvvtv Fyggeldglt FTDPATVKKY ARRAQLGEIF EFDRATLSD GVFRSSPRGW FTFGHATFAL Euglena SIEQvgvtv FYGGELNGVT FTDPATVKKY ARRAQLGEIF EFDRATLGSD GVFRSSPRGW FTFGHAVFAL Prochloro. SFEQTGVTVS FYGGALDGQT FTNPSDVKKF ARRAQLGEIF EFDRETLNSD GVFRSSPRGW FTFGHAVFAL Syn.6803 SWEQTGVTVS FYGGALDGQT FNPSDVKKF ARRAQLGEAF DFDTETINSD GVFRTSPRGW FTFGHAVFAL	Euglena	KGGLFRSGPM	NNGDGIATAW	LGHAVFIDKE	GNSLFVRRMP	TFFETFPVIL	LDQNGVVRAD	IPFRRAESKY
Syn.6803 KGGLFRTGAM NSCDGIAQEW IGHPIFKDKE GRELEVRRMP NFFETFPVIM TDADGVVRAD IPFRSESKF -CD-sensus KGGLFR-G-MCDGIAW -GHF-D G-L-VRR-P -FFE-FFVDG-VRAD -PFRES 391 401 411 421 431 441 451 Tobacco SveQvGvtve FyGGELNGVS YSDPATVKKY ARRAQLGEIF ELDRATLKSD GVFRSSPRGW FTFGHASFAL SPINACH SVEQVGVTVE FYGGELNGVS YSDPATVKKY ARRAQLGEIF ELDRATLKSD GVFRSSPRGW FTFGHASFAL ARRAQLGEIF ELDRATLKSD GVFRSSPRGW FTFGHASFAL GVFRSSPRGW FTFGHAFFAL SVEQVGVTVE FYGGELNGVS YSDPATVKKY ARRAQLGEIF ELDRATLKSD GVFRSSPRGW FTFGHAFFAL Rye/Wheat SVEQVGVTVE FYGGELNGVS YSDPATVKKY ARRAQLGEIF ELDRATLKSD GVFRSSPRGW FTFGHATFAL GVFRSSPRGW FTFGHATFAL SVEQVGVTVE FYGGELNGVS YSDPATVKKY ARRAQLGEIF ELDRATLKSD GVFRSSPRGW FTFGHATFAL Barley SVEQVGVTVE FYGGELNGVS YSDPATVKKY ARRAQLGEIF ELDRATLKSD GVFRSSPRGW FTFGHATFAL Liverwort SVEQVGVTVE FYGGELNGVS YSDPATVKKY ARRAQLGEIF ELDRATLKSD GVFRSSPRGW FTFGHATFAL Chlamydom. SIEQVGVTVE FYGGELDGUS FSDPATVKKY ARRAQLGEIF ELDRATLKSD GVFRSSPRGW FTFGHATFAL GVFRSSPRGW FTFGHATFAL FYGGELNGVS FYGGELDGUS FSDPATVKKY ARRAQLGEIF EFDRATLKSD GVFRSSPRGW FTFGHATFAL GVFRSSPRGW FTFGHATF	Prochloro.	KGGLFRVGAM	DSGDGIAEEW	LGHPVFQDGA	GRALSVRRLP	NFFENFPVIL	TDGDGVVRAD	IPFRRSESQY
Consensus KGGLFR-G-M -GDGIA-W -GH-F-D- G-L-VRR-P -FFE-FFVDG-VRAD -PFR-ES 391 401 411 421 431 441 451 Tobacco SveQvgvtve FyggeLngvs ysdpatvkky arraqlgeif eldratlksd gvfrssprgw fffghasfal Spinach SveQvgvtve fyggeLngvs ysdpatvkky arraqlgeif eldratlksd gvfrssprgw fffghasfal Maize SveQvgvtve fyggeLngvs ysdpatvkky arraqlgeif eldratlksd gvfrssprgw fffghasfal Rice SveQvgvtve fyggeLngvs ysdpatvkky arraqlgeif eldratlksd gvfrssprgw fffghatfal Rye/Wheat SveQvgvtve fyggeLngvs ysdpatvkky arraqlgeif eldratlksd gvfrssprgw fffghatfal Barley SveQvgvtve fyggeLngvs ysdpatvkky arraqlgeif eldratlksd gvfrssprgw fffghatfal Liverwort SveQvgvtve fyggeLngvn ysdpatvkky arraqlgeif eldratlksd gvfrssprgw fffghatfal Chlamydom. SieQvgvtve fyggeLngvn ysdpatvkky arraqlgeif efdratlksd gvfrssprgw fffghatfal Euglena SieQvgvtvr fyggeLngvr ffdpatvkky arraqlgeif efdratlksd gvfrssprgw fffghatfal Prochloro. SfeQtgvtvr fyggeLngvr ffnpatvkky arraqlgeif efdratlgsd gvfrssprgw fffghatfal Prochloro. SfeQtgvtvs fyggaldgqt ffnpsdvkky arraqlgeif Dfnrstlqsd gvfrssprgw fffghatfal Anabaena SfeQsgvtvs fyggaldgqt ffnpsdvkky arraqlgeif efdretlnsd gvfrssprgw fffghavfal Syn.6803 SveQtgvtvs fyggaldgqt ffnpsdvkky arraqlgeif efdretlnsd gvfrtsprgw fffghavfal	Anabaena	KGGLFRTGPM	VKGDGIAQSW	QGHGVFKDAE	GRELTVRRLP	NFFETFPVIL	TDADGVVRAD	IPFRRAESKY
391 401 411 421 431 441 451 Tobacco Sveqvgvtve Fyggelngvs ysdpatvkky arraqlgeif eldratlksd gvfrssprgw ftfghasfal spinach sveqvgvtve fyggelngvs ysdpatvkky arraqlgeif eldratlksd gvfrssprgw ftfghasfal sveqvgvtve fyggelngvs ysdpatvkky arraqlgeif eldratlksd gvfrssprgw ftfghasfal arraqlgeif eldratlksd gvfrssprgw ftfghasfal gvfrce sveqvgvtve fyggelngvs ysdpatvkky arraqlgeif eldratlksd gvfrssprgw ftfghatfal arraqlgeif eldratlksd gvfrssprgw ftfghatfal gvfwheat sveqvgvtve fyggelngvs ysdpatvkky arraqlgeif eldratlksd gvfrssprgw ftfghatfal arraqlgeif eldratlksd gvfrssprgw ftfghatfal gveywheat sveqvgvtve fyggelngvs ysdpatvkky arraqlgeif eldratlksd gvfrssprgw ftfghatfal liverwort sveqvgvtve fyggelngvn ysdpatvkky arraqlgeif eldratlksd gvfrssprgw ftfghatfal chlamydom. Sieqvgvtve fyggeldgt ftdpatvkky arraqlgeif efdratlksd gvfrssprgw ftfghatfal elglena sieqvgvtvr fyggeldgt ftdpatvkky arraqlgeif efdratlksd gvfrssprgw ftfghatfal prochloro. Sfeqtgvtvs fyggeldgt ftdpatvkky arraqlgeif efdratlgsd gvfrssprgw ftfghatfal prochloro. Sfeqtgvtvs fyggaldgqt ftnpsdvkky arraqlgeif bfnrstlqsd gvfrssprgw ftfghatfal arraqlgein sfeqtgrt sfrghatfal gvfrssprgw ftfghatfal arraqlgeif bfnrstlqsd gvfrssprgw ftfghatfal arraqlgeif sfortligsd gvfrssprgw ftfghatfal gvfrshatfal gvfrstrgw ftfghatfal arraqlgeif bfnrstlqsd gvfrssprgw ftfghatfal gvfrshatfal gvfrstrgw ftfghatfal arraqlgeif bfnrstlqsd gvfrssprgw ftfghatfal gvfrshatfal gvfrs	Syn.6803	KGGLFRTGAM	NEGDGIAQEW	IGHPIFKDKE	GRELEVRRMP	NFFETFPVIM	TDADGVVRAD	IPFRRSESKF
Tobacco SVEQVGVTVE FYGGELNGVS YSDPATVKKY ARRAQLGEIF ELDRATLKSD GVFRSSPRGW FTFGHASFAL Spinach SVEQVGVTVE FYGGELNGVS YSDPATVKKY ARRAQLGEIF ELDRATLKSD GVFRSSPRGW FTFGHASFAL Oenothera SVEQVGVTVE FYGGELNGVS YSDPATVKKY ARRAQLGEIF ELDRATLKSD GVFRSSPRGW FTFGHASFAL Rice SVEQVGVTVE FYGGELNGVS YSDPATVKKY ARRAQLGEIF ELDRATLKSD GVFRSSPRGW FTFGHATFAL RYE/Wheat SVEQVGVTVE FYGGELNGVS YSDPATVKKY ARRSQLGEIF ELDRATLKSD GVFRSSPRGW FTFGHATFAL Barley SVEQVGVTVE FYGGELNGVS YSDPATVKKY ARRSQLGEIF ELDRATLKSD GVFRSSPRGW FTFGHATFAL Liverwort SVEQVGVTVE FYGGELNGVN YSDPATVKKY ARRSQLGEIF ELDRATLKSD GVFRSSPRGW FTFGHATFAL Chlamydom. SIEQVGVTVE FYGGELDGVS FSDPATVKKY ARRAQLGEIF EFDRATLKSD GVFRSSPRGW FTFGHATFAL Euglena SIEQVGVTVF FYGGELDGLT FTDPATVKKY ARRAQLGEIF EFDRSTLQSD GVFRSSPRGW FTFGHLSFAL Prochloro. SFEQTGVTVS FYGGALDGQT FTNPSDVKKF ARRAQLGEIF DFNRSTLQSD GVFRSSPRGW FTFGHLSFAL ARRAQLGEIF SFDRSTLQSD GVFRSSPRGW FTFGHLSFAL Syn.6803 SVEQTGVTVS FYGGALDGQT FTNPSDVKKF ARRAQLGEIF EFDRSTLQSD GVFRSSPRGW FTFGHAVFAL Syn.6803 SVEQTGVTVS FYGGALDGQT FSNPSDVKKF ARRAQLGEIF EFDRSTLNSD GVFRTSPRGW FTFGHAVFAL SYN.6803 SVEQTGVTVS FYGGALDGQT FSNPSDVKKF ARRAQLGEGF DFDTETLGSD GVFRTSPRGW FTFGHAVFAL SYN.6803 SVEQTGVTVS FYGGALDGQT FSNPSDVKKF ARRAQLGEGF DFDTETLNSD GVFRTSPRGW FTFGHAVFAL	Consensus	KGGLFR-G-M	GDGIAW	-GHF-D	G-L-VRR-P	-FFE-FPV	-DG-VRAD	-PFRES
Tobacco SVEQVGVTVE FYGGELNGVS YSDPATVKKY ARRAQLGEIF ELDRATLKSD GVFRSSPRGW FTFGHASFAL Spinach SVEQVGVTVE FYGGELNGVS YSDPATVKKY ARRAQLGEIF ELDRATLKSD GVFRSSPRGW FTFGHASFAL Oenothera SVEQVGVTVE FYGGELNGVS YSDPATVKKY ARRAQLGEIF ELDRATLKSD GVFRSSPRGW FTFGHASFAL Rice SVEQVGVTVE FYGGELNGVS YSDPATVKKY ARRAQLGEIF ELDRATLKSD GVFRSSPRGW FTFGHATFAL RYE/Wheat SVEQVGVTVE FYGGELNGVS YSDPATVKKY ARRSQLGEIF ELDRATLKSD GVFRSSPRGW FTFGHATFAL Barley SVEQVGVTVE FYGGELNGVS YSDPATVKKY ARRSQLGEIF ELDRATLKSD GVFRSSPRGW FTFGHATFAL Liverwort SVEQVGVTVE FYGGELNGVN YSDPATVKKY ARRSQLGEIF ELDRATLKSD GVFRSSPRGW FTFGHATFAL Chlamydom. SIEQVGVTVE FYGGELDGVS FSDPATVKKY ARRAQLGEIF EFDRATLKSD GVFRSSPRGW FTFGHATFAL Euglena SIEQVGVTVF FYGGELDGLT FTDPATVKKY ARRAQLGEIF EFDRSTLQSD GVFRSSPRGW FTFGHLSFAL Prochloro. SFEQTGVTVS FYGGALDGQT FTNPSDVKKF ARRAQLGEIF DFNRSTLQSD GVFRSSPRGW FTFGHLSFAL ARRAQLGEIF SFDRSTLQSD GVFRSSPRGW FTFGHLSFAL Syn.6803 SVEQTGVTVS FYGGALDGQT FTNPSDVKKF ARRAQLGEIF EFDRSTLQSD GVFRSSPRGW FTFGHAVFAL Syn.6803 SVEQTGVTVS FYGGALDGQT FSNPSDVKKF ARRAQLGEIF EFDRSTLNSD GVFRTSPRGW FTFGHAVFAL SYN.6803 SVEQTGVTVS FYGGALDGQT FSNPSDVKKF ARRAQLGEGF DFDTETLGSD GVFRTSPRGW FTFGHAVFAL SYN.6803 SVEQTGVTVS FYGGALDGQT FSNPSDVKKF ARRAQLGEGF DFDTETLNSD GVFRTSPRGW FTFGHAVFAL						-		
Spinach SVEQVGVTVE FYGGELNGVS YSDPATVKKY ARRAQLGEIF ELDRATLKSD GVFRSSPRGW FTFGHASFAL Maize SVEQVGVTVE FYGGELNGVS YSDPATVKKY ARRAQLGEIF ELDRATLKSD GVFRSSPRGW FTFGHASFAL Rice SVEQVGVTVE FYGGELNGVS YSDPATVKKY ARRAQLGEIF ELDRATLKSD GVFRSSPRGW FTFGHATFAL Rye/Wheat SVEQVGVTVE FYGGELNGVS YSDPATVKKY ARRAQLGEIF ELDRATLKSD GVFRSSPRGW FTFGHATFAL Barley SVEQVGVTVE FYGGELNGVS YSDPATVKKY ARRSQLGEIF ELDRATLKSD GVFRSSPRGW FTFGHATFAL Liverwort SVEQVGVTVE FYGGELNGVN YSDPATVKKY ARRSQLGEIF ELDRATLKSD GVFRSSPRGW FTFGHATFAL Chlamydom. SIEQVGVTVE FYGGELDGVS FSDPATVKKY ARRAQLGEIF EFDRATLKSD GVFRSSPRGW FTFGHATFAL Chlamydom. SIEQVGVTVE FYGGELDGLT FTDPATVKKY ARRAQLGEIF EFDRSTLQSD GVFRSSPRGW FTFGHATFAL Euglena SIEQVGVTVR FFGGSFDTLS FNDPATVKRY ARRAQLGEIF DFNRSTLQSD GVFRSSPRGW FTFGHLSFAL Prochloro. SFEQTGVTVS FYGGALDGQT FTNPSDVKKF ARRAQLGEIF DFDTETLGSD GVFRTSPRGW FTFGHAVFAL Syn.6803 SVEQTGVTVS FYGGALDGQT FSNPSDVKKF ARRAQLGEIF EFDRETLNSD GVFRTSPRGW FTFGHAVFAL		391	401	411	421	431	441	451
Oenothera SVEQVGVTVE FYGGELNGVS YSDPATVKKY ARRAQLGEIF ELDRATLKSD GVFRSSPRGW FTFGHASFAL Maize SVEQVGVTVE FYGGELNGVS YSDPATVKKY ARRAQLGEIF ELDRATLKSD GVFRSSPRGW FTFGHATFAL Rice SVEQVGVTVE FYGGELNGVS YSDPATVKKY ARRSQLGEIF ELDRATLKSD GVFRSSPRGW FTFGHATFAL Rye/Wheat SVEQVGVTVE FYGGELNGVS YSDPATVKKY ARRSQLGEIF ELDRATLKSD GVFRSSPRGW FTFGHATFAL Barley SVEQVGVTVE FYGGELNGVN YSDPATVKKY ARRSQLGEIF ELDRATLKSD GVFRSSPRGW FTFGHATFAL Liverwort SVEQVGVTVE FYGGELDGVS FSDPATVKKY ARRAQLGEIF EFDRATLKSD GVFRSSPRGW FTFGHATFAL Chlamydom. SIEQVGVTVE FYGGELDGUS FTDPATVKKY ARRAQLGEIF EFDRATLKSD GVFRSSPRGW FTFGHATFAL Euglena SIEQVGVTVR FFGGSFDTLS FNDPATVKRY ARRAQLGEIF EFDRSTLQSD GVFRSSPRGW FTFGHACFAL Prochloro. SFEQTGVTVS FYGGALDGQT FTNPSDVKKF ARRAQLGEAF DFDTETLGSD GVFRSSPRGW FTFGHACFAL Anabaena SFEQSGVTVS FYGGALDGQT FTNPSDVKKF ARRAQLGEIF EFDRETLNSD GVFRTSPRGW FTFGHAVFAL Syn.6803 SVEQTGVTVS FYGGALDGQT FSNPSDVKKF ARRAQLGEF DFDTETLNSD GVFRTSPRGW FTFGHAVFAL	Tobacco	SVEQVGVTVE	FYGGELNGVS	YSDPATVKKY	ARRAQLGEIF	ELDRATLKSD	GVFRSSPRGW	FTFGHASFAL
Maize SVEQVGVTVE FYGGELNGVS YSDPATVKKY ARRAQLGEIF ELDRATLKSD GVFRSSPRGW FTFGHATFAL Rice SVEQVGVTVE FYGGELNGVS YSDPATVKKY ARRSQLGEIF ELDRATLKSD GVFRSSPRGW FTFGHATFAL Rye/Wheat SVEQVGVTVE FYGGELNGVS YSDPATVKKY ARRSQLGEIF ELDRATLKSD GVFRSSPRGW FTFGHATFAL Barley SVEQVGVTVE FYGGELNGVN YSDPATVKKY ARRSQLGEIF ELDRATLKSD GVFRSSPRGW FTFGHATFAL Liverwort SVEQVGVTVE FYGGELDGVS FSDPATVKKY ARRAQLGEIF EFDRATLKSD GVFRSSPRGW FTFGHATFAL Chlamydom. SIEQVGVVVT FYGGELDGLT FTDPATVKKY ARRAQLGEIF EFDRSTLQSD GVFRSSPRGW FTFGHAVFAL Euglena SIEQVGVTVR FFGGSFDTLS FNDPATVKRY ARRAQLGEIF DFNRSTLQSD GVFRSSPRGW FTFGHLSFAL Prochloro. SFEQTGVTVS FYGGALDGQT FTNPSDVKKF ARRAQLGEIF DFDTETLGSD GVFRSSPRGW FTFGHAVFAL Anabaena SFEQSGVTVS FYGGALDGQT FSNPSDVKKF ARRAQLGEIF EFDRETLNSD GVFRTSPRGW FTFGHAVFAL Syn.6803 SVEQTGVTVS FYGGALDGQT FSNPSDVKKF ARRAQLGEGF DFDTETFNSD GVFRTSPRGW FTFGHAVFAL	Spinach	SVEQVGVTVE	FYGGELNGVS	YSDPATVKKY	ARRAQLGEIF	ELDRATLKSD	GVFRSSPRGW	FTFGHASFAL
Rice SVEQVGVTVE FYGGELNGVS YSDPATVKKY ARRSQLGEIF ELDRATLKSD GVFRSSPRGW FTFGHATFAL Rye/Wheat SVEQVGVTVE FYGGELNGVS YSDPATVKKY ARRSQLGEIF ELDRATLKSD GVFRSSPRGW FTFGHATFAL Barley SVEQVGVTVE FYGGELNGVN YSDPATVKKY ARRSQLGEIF ELDRATLKSD GVFRSSPRGW FTFGHATFAL Liverwort SVEQVGVTVE FYGGELDGVS FSDPATVKKY ARRAQLGEIF EFDRATLKSD GVFRSSPRGW FTFGHATFAL Chlamydom. SIEQVGVTVE FYGGELDGLT FTDPATVKKY ARRAQLGEIF EFDRSTLQSD GVFRSSPRGW FTFGHATFAL Euglena SIEQVGVTVR FFGGSFDTLS FNDPATVKRY ARRAQLGEIF DFNRSTLQSD GVFRSSPRGW FTFGHLSFAL Prochloro. SFEQTGVTVS FYGGALDGQT FTNPSDVKKF ARRAQLGEIF DFDTETLGSD GVFRTSTRGW FTFGHACFAL ANAbaena SFEQSGVTVS FYGGALDGQT FTNPSDVKKF ARRAQLGEIF EFDRETLNSD GVFRTSPRGW FTFGHAVFAL Syn.6803 SVEQTGVTVS FYGGALDGQT FSNPSDVKKF ARRAQLGEGF DFDTETFNSD GVFRTSPRGW FTFGHAVFAL	Oenothera	SVEQVGVTVE	FYGGELNGVS	YSDPATVKKY	ARRAQLGEIF	ELDRATLKSD	GVFRSSPRGW	FTFGHASFAL
Rye/Wheat SVEQVGVTVE FYGGELNGVS YSDPATVKKY ARRSQLGEIF ELDRATLKSD GVFRSSPRGW FTFGHATFAL Barley SVEQVGVTVE FYGGELNGVN YSDPATVKKY ARRSQLGEIF ELDRATLKSD GVFRSSPRGW FTFGHATFAL Liverwort SVEQVGVTVE FYGGELDGVS FSDPATVKKY ARRAQLGEIF EFDRATLKSD GVFRSSPRGW FTFGHATFAL Chlamydom. SIEQVGVVVF FYGGELDGLT FTDPATVKKY ARRAQLGEIF EFDRSTLQSD GVFRSSPRGW FTFGHAVFAL Euglena SIEQVGVTVR FFGGSFDTLS FNDPATVKRY ARRAQLGEIF DFNRSTLQSD GVFRSSPRGW FTFGHLSFAL Prochloro. SFEQTGVTVS FYGGALDGQT FTNPSDVKKF ARRAQLGEAF DFDTETLGSD GVFRTSTRGW FTFGHACFAL Anabaena SFEQSGVTVS FYGGALDGQT FTNPSDVKKF ARRAQLGEAF DFDTETLNSD GVFRTSPRGW FTFGHAVFAL Syn.6803 SVEQTGVTVS FYGGALDGQT FSNPSDVKKF ARRAQLGEGF DFDTETFNSD GVFRTSPRGW FTFGHAVFAL	Maize	SVEQVGVTVE	FYGGELNGVS	YSDPATVKKY	ARRAQLGEIF	ELDRATLKSD	GVFRSSPRGW	FTFGHATFAL
Barley SVEQVGVTVE FYGGELNGVN YSDPATVKKY ARRSQLGEIF ELDRATLKSD GVFRSSPRGW FTFGHATFAL Liverwort SVEQVGVTVE FYGGELDGVS FSDPATVKKY ARRAQLGEIF EFDRATLKSD GVFRSSPRGW FTFGHATFAL Chlamydom. SIEQVGVVV FYGGELDGLT FTDPATVKKY ARRAQLGEIF EFDRSTLQSD GVFRSSPRGW FTFGHVCFAL Euglena SIEQVGVTVR FFGGSFDTLS FNDPATVKRY ARRAQLGEIF DFNRSTLQSD GVFRSSPRGW FTFGHLSFAL Prochloro. SFEQTGVTVS FYGGALDGQT FTNPSDVKKF ARRAQLGEAF DFDTETLGSD GVFRTSTRGW FTFGHACFAL Anabaena SFEQSGVTVS FYGGDLDGKT FTDPADVKKY ARRAQLGEIF EFDRETLNSD GVFRTSPRGW FTFGHAVFAL Syn.6803 SVEQTGVTVS FYGGALDGQT FSNPSDVKKF ARRAQLGEGF DFDTETFNSD GVFRTSPRGW FTFGHAVFAL	Rice	SVEQVGVTVE	FYGGELNGVS	YSDPATVKKY	ARRSQLGEIF	ELDRATLKSD	GVFRSSPRGW	FTFGHATFAL
Liverwort SVEQVGVTVE FYGGELDGVS FSDPATVKKY ARRAQLGEIF EFDRATLKSD GVFRSSPRGW FTFGHATFAL Chlamydom. SIEQVGVSVT FYGGELDGLT FTDPATVKKY ARKAQLGEIF EFDRSTLQSD GVFRSSPRGW FTFGHVCFAL Euglena SIEQVGVTVR FFGGSFDTLS FNDPATVKRY ARHAQLGEIF DFNRSTLQSD GVFRSSPRGW FTFGHLSFAL Prochloro. SFEQTGVTVS FYGGALDGQT FTNPSDVKKF ARRAQLGEAF DFDTETLGSD GVFRTSTRGW FTFGHACFAL Anabaena SFEQSGVTVS FYGGDLDGKT FTDPADVKKY ARKAQLGEIF EFDRETLNSD GVFRTSPRGW FTFGHAVFAL Syn.6803 SVEQTGVTVS FYGGALDGQT FSNPSDVKKF ARKAQLGEGF DFDTETFNSD GVFRTSPRGW FTFGHAVFAL	Rye/Wheat	SVEQVGVTVE	FYGGELNGVS	YSDPATVKKY	ARRSQLGEIF	ELDRATLKSD	GVFRSSPRGW	FTFGHATFAL
Chlamydom. SIEQVGVSVT FYGGELDGLT FTDPATVKKY ARKAQLGEIF EFDRSTLQSD GVFRSSPRGW FTFGHVCFAL Euglena SIEQVGVTVR FFGGSFDTLS FNDPATVKRY ARHAQLGEIF DFNRSTLQSD GVFRSSPRGW FTFGHLSFAL Prochloro. SFEQTGVTVS FYGGALDGQT FTNPSDVKKF ARRAQLGEAF DFDTETLGSD GVFRTSTRGW FTFGHACFAL Anabaena SFEQSGVTVS FYGGDLDGKT FTDPADVKKY ARKAQGGEIF EFDRETLNSD GVFRTSPRGW FTFGHAVFAL Syn.6803 SVEQTGVTVS FYGGALDGQT FSNPSDVKKF ARKAQLGEGF DFDTETFNSD GVFRTSPRGW FTFGHAVFAL	Barley	SVEQVGVTVE	FYGGELNGVN	YSDPATVKKY	ARRSQLGEIF	ELDRATLKSD	GVFRSSPRGW	FTFGHATFAL
Euglena SIEQVGVTVR FFGGSFDTLS FNDPATVKRY ARHAQLGEIF DFNRSTLQSD GVFRSSPRGW FTFGHLSFAL Prochloro. SFEQTGVTVS FYGGALDGQT FTNPSDVKKF ARRAQLGEAF DFDTETLGSD GVFRTSTRGW FTFGHACFAL Anabaena SFEQSGVTVS FYGGDLDGKT FTDPADVKKY ARKAQGGEIF EFDRETLNSD GVFRTSPRGW FTFGHAVFAL Syn.6803 SVEQTGVTVS FYGGALDGQT FSNPSDVKKF ARKAQLGEGF DFDTETFNSD GVFRTSPRGW FTFGHAVFAL	Liverwort	SVEQVGVTVE	FYGGELDGVS	FSDPATVKKY	ARRAQLGEIF	EFDRATLKSD	GVFRSSPRGW	FTFGHATFAL
Prochloro. SFEQTGVTVS FYGGALDGQT FTNPSDVKKF ARRAQLGEAF DFDTETLGSD GVFRTSTRGW FTFGHACFAL Anabaena SFEQSGVTVS FYGGDLDGKT FTDPADVKKY ARKAQGGEIF EFDRETLNSD GVFRTSPRGW FTFGHAVFAL Syn.6803 SVEQTGVTVS FYGGALDGQT FSNPSDVKKF ARKAQLGEGF DFDTETFNSD GVFRTSPRGW FTFGHAVFAL	Chlamydom.	SIEQVGVSVT	FYGGELDGLT	FTDPATVKKY	ARKAQLGEIF	EFDRSTLQSD	GVFRSSPRGW	FTFGHVCFAL
Anabaena SFEQSGVTVS FYGGDLDGKT FTDPADVKKY ARKAQGGEIF EFDRETLNSD GVFRTSPRGW FTFGHAVFAL Syn.6803 SVEQTGVTVS FYGGALDGQT FSNPSDVKKF ARKAQLGEGF DFDTETFNSD GVFRTSPRGW FTFGHAVFAL	Euglena	SIEQVGVTVR	FFGGSFDTLS	FNDPATVKRY	ARHAQLGEIF	DFNRSTLQSD	GVFRSSPRGW	FTFGHLSFAL
Syn.6803 SVEQTGVTVS FYGGALDGQT FSNPSDVKKF ARKAQLGEGF DFDTETFNSD GVFRTSFRGW FTFGHAVFAL	Prochloro.	SFEQTGVTVS	FYGGALDGQT	FTNPSDVKKF	ARRAQLGEAF	DFDTETLGSD	GVFRTSTRGW	FTFGHACFAL
	Anabaena	SFEQSGVTVS	FYGGDLDGKT	FTDPADVKKY	ARKAQGGEIF	EFDRETLNSD	GVFRTSPRGW	FTFGHAVFAL
Consensus SEQ-GV-V- F-GGP-VK ARQ-GE-FTSD GVFR-S-RGW FTFGHFAL	Syn.6803	SVECTGVTVS	FYGGALDGQT	FSNPSDVKKF	ARKAQLGEGF	DFDTETFNSD	GVFRTSPRGW	FTFGHAVFAL
	Consensus	SEQ-GV-V-	F-GG	P-VK	ARQ-GE-F	TSD	GVFR-S-RGW	FTFGHFAL

FIGURE 1: Alignment of the amino acid sequence from residue 251 to residue 460 of CP47 from twelve different organisms, including cyanobacteria, algae, bryophytes, and monocotyledonous and dicotyledonous spermatophytes. The bottom line shows the consensus of all sequences listed. Nonconserved residues are replaced by dashes. The locations of the twelve deletions presented in this study are indicated by solid-line frames; the locations of two deletions presented in a previous study (Eaton-Rye & Vermaas, 1991) are indicated by dashed-line frames. References for the sequences: tobacco, Shinozaki et al. (1986); spinach, Morris and Herrmann (1984); Oenothera, Offermann-Steinhard and Herrmann (1990); maize, Rock et al. (1987); rice, Hiratsuka et al. (1989); rye, Bukharov et al. (1988); wheat, Hird et al. (1991); barley, Andreeva et al. (1989); liverwort, Ohyama et al. (1986); Chlamydomonas reinhardtii, Berry-Lowe et al. (1992); Euglena gracilis, Keller et al. (1989); Prochlorothrix hollandica, Greer and Golden (1991); Anabaena 7120, Lang and Haselkorn, (1989); and Synechocystis 6803, Vermaas et al. (1987).

Table II: Compilation of the Characterization of the Wild Type, the Deletion Mutants $\Delta 1-\Delta 12$, and the Single-Site Mutant H343L^a

	mutation		phototrophic growth ^b (doubling time/h)		abundance of PSII proteins ^c (immunoblot)				PSII activity ^d	herbicide bindinge
	deletion	length	-glucose	+glucose	D1	D2	CP47	CP43	$(O_2 \text{ rate}, \%)$	(Chl/PSII)
WT			11–12	11–12	++	++	++	++	100′	750–850
$\Delta 1$	I265-F268	4	*	16–18	*	*	*	+	*	*
$\Delta 2$	T271-K277	7	*	21-23	(+)	(+)	*	+	*	*
$\Delta 3$	K277-E283	7	41-45	11-12	÷´	÷´	(+)	+	25-358	3000-4500
$\Delta 4$	T304-L309	6	*	17-20	*	*	*	+	*	*
$\Delta 5$	F311-N317	7	*	17-19	*	•	*	+	*	•
$\Delta 6$	G333-I336	4	12-13	11-12	++	++	++	++	80–90	1200-1500
$\Delta 7$	K347-R352	6	13-14	11-12	++	++	+	++	85-95	1000-1400
$\Delta 8$	A373-D380	8	12-13	10-13	++	++	++	++	50-60g	1000-1200
$\Delta 9$	V392-Q394	3	10-11	12-13	++	++	++	++	85-95	900-1000
$\Delta 10$	D416-F420	5	13-14	11-12	++	++	+	++	≤80	2000-2500
$\Delta 11$	R422-E428	7	>60	12-14	+	+	(+)	+	20-308	5600-7500
$\Delta 12$	D440-P447	8	*	19-22	*	*	* í	++	*	*
H343L		·	16–19	11–12	++	++	+	++	75–85	1400-1500

^a All results listed represent an average of at least three experiments. ^b The doubling times were calculated from growth curves taken in the absence and presence of 5 mM glucose. ^c The abundance of PS II core proteins D1, D2, CP43, and CP47 was estimated from immunoreactions on western blots; (++) wild type level or close to it, (+) noticeably reduced level, ((+)) strikingly reduced level, (*) barely detectable. ^d The oxygen evolution rate was measured with whole cells. ^e Chl/PS II was determined by [14 C]diuron binding in cell suspensions with 50 μ g of Chl/mL. ^f The wild-type rate was in the range of 475 ± 25 μ mol of O₂/[(mg of Chl)·h]. ^g Unstable oxygen evolution rate, with half-times of 120 s for mutants Δ 3 and Δ 8, 75 s for mutant Δ 11, and >300 s for WT cells at 5000 μ E/(m²s) illumination.

medium without addition of glucose. The observed doubling times are listed in Table II. Two interesting aspects stand out: (i) the effect on photoautotrophic growth depends on the position of the deletion, but is independent of its length (compare $\Delta 1$ with $\Delta 6$, $\Delta 4$ with $\Delta 7$, and $\Delta 12$ with $\Delta 8$, the pairs carrying deletions of four, six, and eight amino acid residues, respectively); and (ii) except $\Delta 3$, $\Delta 11$, and H343L, all mutant strains can be roughly divided into two catagories, those exhibiting growth rates close to wild type and those unable to grow in the absence of glucose. The latter group of obligate photoheterotrophs is composed of mutant strains $\Delta 1$, $\Delta 2$, $\Delta 4$, $\Delta 5$, and $\Delta 12$. In contrast, both mutants $\Delta 3$ and $\Delta 11$ exhibit a very slow photoautotrophic growth with initial doubling times of 43 and >60 h, respectively. However, under our growth conditions $\Delta 3$ and $\Delta 11$ perform only about one doubling cycle before further cell propagation is stopped. It is conceivable that this is due to residual glucose since the cultures assayed for photoautotrophic growth were inoculated from glucose-grown precultures. However, precultures of all mutant strains tested were thoroughly washed before subjecting them to photoautotrophic growth conditions, and the behavior of the obligate photoheterotrophic mutant strains $\Delta 1$, $\Delta 2$, $\Delta 4$, $\Delta 5$, and $\Delta 12$ does not indicate residual internal or external glucose. Therefore, the above explanation seems less likely and would imply a different internal glucose storage capacity. At present, the nature of the factor limiting photoautotrophic growth in mutant strains $\Delta 3$ and $\Delta 11$ remains obscure. Figure 2 shows representative growth curves for the wild type, the impaired $\Delta 3$ and $\Delta 11$ mutants, and the obligate photoheterotrophic $\Delta 12$ mutant.

The findings show that differential effects, from nonsignificant to severe, can be observed upon deletion of several amino acid residues from loop E, which is presumably exposed to the lumenal side according to hydropathy analysis of the CP47 sequence (Vermaas et al., 1987; Bricker, 1990). On the basis of cross-linking experiments CP47 is inferred to bind the extrinsic PS II-O polypeptide (Odom & Bricker, 1992; Enami et al., 1992), which plays an important regulatory role for the water oxidase. Therefore, it was interesting to determine whether mutations in loop E affected the interaction of CP47 with PS II-O. According to a study of Philbrick et al. (1991), psbO mutants of Synechocystis sp. PCC 6803 exhibit only slightly reduced growth rates in normal BG-11 medium but are unable to grow in Ca²⁺-depleted media. Thus, investigation of the growth characteristics in Ca²⁺-depleted

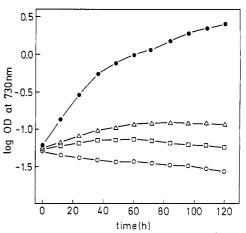


FIGURE 2: Photoautotrophic growth curve of wild-type (\bullet), mutant $\Delta 3$ (Δ), mutant $\Delta 11$ (\square), and mutant $\Delta 12$ (\bigcirc) cells as shown by semilogarithmic plots of the optical density measured at 730 nm of cultures grown in the absence of glucose.

BG-11 could indicate whether a functional binding of the PS II-O protein has been disturbed due to any of the short deletions introduced into the large extrinsic loop of CP47. The growth rates of $\Delta 3$, $\Delta 6$, $\Delta 7$, $\Delta 8$, $\Delta 9$, and $\Delta 10$ were assayed under omission of Ca²⁺ from the medium. A psbO⁻ mutant strain, kindly provided by Dr. B. Zilinskas, was run as a control and behaved as described previously (Philbrick et al., 1991). However, it was found that none of the CP47 mutants tested showed any change in their ability to grow photoautotrophically. Doubling times in the absence or presence of Ca²⁺ were virtually identical (not shown). This suggests that the PS II-O polypeptide is still functionally bound in all photoautotrophic mutants. In addition, it was tested whether any of the photoautotrophic deletion mutants exhibited an increased demand for manganese, for example, due to unstable binding of the manganese cluster. For these experiments, the mutants $\Delta 3$, $\Delta 6$, $\Delta 7$, $\Delta 8$, $\Delta 9$, and $\Delta 10$ were grown in BG-11 medium with a 20-fold reduction in manganese concentration. Again, no significant response to a reduction of the exogenous manganese concentration was observed.

Abundance of the Main Polypeptides and Stability of the PS II Complex. The relative abundance of the four major membrane-integral PS II core proteins was tested by immunoblot analysis employing antibodies raised against D1, D2, CP47, and CP43. To avoid protein degradation, thylakoids

subjected to SDS-polyacrylamide gel electrophoresis were prepared in the presence of protease inhibitors (1 mM PMSF, 2 mM benzamidine, 2 mM ϵ -aminocaproic acid). The results of the immunoblot analysis (see Table II) reveal a good correlation between the abundance of CP47 and the reaction center proteins D1 and D2 in agreement with previous findings (Vermaas et al., 1988). A reduced level of CP47 in psbB mutant thylakoids is probably due to an increased rate of protein degradation caused by the failure of correct assembly with the reaction center proteins (Yu & Vermaas, 1990). In general, the lack of stable assembly leads to degradation of CP47, D1, and D2, with a decrease in the level of CP47 normally paralleled by a corresponding decrease of D1 and D2. However, in some mutant thylakoids, especially $\Delta 7$, $\Delta 10$, and H343L, the relative level of CP47 seems to be markedly reduced compared to that of other PS II proteins. This observation could be explained by either (1) an altered antigenicity caused by the introduced mutation; (2) an increased instability of CP47 during thylakoid preparation despite the presence of protease inhibitors; or (3) the possibility that, after assembly, CP47 is more unstable than D1 and D2 in these mutants, leading to a relative decrease in the amount of CP47. At present, we cannot distinguish between these possibilities. In contrast to CP47, D1, and D2, the abundance of CP43 appears to be largely independent of the presence of other PS II core proteins. As shown in Table II, CP43 is present at levels close to wild type in all mutant thylakoids. This is in agreement with earlier data obtained with Synechocystis sp. PCC 6803 (Vermaas et al., 1986, 1988) and is somewhat similar to the situation in Chlamydomonas reinhardtii where low levels of CP43 can accumulate in thylakoids that exhibit missing or very low levels of D1, D2, and CP47 (de Vitry et al., 1989). It is important to note that D1, D2, and CP47 are virtually absent in the fully photoheterotrophic mutants $\Delta 1$, $\Delta 2$, $\Delta 4$, $\Delta 5$, and $\Delta 12$. This suggests that the reaction center complex is very unstable in these mutants.

The relative concentration of PS II centers in the thylakoid membrane was determined by measuring the binding of ¹⁴Clabeled diuron to the D1 protein of PS II on a Chl basis (Vermaas, 1990b). In cyanobacteria the distal antenna system is provided by chlorophyll-less phycobilisomes, and therefore, all Chl present in cyanobacterial thylakoids is associated with either PS II proteins or PS I proteins. As a consequence, the concentration of PS II per Chl predominantly depends on the ratio of PS II to PS I. Since this ratio is strongly influenced by the physiological state of the cells, the cultures assayed for herbicide binding were grown in parallel and to the same optical density in the presence of 5 mM glucose. Under the conditions applied in this study, roughly 1 PS II per 800 Chl molecules is detected in wild-type cells. Representative doublereciprocal herbicide-binding data are shown in Figure 3 for wild type, $\Delta 3$, $\Delta 8$, $\Delta 10$, and $\Delta 11$. Diuron-binding affinity was similar in all strains (roughly 20 nM). Calculated Chl/ PS II ratios of wild type and all mutants led to the results given in Table II. A general feature emerging from these data is the finding that the photoautotrophic mutants with rather normal growth rates exhibit a slightly reduced PS II concentration compared to wild-type cells as shown by an up-to-2-fold increase in the number of Chl molecules per PS II. This slight decrease in PS II concentration can be interpreted in terms of a generally enhanced destabilization of PS II in the mutant strains. A significant decrease in PS II concentration is observed in the case of $\Delta 3$ and $\Delta 11$, indicative of a greatly reduced PS II stability. This is in line with the reduced level of PS II proteins detected on immunoblots for $\Delta 3$ and $\Delta 11$ thylakoids. In all obligate photo-

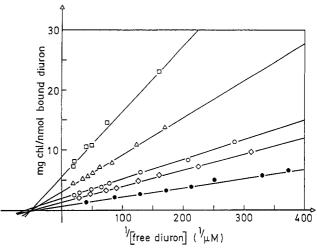
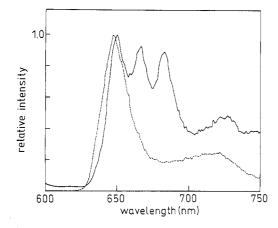


FIGURE 3: Double-reciprocal plot of binding of ¹⁴C-labeled diuron on a chlorophyll basis in wild-type (\bullet), mutant $\Delta 8$ (\diamond), mutant $\Delta 10$ (O), mutant $\Delta 3$ (Δ), and mutant $\Delta 11$ cells (\Box).

heterotrophic mutants the concentration of functional PS II is below a detectable level.

Photosystem II Activity. The PS II activity of wild-type and mutant strains was determined by measuring the oxygen evolution rate. Whole cells were illuminated at a light intensity of 5000 $\mu E/(m^2 \cdot s)$, which proved to be saturating in all cases. Wild-type cells exhibited a rate of $475 \pm 25 \,\mu\text{mol}$ of $O_2/[(\text{mg})]$ of Chl)·h]. The oxygen evolution activities of mutant cells relative to the wild-type rate are listed in Table II. A comparison of the relative PS II activity with the relative PS II concentration reveals that most of the accumulated PS II centers are functional in the photoautotrophic mutants. In general, a reduced PS II concentration correlates with a reduced oxygen evolution rate. However, a close inspection of the data reveals that in most photoautotrophic mutants the PS II activity seems less reduced than the PS II concentration. The origin of this apparent discrepancy, which is especially pronounced in the case of mutant $\Delta 10$, is unknown, but might be related either to (1) the enhanced destabilization of PS II, which could be of different consequence during different characterization procedures, or (2) a limitation of the oxygen evolution rate due to the rate-limiting reoxidation of reduced plastoquinone. The latter would be of greater consequence for the wild type and for those mutants that have relatively more PS II centers per PQ pool (Siggel et al., 1972). Beyond that, it must be emphasized that in some mutants the correlation between PS II abundance and activity exists only for the initial rate of oxygen evolution. In contrast to the wild type, cells of $\Delta 3$, $\Delta 8$, and $\Delta 11$ are not able to maintain a stable oxygen evolution rate during the illumination period of 5 min. The half-times for the decrease in oxygen evolution are 120 s for $\Delta 3$ and $\Delta 8$ and 75 s for $\Delta 11$. A decrease in light intensity only slightly affected the initial oxygen evolution rate over a wide range [1000–5000 μ E/(m²·s)], but increased the stability of the rate. At 500 $\mu E/(m^2 \cdot s)$ all mutants exhibited a stable rate over the entire measuring period of 5 min. A further decrease in light intensity was accompanied by a rapid decrease in the oxygen evolution rate due to subsaturation. At the light intensity provided in the growth chamber $[60 \,\mu\text{E}/(\text{m}^2\cdot\text{s})]$ the rate was roughly 5% of the initial rate under saturating conditions for all samples.

Fluorescence Emission Spectra. CP47 is known to function as a core antenna and is estimated to contain 10-15 (Barbato et al., 1991) or 20–25 Chl a molecules (de Vitry et al., 1984). Low-temperature fluorescence emission spectra provide a sensitive tool to monitor changes in the antenna system.



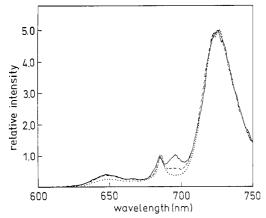


FIGURE 4: Normalized 77 K fluorescence emission spectra. (A) Spectra of mutant Δ8 cells (—) and thylakoids (- - -) excited at 590 nm. (B) Spectra of wild-type (—), mutant $\Delta 3$ (---), and mutant $\Delta 12$ thylakoids (...) excited at 440 nm. The bandwidth of the excitation and emission monochromators was 2 nm.

Substantial evidence indicates that the 77 K fluorescence peaking at 695 nm (F_{695}) originates from a Chl a species bound to CP47, and likewise, a second peak at 685 nm (F_{685}) is attributed for the most part to the Chl a-binding CP43 (Nakatani et al., 1984; van Dorssen et al., 1987). To check whether the modifications in loop E of CP47 generated in this study have any influence on the environment of the antenna Chl molecules, fluorescence emission spectra at 77 K were recorded. Unfortunately, in cyanobacteria a considerable contribution to the fluorescence emitted between 630 and 700 nm is due to the phycobilin pigments, phycocyanin and allophycocyanin, associated with the distal antenna system. In order to minimize this effect, the emission spectra were measured in thylakoids which are highly depleted of the peripheral antenna. Since phycobilisomes are extrinsic watersoluble pigment-protein complexes, they are readily detached and lost during a simple thylakoid preparation even without employing any detergent treatment that might alter fluorescence characteristics (Irrgang et al., 1988). Figure 4A compares the fluorescence emission spectra of whole cells and thylakoids excited at 590 nm, where excitation of phycobilisomes is maximal and that of Chl a is minimal. It is obvious that the fluorescence emitted between 670 and 700 nm is remarkably reduced in thylakoids, indicative of a loss of allophycocyanin. Furthermore, the fluorescence peak at 650 nm, mostly originating from remaining phycocyanin, is slightly blue-shifted, which implies that the remnant has a changed conformation, e.g., due to detachment from PS II. Figure 4B shows the 77 K fluorescence emission spectra of thylakoids isolated from wild-type cells and from mutants $\Delta 3$ and $\Delta 12$ excited at 440 nm within the Soret band of Chl a. The spectra exhibit the three characteristic peaks at 685, 695, and 725 nm

Table III:	Evaluation of 77K Fluorescence Emission Spectra ^a					
	deletion	F ₆₈₅	F ₆₉₅	F ₇₂₅		
WT		1.00	1.04	5.0		
$\Delta 1$	I265-F268	0.90	0.40	5.0		
$\Delta 2$	T271-K277	1.33	0.40	5.0		
$\Delta 3$	K277-E283	0.93	0.61	5.0		
$\Delta 4$	T304-L309	0.95	0.43	5.0		
$\Delta 5$	F311-N317	1.03	0.40	5.0		
$\Delta 6$	G333–I336	0.65	0.61	5.0		
$\Delta 7$	K347-R352	0.61	0.68	5.0		
$\Delta 8$	A373-D380	0.60	0.67	5.0		
$\Delta 9$	V392-Q394	0.88	0.77	5.0		
$\Delta 10$	D416-F420	0.61	0.54	5.0		
$\Delta 11$	R422-E428	1.14	0.55	5.0		
$\Delta 12$	D440-P447	1.04	0.38	5.0		
H343L		0.91	0.84	5.0		

^a Spectra were recorded using thylakoids excited at 440 nm. For convenience, all spectra were normalized to the PS I fluorescence peaking at 725 nm and the relative intensity of F_{725} was set to 5, resulting in a relative intensity F_{695} of roughly 1 for wild-type thylakoids. The normalized amplitudes of F_{685} , F_{695} , and F_{725} were measured from the

and a minor broad band around 650 nm which is due mainly to some remaining phycobilin pigments. For the sake of direct comparison the spectra were normalized to the emission peak at 725 nm that originates predominantly from PS I. The relative amplitudes of the emission peaks at 685, 695, and 725 nm of thylakoids from wild-type cells and from all mutants are summarized in Table III. Two important observations can be made from these data: (i) the normalized extent of the 695-nm peak, referred to as F_{695} , qualitatively correlates with the content of stably assembled PS II; and (ii) all mutants exhibit a marked fluorescence at 685 nm, irrespective of the presence of assembled PS II centers. A closer inspection of the F_{695} data shows that all mutants lacking a stable PS II complex have an F_{695} value of about 0.4. This amplitude we interpret to be composed of the overlapping flanks of the fluorescence bands at 685 and 725 nm. Subtraction of this background level results in corrected values, $(F_{695})_{corr}$, that correlate with the abundance of the PS II polypeptides, CP47, D1, and D2 in wild-type cells and in the mutant strains $\Delta 6$, $\Delta 7$, $\Delta 8$, $\Delta 9$, and H343L. This finding suggests that the properties of the Chl a molecules bound to CP47 and the excitation energy transfer to the reaction center are not markedly affected by the deletions in loop E of these mutants.

Predictions of the Secondary Structure of CP47. The results presented imply that the effects of a deletion mutation are highly dependent on the location of the 3-8-residue deletion. This indicates that the deletions affect elements of quite different functional or structural relevance. Unfortunately, direct information on the structure of CP47 is not available. The current structure model of CP47 is based mainly on hydropathy plot analysis and predicts the topology and orientation of CP47 within the thylakoid membrane (Vermaas et al., 1987; Bricker, 1990). General features of this model include the number and location of membranespanning α -helices, the orientation of the N- and the C-terminus, and the existence of a large loop between the fifth and sixth putative helices. However, the secondary structure of the large extrinsic loop is unresolved. All methods available for predictions of the secondary structures from the amino acid sequence of a protein are only weakly correlated with actual structures. Nevertheless, an attempt to evaluate predictions for the secondary structure of loop E of CP47 is made in this study. However, it has to be emphasized that the intention is not to establish a detailed model for the secondary structure of loop E, but to analyze whether mutationally induced functional deficiencies can be correlated 251

		AAVGLWAAGUVIILLTETKGLIKAÄMDYGALÄEETÄKKADSÄTVERVOPPEVMOTILESTTULIDIAAUOLY
	wild-type	HHHTTBBBBTTTTTTTTTTTTTTTT
Δ1	I265-F268	HHHTT-TTBBBBBTTTT
Δ 2	T271-K277	HHHTTBBBBTTTTTT
Δ 3	K277-E283	HHHTTBBBBTTTTT-TTTT-TTTT-TTTT
Δ 4	T304-L309	НННТТВВВВТТТТТТТТТТТТНННИНИНННН
Δ5	F311-N317	HHHTTBBBBTTTTTTTTTTTTTT
Δ6	G333-I336	HHHTTBBBBTTTTTTTTTTTTT
Δ7	K347-R352	HHHTTBBBBTTTTTTTTTTTTTTTT
Δ8	A373-D380	HHHTTBBBBTTTTTTTTTTTTTTTTTT
Δ9	V392-Q394	HHHTTBBBBTTTTTTTTTTTTTTTT
Δ 10	D416-F420	$\tt HHHTTBBBBTTTTTTTTTTTTT$
Δ 11	R422-E428	HHHTTBBBBTTTTTTTTTTTTTTTT
Δ 12	D440-P447	HHHTTBBBBTTTTTTTTTTTTTTTBBBBBTTTT
		321
		KGGLFRTGAMNSGDGIAQEWIGHPIFKDKEGRELEVRRMPNFFETFPVIMTDADGVVRADIPFRRSESKF
	wild-type	ТТВВВВВННННННННТТНКНННННННВВВВВВВТТТ
Δ1	I265-F268	ТТВВВВВнининининТТнининининыввввввтТТ
Δ 2	T271-K277	ТТВВВВВннинининттининининыввввввттт
Δ 3	K277-E283	ТТВВВВВнининининттининининин
Δ4	T304-L309	ТТВВВВВннининининттнининининыввввввттт
Δ5	F311-N317	ТТВВВВВ
Δ 6	G333-I336	ТТВВВВВ
Δ7	K347-R352	ттвввввннининттининининвввввввттт
Δ8	A373-D380	ТТВВВВВ
Δ9	V392-Q394	TTBBBBBHHHHHHHHHHTTHHHHHHHHBBBBBBBTTTTTTTT
Δ 10	D416-F420	ТТВВВВВ
Δ 11	R422-E428	ТТВВВВВ
Δ 12	D440-P447	ТТВВВВВ
		391
		SVEOTGVTVSFYGGALDGOTFSNPSDVKKFARKAQLGEGFDFDTETFNSDGVFRTSPRGWFTFGHAVFAL
	wild-type	TTTBBBBBTTTTTTTHHHHHHHHTTTTTTT-TTTTT
Δ1	1265-F268	TTTBBBBBTTTTTTTHHHHHHHHTTTTTT-TTTTTHHHH
Δ 2	T271-K277	TTTBBBBBTTTTTTHHHHHHHHTTTTTT-TTTTT
Δ 3	K277-E283	ТТТВВВВВТТТТТННН
Δ 4	T304-L309	TTTBBBBBTTTTTTHHHHHHHHTTTTTTT-TTTTT
Δ 5	F311-N317	TTTBBBBBTTTTTTHHHHHHHH
Δ6	G333-I336	TTTBBBBBTTTTTTHHHHHHHHTTTTTT-TT
Δ7	K347-R352	TTTBBBBBTTTTTTHHHHHHHHTTTTTTTT
Δ8	A373-D380	TTTBBBBBTTTTTTHHHHHHHHTTTTTT-TTTTT
Δ9	V392-Q394	
Δ10	D416-F420	ТТТВВВВВТТТТТННН
Δ 11	R422-E428	TTTBBBBBTTTTTTT-BBBB
Δ12	D440-P447	ТТТВВВВВТТТТТТТНИННИННТТТ-
4	FTT/	

FIGURE 5: Comparison of the secondary structure prediction for wild-type and mutant sequences in the region between residues 251 and 460 of CP47 (Jameson & Wolf, 1988). The deletions are replaced by empty frames. Other symbols: (H) indicates a predicted α -helix; (T) indicates a predicted β -turn; (B) indicates a predicted β -sheet; (-) indicates that no secondary structure element is predicted.

with the disappearance of predicted structural features of loop E. For this purpose, wild-type and mutagenic sequences of loop E were analyzed according to Jameson and Wolf (1988) using the GCG software system, version 7 (Garnier et al., 1978; Devereux et al., 1984). A compilation of the data is

shown in Figure 5. For the photoheterotrophic mutant strains $\Delta 1$, $\Delta 2$, $\Delta 4$, $\Delta 5$, and $\Delta 12$, the deletion has either completely removed or replaced a predicted β -sheet or a longer series of β -turns, while deletions introduced in the photoautotrophic mutant strains $\Delta 6$, $\Delta 7$, and $\Delta 9$ do not interfere very much

with predicted structural elements. However, less correlation between possible structural disturbances and function of the PS II complex is observed for the other mutants, i.e., $\Delta 3$, $\Delta 8$, $\Delta 10$, and $\Delta 11$. Part of the apparent lack of correlation could be due to (1) limitation in secondary structure prediction, (2) possible modifications of the secondary structure by other proteins, and (3) possible importance of selected regions of the loop for PS II structure, function, and assembly.

DISCUSSION

The aim of the present work is to provide a map of functionally important domains or residues of the large hydrophilic loop of CP47 (loop E) which is thought to be exposed to the lumenal side and roughly involves 200 amino acid residues (Vermaas et al., 1987; Bricker, 1990). To address this question, loop E of CP47 was modified using directedmutagenesis techniques. For several reasons loop E is an attractive target to study the consequences of changes in its primary structure on PS II function: (i) within the large evolutionary gap between cyanobacteria and higher plants, the amino acid sequence of loop E is largely conserved, with 60% identical residues and an overall homology of roughly 80%, taking into account conservative changes; (ii) loop E of CP47 might interact with lumenally exposed loops and/or the C-termini of the reaction center proteins containing amino acid residues of relevance for the assembly of a functionally competent water oxidase [for a review, see Debus (1992)]; and (iii) cross-linking experiments have led to the conclusion that the domain between amino acid residues 364 and 440 of loop E interacts with the N-terminus of the extrinsic PS II-O polypeptide (Odom & Bricker, 1992).

The 12 deletion mutants characterized in this study were designed to delete not more than 3-8 amino acid residues from different domains of loop E. As shown previously, this is a length which is not necessarily deleterious for CP47 and PS II assembly and function (Eaton-Rye & Vermaas, 1991). The results of this study demonstrate that the effect of a deletion on PS II is indeed not correlated with the length per se, but is strongly dependent on the position of the deletion on loop E. The mutants obtained cover a wide range of phenotypes which includes those that are photoheterotrophic $(\Delta 1, \Delta 2, \Delta 4, \Delta 5, \Delta 12)$ or photoautotrophic but clearly impaired $(\Delta 3, \Delta 11)$ and those retaining photoautotrophic growth rates similar to wild type ($\Delta 6$, $\Delta 7$, $\Delta 8$, $\Delta 9$, $\Delta 10$). This finding suggests that different regions of loop E are of quite different functional and/or structural importance for PS activity and stability. Two striking phenomena emerge from the results: (i) deletion of amino acids from the regions close to transmembrane helices V and VI prevents normal assembly of the PS II complex; and (ii) a region of about 50 amino acid residues starting from the putative helix V is especially susceptible to changes in the primary structure, whereas the corresponding region adjacent to helix VI is markedly shorter. At present, it is impossible to decide whether the serious damage caused in some of the mutants is due to the deletion of specific amino acid residues of functional relevance (e.g., as a ligand of manganese or another cofactor, like Ca²⁺) or due to a more general structural effect. Interestingly, predictions of the secondary structure indicate that in the case of all deletions leading to photoheterotrophic phenotypes (mutant strains $\Delta 1$, $\Delta 2$, $\Delta 4$, $\Delta 5$, and $\Delta 12$) a predicted β -sheet or a longer series of β -turns is removed or replaced. This finding might favor the latter structural effect to be responsible for the severe impairment of PS II. Nonetheless, a specific functional role of single amino acid residues in these domains cannot be excluded.

A further result is the observation that for most mutant strains the photoautotrophic growth rate correlates with the accumulation of stable PS II centers. Evidence has been provided that mutagenesis of the *psbB* gene does not alter the level of transcription or translation (Yu & Vermaas, 1990). Therefore, a reduced or vanishing level of PS II is very likely due to a reduced stability of PS II caused by an increased turnover of CP47 and as a consequence an increased turnover of D1 and D2. The origin of this instability of CP47, which is induced by a lack of 3–8 amino acids in specific regions of loop E, remains to be clarified. One promising approach could be the introduction of single-site mutations in those regions of loop E which are shown to be of general importance for the assembly of a PS II complex with normal stability.

One of the central points of interest with respect to the function of loop E of CP47 is the mode of interaction with the extrinsic PS II-O protein and/or the redox-active manganese cluster of the water-oxidizing complex. On the basis of crosslinking experiments in PS II membrane fragments from higher plants, the region between E364 and D440 of loop E was inferred to be involved in binding of PS II-O (Odom & Bricker, 1992). Likewise, removal of this polypeptide and the manganese cluster was necessary to render the site of interaction accessible to the monoclonal antibody FAC2. The antigenic determinant for FAC2 was identified as the region between P360 and S391 of loop E (Bricker & Frankel, 1987; Frankel & Bricker, 1990). Provided the mode of interaction between CP47 and the PS II-O protein is very similar in cyanobacteria and higher plants, one might expect that deletions in loop E between P360 and D440 could lead to distortion of this interaction. Therefore, it would be tempting to assume that certain deletions in loop E could lead to changes in the dissociation constant of the extrinsic PS II-O protein to PS II and could result in a phenotype similar to the psbO-deletion mutant strain (Philbrick et al., 1991; Burnap & Sherman, 1991; Mayes et al., 1991). Since the psbO deletion mutant was shown to exhibit enhanced susceptibility to depletion of Ca²⁺ from the growth medium (Philbrick et al., 1991), mutant strains $\Delta 6$, $\Delta 7$, $\Delta 8$, $\Delta 9$, and $\Delta 10$ were screened for enhanced Ca²⁺ demand in the growth medium. However, none of the mutants tested responded to Ca²⁺ depletion by reduction of the photoautotrophic growth rate. Mutant $\Delta 11$ was not subjected to this assay because the autotrophic growth rate was too low to permit experiments leading to conclusive results. Thus, with respect to the Ca²⁺ dependence of photoautotrophic growth, mutant strains $\Delta 6$, $\Delta 7$, $\Delta 8$, $\Delta 9$, and $\Delta 10$ differed from the psbO- deletion mutant. Another characteristic feature of the psbO- mutant is its susceptibility to photoinhibiton at higher light intensities (Philbrick et al., 1991). A similar susceptibility could be observed in the present study for some of the deletion mutants ($\Delta 3$, $\Delta 8$, $\Delta 11$). In this respect, mutant $\Delta 8$, with a deletion that lies within the domain containing the antigenic determinant for FAC2 (Bricker & Frankel, 1987; Frankel & Bricker, 1990), is of special interest. Mutant $\Delta 8$ maintained a stable photoautotrophic growth rate comparable to that of wild-type cells under low-light conditions $(60 \mu E/(m^2 \cdot s)]$, but exhibited an unstable oxygen-evolution rate at higher light intensities, with a half-time of roughly 2 min at 5000 $\mu E/(m^2 s)$. A similar sensitivity to light has been previously observed for a number of single-site mutations in the D2 protein (van der Bolt & Vermaas, 1992). Different lines of evidence indicate that destruction of the wateroxidizing complex causes an increased susceptibility to photoinhibition (Blubaugh et al., 1991; Eckert et al., 1991). Therefore, the enhanced sensitivity to light could reflect an impaired water-oxidizing complex in mutant $\Delta 8$. However,

an unambiguous answer will require more detailed mechanistic studies. In conclusion, measurements of the photoautotrophic growth and oxygen evolution rates point to severe effects on the integrity of the water-oxidizing system by deletions $\Delta 3$ (K277–E283), $\Delta 8$ (A373–D380), and $\Delta 11$ (R422–E428), but do not permit us to specify the origin of this impairment, e.g., due to modified interaction of CP47 with the extrinsic PS II-O protein or changes in ion requirements.

A last point refers to possible effects on the microenvironment of the Chl a molecules bound to CP47 caused by mutations in loop E. The 695-nm peak of the fluorescence emission at 77 K was used as a specific marker for assembled PS II centers. Within reasonable limits, the corrected F_{695} correlates with the concentration of assembled PS II centers for mutants $\Delta 6$, $\Delta 7$, $\Delta 8$, $\Delta 9$, and H343L. This indicates that the decrease of F_{695} is predominantly due to a reduction in assembled PS II centers, and CP47 function as core antenna remains largely unaffected. Therefore, neither the binding environment of the Chl a emitting at 695 nm nor the excitation energy transfer to the reaction center seem markedly affected, at least not by those mutations in loop E that permit synthesis of a stable CP47. However, it is possible that the intensity of F_{695} is, in addition, slightly reduced due to mutationally induced structural changes which influence the rate constants for all decay processes. A similar explanation has been discussed for the decrease of F_{695} in psbO-mutant cells (Burnap & Sherman, 1991). In contrast, fluorescence emission at 685 nm is prominent in all mutant thylakoids. Among the spectra recorded for deletion mutant strains, F_{685} is more pronounced in those with reduced or vanishing PS II levels. This is in line with the detection of CP43 in all mutant thylakoids and with the assumption of an enhanced fluorescence emission from disconnected CP43 due to interruption of energy transfer in mutants largely devoid of assembled PS II.

In summary, the presented results show clearly that the large hydrophilic loop of CP47 is of crucial importance for the functional and structural integrity of the PS II complex. The regions directly adjacent to the predicted transmembrane helices V and VI appear to be of special relevance for structural stability, and this highly sensitive region seems to be especially extended in the N-terminal half of loop E (residues 260-320). In contrast, all short deletions generated in this study in the region between residues 330 and 420 did not cause severe changes of photoautotrophic capacity, although deletion $\Delta 8$ (A373-D380) had a significant detrimental effect on the oxygen-evolving complex. On the other hand, in a previous study deletion of G351-T365 produced a photoheterotrophic phenotype lacking stable assembled PS II centers (Eaton-Rye & Vermaas, 1991). In the latter case, however, the deletion is significantly longer than all deletions described above, and therefore, a length effect cannot be completely excluded. To clarify this question, shorter deletions or singlesite mutations should be introduced into the region between residues 350 and 370. Finally, the C-terminal region of loop E between E364 and D440 has been inferred to be involved in binding of the PS II-O protein (Odom & Bricker, 1992). Concerning the importance of selected charged domains within this region for the interaction of the PS II-O protein with PS II, our results suggest that domains deleted in $\Delta 8$, $\Delta 9$, and $\Delta 10$ are not absolutely indispensable or irreplaceable for an at least partly functional assembly and therefore do not seem to represent unique binding domains for PS II-O. However, it is conceivable that the binding properties are changed.

ACKNOWLEDGMENT

We thank Dr. B. Zilinskas for the gift of the Synechocystis sp. PCC 6803 psbO deletion mutant.

REFERENCES

- Akabori, K., Tsukamoto, H., Tsukihara, J., Nagatsuka, T., Motokawa, O., & Toyoshima, Y. (1988) Biochim. Biophys. Acta 932, 345-357.
- Andreeva, A. V., Buryakova, A. A., Reverdatto, S. V., Chakhmakhcheva, O. G., & Efimov, V. A. (1989) Nucleic Acids Res. 17, 2859-2860.
- Babcock, G. T. (1987) in New Comprehensive Biochemistry (Amesz, J., Ed.) Vol. 15, pp 125-158, Elsevier, Amsterdam.
- Barbato, R., Race, H. L., Friso, G., & Barber, J. (1991) FEBS Lett. 286, 86-90.
- Berry-Lowe, S. L., Johnson, C. H., & Schmidt, G. W. (1992) Plant Physiol. 98, 1541-1543.
- Blubaugh, D. J., Atamien, M., Babcock, G. T., Golbeck, J. H., & Cheniae, G. M. (1991) Biochemistry 30, 7586-7597.
- Bricker, T. M. (1990) Photosynth. Res. 24, 1-13.
- Bricker, T. M. (1992) Biochemistry 31, 4628-4636.
- Bricker, T. M., & Frankel, L. K. (1987) Arch. Biochem. Biophys. 256, 295-301.
- Bricker, T. M., Odom, W. R., & Queirolo, C. B. (1988) FEBS Lett. 231, 111-117.
- Bukharov, A. A., Kolosov, V. L., & Zolotarev, A. S. (1988) Nucleic Acids Res. 16, 8737.
- Burnap, R. L., & Sherman, L. A. (1991) Biochemistry 30, 440-
- Chua, N.-H. (1980) Methods Enzymol. 69, 434-446.
- Debus, R. (1992) Biochim. Biophys. Acta 1102, 269-353.
- Devereux, J., Haberli, P., & Smithies, O. (1984) Nucleic Acids Res. 12, 385-398.
- de Vitry, C., Wollman, F.-A., & Delepelaire, P. (1984) Biochim. Biophys. Acta 767, 415-422.
- de Vitry, C., Olive, J., Drapier, D., Recouvreur, M., & Wollman, F.-A. (1989) J. Cell Biol. 109, 991-1006.
- Eaton-Rye, J. J., & Murata, N. (1989) Biochim. Biophys. Acta 977, 219-226.
- Eaton-Rye, J. J., & Vermaas, W. F. J. (1991) *Plant Mol. Biol.* 17, 1165-1177.
- Eckert, H. J., Geiken, B., Bernarding, J., Napiwotzki, A., Eichler, H. J., & Renger, G. (1991) *Photosynth. Res.* 27, 97-108.
- Enami, I., Satoh, K., & Katoh, S. (1987) FEBS Lett. 226, 161-165.
- Enami, I., Miyaoka, T., Mochizuki, Y., Shen, J.-R., Satoh, K., & Katoh, S. (1989) Biochim. Biophys. Acta 973, 35-40.
- Enami, I., Kaneko, M., Kitamuro, N., Koike, H., Souoike, K., Inoue, Y., & Katoh, S. (1991) Biochim. Biophys. Acta 1060, 224-232.
- Enami, I., Ohta, S., Mitsuhashi, S., Takahashi, S., Ikeuchi, M., & Katoh, S. (1992) Plant Cell Physiol. 33, 291-297.
- Frankel, L. K., & Bricker, T. M. (1990) in Current Research in Photosynthesis (Baltscheffsky, M., Ed.) Vol. 1, pp 639-642, Kluwer Academic Publishers, Dordrecht.
- Garnier, J., Osguthorpe, D. J., & Robson, B. (1978) J. Mol. Biol. 120, 97-120.
- Ghanotakis, D. F., de Paula, J. C., Dementriou, D. M., Bowlby, N. R., Petersen, J., Babcock, G. T., & Yocum, C. F. (1989) Biochim. Biophys. Acta 974, 44-53.
- Greer, K. L., & Golden, S. S. (1991) Plant Mol. Biol. 17, 915-917.
- Hansson, Ö., & Wydrzynski, T. (1990) Photosynth. Res. 23, 131-162.
- Hiratsuka, J., Shimada, H., Whittier, R., Ishibashi, T., Sakamoto, M., Mori, M., Kondo, C., Honji, Y., Sun, C. R., Meng, B. Y., Li, Y. Q., Kanno, A., Nishizawa, Y., Hirai, A., Shinozaki, K., & Sugiura, M. (1989) MGG, Mol. Gen. Genet. 217, 185-194.
- Hird, S. M., Webber, A. N., Wilson, R. J., Dyer, T. A., & Gray, J. C. (1991) Curr. Genet. 19, 199-206.
- Ikeuchi, M. (1992) Bot. Mag. 105, 327-373.
- Irrgang, K.-D., Hanssum, B., Renger, G., & Vater, J. (1988) Ber. Bunsen.-Ges. Phys. Chem. 92, 1050-1056.
- Jameson, B., & Wolf, H. (1988) Comput. Appl. Biosci. 4, 181– 186.

- Keller, M., Weil, J. H., & Nair, C. K. K. (1989) Plant Mol. Biol. 13, 723-725
- Laemmli, U. K. (1970) Nature 277, 680-685.
- Lang, J. D., & Haselkorn, R. (1989) Plant Mol. Biol. 13, 441-456.
- MacKinney, G. (1941) J. Biol. Chem. 140, 315-322.
- Mavankal, G., McCain, D. C., & Bricker, T. M. (1986) FEBS Lett. 202, 235-239.
- Mayes, S. R., Cook, K. M., Self, S. J., Zhang, Z., & Barber, J. (1991) Biochim. Biophys. Acta 1060, 1-12.
- Mayfield, S. P., Bennoun, P., & Rochaix, J.-D. (1987) EMBO *J. 6*, 313–318.
- Michel, H., & Deisenhofer, J. (1988) Biochemistry 27, 1-7.
- Miyao, M., & Murata, N. (1984a) FEBS Lett. 170, 350-354. Miyao, M., & Murata, N. (1984b) Biochim. Biophys. Acta 765, 253–257.
- Morris, J., & Herrmann, R. G. (1984) Nucleic Acids Res. 12, 2837–2850.
- Nakatani, H. Y., Ke, B., Dolan, E., & Arntzen, C. J. (1984) Biochim. Biophys. Acta 765, 347-352.
- Nixon, P. J., & Diner, B. A. (1992) Biochemistry 31, 942-948. Odom, W. R., & Bricker, T. M. (1992) Biochemistry 31, 5616-5620.
- Offermann-Steinhard, K., & Herrmann, R. G. (1990) Nucleic Acids Res. 18, 6452.
- Ohyama, K., Kohchi, T., Sano, T., Sano, S., Umesono, K., Shiki, Y., Takeuchi, M. M., Chang, Z., Aota, S., Inokuchi, H., & Ozeki, H. (1986) Nature 322, 572-574.
- Philbrick, J. B., Diner, B. A., & Zilinskas, B. A. (1991) J. Biol. Chem. 266, 13370-13376.
- Renger, G. (1987) Angew. Chem., Int. Ed. Engl. 26, 660-678. Renger, G., & Wydrzynski, T. (1991) Biol. Met. 4, 73-80.
- Rippka, R., Deruelles, J., Waterbury, J. B., Herdman, M., & Stanier, R. Y. (1979) J. Gen. Microbiol. 111, 1-61.
- Rock, C. D., Barkan, A., & Taylor, W. C. (1987) Curr. Genet. 12, 69-77.
- Rögner, M., Chrisholm, D. A., & Diner, B. (1991) Biochemistry *30*, 5387–5395.

- Rutherford, A. W., Zimmermann, J.-L., & Boussac, A. (1992) in The Photosystems: Structure, Function and Molecular Biology (Barber, J., Ed.) pp 179-229, Elsevier, Amsterdam.
- Shen, J.-R., Ikeuchi, M., & Inoue, Y. (1992) FEBS Lett. 301, 145-149.
- Shinozaki, K., Ohme, M., Tanaka, M., Wakasugi, T., Hayashida, N., Matsubayashi, T., Zaita, N., Chunwongse, J., Obokata, J., Yamaguchi-Shinozaki, K., Ohto, C., Torozawa, K., Meng, B. Y., Sugita, M., Deno, H., Kamogashira, T., Yamada, K., Kusuda, J., Takaiwa, F., Kato, A., Tohdoh, N., Shimada, H., & Sugiura, M. (1986) EMBO J. 5, 2043-2049.
- Siggel, U., Renger, G., Stiehl, H. H., & Rumberg, B. (1972) Biochim. Biophys. Acta 256, 328-335.
- Trebst, A. (1986) Z. Naturforsch. 41C, 240-245.
- van der Bolt, F., & Vermaas, W. F. J. (1992) Biochim. Biophys. Acta 1098, 247-254.
- van Dorssen, R. J., Breton, J., Plijter, J. J., Satoh, K., van Gorkom, H. J., & Amesz, J. (1987) Biochim. Biophys. Acta 893, 267-
- Vermaas, W., Charité, J., & Eggers, B. (1990a) in Current Research in Photosynthesis (Baltscheffsky, M., Ed.) Vol. 1, pp 231-238, Kluwer Academic Publishers, Dordrecht.
- Vermaas, W., Charité, J., & Shen, G. (1990b) Biochemistry 29, 5325-5332.
- Vermaas, W. F. J., & Ikeuchi, M. (1991) in The Photosynthetic Apparatus: Molecular Biology and Operation. Cell Culture and Somatic Cell Genetics of Plants. (Bogorad, L., & Vasil, I. K., Eds.) Vol. 7B, pp 5-111, Academic Press, San Diego.
- Vermaas, W. F. J., Williams, J. G. K., Rutherford, A. W., Mathis, P., & Arntzen, C. J. (1986) Proc. Natl. Acad. Sci. U.S.A. 83, 9474-9477.
- Vermaas, W. F. J., Williams, J. G. K., & Arntzen, C. J. (1987) Plant Mol. Biol. 8, 317-326.
- Vermaas, W. F. J., Ikeuchi, M., & Inoue, Y. (1988) Photosynth. Res. 17, 97-113.
- Williams, J. G. K. (1988) Methods Enzymol. 167, 766-778.
- Yu, J., & Vermaas, W. F. J. (1990) Plant Cell 2, 315-322.