Plant Molecular Biology Reporter
Volume 7(4) 1989

pages 266-275

Genetic Resources

Proposals for the Naming of Chloroplast Genes. II. Update to the Nomenclature of Genes for Thylakoid Membrane Polypeptides

Richard B. Hallick

Department of Biochemistry, Bioscience West 524,
University of Arizona, Tucson, Arizona 85721, USA

Abstract: The nomenclature for genes for components of the photosynthetic membranes has been reviewed and updated. Newly discovered genes have been added to the existing convention for gene nomenclature. Genes designated *petA* through *petI* are described for components of the photosynthetic electron transport systems, *psaA* through *psaK* for photosystem I components, and *psbA* through *psbR* for photosystem II, including the extrinsic polypeptides of the oxygen-evolving complex. References for representative examples of each gene are given.

In 1983, Hallick and Bottomley proposed a series of general principles for naming both chloroplast genes and nuclear genes for chloroplast proteins, with an emphasis on chloroplast and nuclear genes for components of the thylakoid membranes. Specific gene names were recommended for loci that were known at the time. During the ensuing years, this nomenclature proposal has been almost universally adopted among chloroplast molecular biologists. It has been equally useful in naming cyanobacterial genes with higher plant counterparts. These guidelines have also been cited as a model for a recently proposed *Mitochondrial Gene Nomenclature* (Lonsdale and Leaver, 1988).

Abbreviations: ctDNA, chloroplast DNA; nucDNA, nuclear DNA.

Keywords: nomenclature, chloroplast, genes, thylakoid, photosynthesis

Since 1983, a number of new genes for chloroplast proteins have been characterized. A workshop on gene nomenclature was held in Stockholm during the August, 1989 meeting of the *VIIIth International Congress on Photosynthesis*. The purpose was to update the nomenclature of genes for components of the photosynthetic membranes. In this article I shall report on the recommendations from this workshop.

General Principles

In the original proposal for chloroplast gene nomenclature, some general principles were articulated. It is worthwhile to restate and expand on those concepts most relevant to thylakoid protein genes.

- Genes should be named.

- The gene name should contain the maximum information about the gene product to allow ease of identification. As a corollary to this principle, the same gene from different sources should always have the same name.

- The name for a gene product should not be identical to the name for the corresponding gene. DNA should not be given a polypeptide name, and polypeptides should not be given a DNA name. Some investigators, however, have found it useful to use gene names as a guide to identify newly characterized photosystem polypeptides; e.g., the product of the chloroplast *psbI* locus has been referred to as the "PSII - I polypeptide."

- The gene name should consist of two parts. The first part, a three-letter code in lower-case, letters are used to designate the group to which the gene product belongs. The second part, one or more capitalized letters or numbers, are used to designate specific genes. The gene name is italicized.

- Groups of genes that are related in coding function should have the same three-letter code. In this system, *"psa"* is used for genes that are components of photosystem I, *"psb"* for components of photosystem II, *"pet"* for polypeptides involved in photosynthetic electron transport, and *"atp"* for components of ATP synthase. For groups of polypeptides genes, the capitalized letters or numbers used to designate specific genes within a group do not necessarily

carry any connotation about polypeptide MW hierarchy or relative gel electrophoretic mobility.

- Although the principles of gene nomenclature were originally designed for chloroplast DNA-encoded chloroplast protein genes, in some cases there are obvious extrapolations to nuclear genes, e.g., those coding for thylakoid polypeptides. Some nuclear gene names were proposed by Hallick and Bottomley (1983), including *atpC* and *atpD* for ATP synthase γ- and δ-subunit genes, and *petE* and *petF* for plastocyanin and ferredoxin genes, respectively. This concept has been extended to nuclear genes for PSI and PSII polypeptides.

Photosynthetic Electron Transport

When the convention for gene nomenclature was first proposed in 1983, the gene names *petA*, *petB*, *petC*, *petD*, *petE* and *petF* were reserved for chloroplast- and nuclear-encoded components of the photosynthetic electron transport system (excluding PSI and PSII) as shown in Table I. A conserved locus (orf37) has been recently identified as the gene for a subunit V of low-molecular-mass of the cytochrome b_6-f complex (Haley and Bogorad, 1989). The name *petE* was suggested for this locus, but since this is in conflict with the previous assignment of *petE* for plastocyanin, it is recommended that the subunit V gene be designated *petG*. The names *petH* and *petI* are recommended for the genes for ferredoxin-NADP+ reductase and flavodoxin, respectively.

Table I. Genes for Photosynthetic Electron Transport Components

Gene	Locus	Gene Product
petA	ctDNA	cytochrome *f*
petB	ctDNA	cytochrome b_6
petC	nucDNA	Rieske Fe-S polypeptide, subunit III
petD	ctDNA	subunit IV
petE	nucDNA	plastocyanin
petF	nucDNA	ferredoxin
petG	ctDNA	"orf37" gene product, subunit V
petH	nucDNA	ferredoxin-NADPH reductase (FNR)
petI	nucDNA	flavodoxin

Table II. Genes for Photosystem I Polypeptides.

Gene	Locus	Gene Product
psaA	ctDNA	P700 apoprotein, subunit Ia
psaB	ctDNA	P700 apoprotein, subunit Ib
psaC	ctDNA	9-kDa Fe-S polypeptide
psaD	nucDNA	ferredoxin-binding, subunit II
psaE	nucDNA	18-20 kDa subunit IV
psaF	nucDNA	plastocyanin-binding subunit III
psaG	nucDNA	14-16 kDa subunit V
psaH	nucDNA	10-12 kDa subunit VI
psaI	ctDNA	PSI - I polypeptide
psaJ	ctDNA	PSI - J polypeptide
psaK	nucDNA	PSI- K polypeptide ("P37")

Photosystem I

There are at least five subunits of photosystem I encoded by chloroplast DNA (Table II). The *psaA* and *psaB* genes encode the large, P700 chlorophyll apoproteins. *psaC*, which has also been known as *frxA* (Kohchi et al, 1988), is the gene for the PSI iron-sulfur apoprotein. Two newly identified chloroplast genes for PSI polypeptides of low-molecular mass are *psaI* and *psaJ*. The barley *psaI*-gene product and chloroplast-localized *psaI* gene have been described (Møller et al., 1989). The *psaI* of liverwort chloroplast DNA is orf36b (Fukuzawa et al., 1988). *psaI* of rice chloroplast DNA is orf36 located between orf106 and orf 185 (Hiratsuka et al., 1989). *psaJ* is orf42b of liverwort chloroplast DNA and orf44 of rice chloroplast DNA, located between *trnP* and *rpl33*.

Photosystem II

There are several new genes for low-molecular-mass components of photosystem II encoded by chloroplast DNA (Table III). *psbI* and *psbK* have recently been identified as genes for small polypeptides of photosystem II. They are located in the order *psbK-psbI* on the liverwort, tobacco, and rice chloroplast genomes between *trnS* and *trnQ*, on the opposite strand from the tRNA genes. *psbK* has also been called *lhcA*

Table III. Genes for Photosystem II Polypeptides.

Gene	Locus	Gene Product
psbA	ctDNA	D_1-reaction center polypeptide
psbB	ctDNA	"CP47" chlorophyll apoprotein
psbC	ctDNA	"CP43" chlorophyll apoprotein
psbD	ctDNA	D_2-reaction center polypeptide
psbE	ctDNA	cytochrome b_{559} α-subunit
psbF	ctDNA	cytochrome b_{559} β-subunit
psbH	ctDNA	10-kDa phosphoprotein
psbI	ctDNA	4.8-kDa I-polypeptide
psbJ	ctDNA	reserved; see comments
psbK	ctDNA	3.9-kDa K-polypeptide
psbL	ctDNA	PSII - L polypeptide
psbM	ctDNA	PSII - M polypeptide
psbN	ctDNA	PSII - N polypeptide
psbO	nucDNA	33-kDa polypeptide of O.E.C.[1]
psbP	nucDNA	23-kDa polypeptide of O.E.C.
psbQ	nucDNA	16-kDa polypeptide of O.E.C.
psbR	nucDNA	10-kDa polypeptide

[1]O.E.C. = oxygen-evolving complex

(Fukuzawa et al., 1988). *psbI* was designated orf36a (Fukuzawa et al., 1988). *psbL* is orf38 of liverwort, tobacco, *Euglena* and rice chloroplast DNAs, located immediately downstream of *psbF*. Distal to *psbL* is a conserved orf40 (orf42 of *Euglena*) that is co-transcribed with *psbE-psbF-psbL*. An orf40 gene product has not yet been identified as a photosystem II component, but since some investigators have already used *"psbJ"* to describe this gene, the name *"psbJ"* will be reserved for this locus, to be used if and when an orf40 gene product in photosystem II is identified. *psbM* is orf34 of liverwort, tobacco, and rice chloroplast DNAs, located one to two kb upstream of *rpoB*. *psbN* is orf43 of liverwort, tobacco, and rice chloroplast DNAs, located between *psbB* and *psbH*, on the opposite DNA strand. The gene designation *"psbG"* has been left out of the compilation of genes for components for photosystem II. To avoid confusion, *"psbG"* will not be used for any other PSII component. When the function of the *"psbG"* gene product is defined, it might be desirable to rename this locus.

Photosystem II contains an extrinsic water-oxidizing complex, situated on the lumen side of the thylakoid membranes. It is composed of three nuclear DNA-encoded polypeptides of 33-, 23-, and 16-kDa, and

other smaller polypeptides. There has been no consistency in nomenclature for the genes encoding these polypeptides. Genes names that have been variously used began with *"wox"*, *"oee"*, *"oec"*, or no name at all. To avoid further confusion and the proliferation of different names for the same gene, it is proposed that all of these genes be recognized and named as photosystem II components. The recommended names for the genes for the 33-, 23-, and 16-kDa polypeptides are therefore *psbO*, *psbP*, and *psbQ*, respectively. Note that the *"O"* of *psbO* is a useful mnemonic for "oxygen." *psbR* is recommended for the nuclear DNA-encoded gene for the 10.2-kDa PSII polypeptide.

Central Registrar

In order to avoid duplication of nomenclature and maintain consistency with this proposal, those wishing to publish new gene names are encouraged to contact me at the University of Arizona. Electronic mail addressed to "hallick%biotec@rvax.ccit.arizona.edu" or fax messages to 1-602-621-9288 will receive a prompt response. Copies of this proposal are available by electronic mail.

References

General References for Chloroplast Genes

Hallick, R.B. and W. Bottomley. 1983. Proposals for the naming of chloroplast genes. Plant Mol. Biol. Reporter 1: 38-43.

Crouse, E.J., J.M. Schmitt, and H.J. Bohnert. 1985. Chloroplast and cyanobacterial genomes, genes and RNAs: a compilation. Plant Mol. Biol. Reporter 3: 43-89.

Lonsdale, D.M., and C.J. Leaver. 1988. Mitochondrial gene nomenclature. Plant Mol. Biol. Reporter: 6:14-21.

Ohyama, K., Fukuzawa, H., Kohchi, T., Shirai, H., Sano, T., Sano, S., Umesono, K., Shiki, Y., Takeuchi, M., Chang, Z., Aota, S., Inokuchi, H., and Ozeki, H. 1986. Chloroplast gene organization deduced from complete sequence of liverwort *Marchantia polymorpha* chloroplast DNA. Nature 322: 572-574.

Shinozaki, K., M. Ohme, M. Tanaka, T. Wakasugi, N. Hayashida, T. Matsubayashi, N. Zaita, J. Chunwongse, J. Obokata, K. Yamaguchi-Shinozaki, C. Ohto, K. Torazawa, B.Y. Meng, M. Sugita, H. Deno, T. Kamogashira, K. Yamada, J. Kusuda, F. Takaiwa, A. Kato, N. Tohdoh, H. Shimada and M. Sugiura. 1986. The complete nucleotide sequence of the tobacco chloroplast genome: its gene organization and expression. EMBO J. 5: 2043-2049.

Hiratsuka, J., H. Shimada, R. Whittier, T. Ishibashi, M. Sakamoto, M. Mori, C. Kondo, Y. Honji, C. Sun, B. Meng, Y. Li, A. Kanno, Y. Nishizawa, A. Hirai, K. Shinozaki and M. Sugiura. 1989. The complete sequence of the rice (*Oryza sativa*) chloroplast genome: Intermolecular recombination between distinct tRNA genes accounts for a major plastid DNA inversion during the evolution of the cereals. Mol. Gen. Genet. 217: 185-194.

Umesono, K., H. Inokuchi, Y. Shiki, M. Takeuchi, Z. Chang, H. Fukuzawa, T. Kohchi, H. Shirai, K. Ohyama and H. Ozeki. 1988. Structure and organization of *Marchantia*

polymorpha chloroplast genome. II. Gene organization of the large single copy region from *rps'12* to *atpB*. J. Mol. Biol. 203: 299-331.
Fukuzawa, H., T. Kohchi, T. Sano, H. Shirai, K. Umesono, H. Inokuchi, H. Ozeki, and K. Ohyama. 1988. Structure and organization of *Marchantia polymorpha* chloroplast genome. III. Gene organization of the large single copy region from *rbcL* to *trnI*(CAU). J. Mol. Biol. 203: 333-351.
Kohchi, T., H. Shirai, H. Fukuzawa, T. Sano, T. Komano, K. Umesono, H. Inokuchi, H. Ozeki, and K. Ohyama. 1988. Structure and organization of *Marchantia polymorpha* chloroplast genome. IV. Inverted repeat and small single copy regions. J. Mol. Biol. 203: 353-372.

Photosynthetic Electron Transport

petE

Smeekens, S., M. DeGroot, J. Van Binsbergen and P. Weisbeek. 1985. Sequence of the precursor of the chloroplast lumen protein, plastocyanin. Nature 317: 456-458.
Rother, C., T. Jansen, A. Tyagi, J. Tittgen, and R.G. Herrmann. 1986. Plastocyanin is encoded by an uninterrupted nuclear gene in spinach. Curr. Genet. 11: 171-176.
Vorst, O., R. Oosterhoff-Teerstra, P. Vankan, S. Smeekens and P. Weisbeek. 1988. Plastocyanin of *Arabidopsis thaliana*: isolation and characterization of the gene and chloroplast import of the precursor protein. Gene 65: 59-69.
Last, D.I. J.C. and Gray. 1989. Plastocyanin is encoded by a single-copy gene in the pea haploid genome. Plant Mol. Biol. 12: 656-666.

petF

Smeekens, S., J. Van Binsbergen and P. Weisbeek. 1985. The plant ferredoxin precursor: nucleotide sequence of a full length cDNA clone. Nucl. Acids Res. 13: 3179-3194.
Wedel, N., D. Bartling, and R.G. Herrmann. 1987. Analysis of cDNA clones encoding the entire ferredoxin I precursor polypeptide from spinach. Botanica Acta 101: 295-300.

petG

Haley, J., and L. Bogorad. 1989. A 4-kDa maize chloroplast polypeptide associated with the cytochrome b_6-f complex: Subunit 5, encoded by the chloroplast *petE* gene. Proc. Natl. Acad. Sci. USA 86: 1534-1538.

petH

Newman, B.J. and J.C. Gray. 1988. Characterization of a full-length cDNA clone for pea ferredoxin-NADP+ reductase. Plant Mol. Biol. 10: 511-520.
Jansen, T., H. Reiländer, J. Steppuhn, and R.G. Herrmann. 1988. Analysis of cDNA clones encoding the entire precursor polypeptide for ferredoxin:$NADP^+$ oxidoreductase from spinach. Curr. Genet. 13: 517-522.

Photosystem I

Franzen, L.G., G. Frank, H. Zuber and J.D. Rochaix. 1989. Isolation and characterization of cDNA clones encoding five subunits of photosystem I from the green alga *Chlamydo-*

monas reinhardtii. Proceedings of the VIIIth International Congress on Photosynthesis, Kluwer Academic Publishers, in press.

Møller, B.L., H.V. Scheller, J.S. Okkels, B. Koch, B. Andersen, H.L. Nielsen, I. Olsen, B.A. Halkier and P.B. Høj. 1989. Chloroplast encoded photosystem I polypeptides of barley. Proceedings of the VIIIth International Congress on Photosynthesis, Kluwer Academic Publishers, in press.

psaC

Dunn, P.P.J., L.C. Packman, D. Pappin and J.C. Gray. 1988. N-terminal amino acid sequence analysis of the subunits of pea photosystem I. FEBS Lett. 228, 157-161.

Hayashida, N., T. Matsubayashi, K. Shinozaki, M. Sugiura, K. Inoue and T. Hiyama. 1987. The gene for the 9 kd polypeptide, a possible apoprotein for the iron-sulfur centers A and B of the photosystem I complex, in tobacco chloroplast DNA. Curr. Genet. 12: 247-250.

Høj, P.B., I. Svendsen, H.V. Scheller and B.L. Møller. 1987. Identification of a chloroplast-encoded 9-kDa polypeptide as a 2[4Fe-4S] protein carrying centers A and B of photosystem I. J. Biol. Chem. 262: 12676-12684.

Oh-oka, H., Y. Takahashi, K. Wada, H. Matsubara, K. Ohyama and H. Ozeki. 1987. The 8 kDa polypeptide in photosystem I is a probable candidate of an iron-sulfur center protein coded by the chloroplast gene *frxA*. FEBS Lett. 218: 52-54.

psaD

Hoffman, N.E., E. Pichersky, V.S. Malik, K. Ko and A.R. Cashmore. 1988. Isolation and sequence of a tomato cDNA clone encoding subunit II of the photosystem I reaction center. Plant Mol. Biol. 10: 435-445.

Munch, S., U. Ljunberg, J. Steppuhn, A. Schneiderbauer, R. Nechushtai, K. Beyreuther and R.B. Herrmann. 1988. Nucleotide sequences of cDNAs encoding the entire precursor polypeptides for subunits II and III of the photosystem I reaction center from spinach. Curr. Genet. 14: 511-518.

Lagoutte, B. 1988. Cloning and sequencing of spinach cDNA clones encoding the 20 kDa PSI polypeptide. FEBS Lett. 232: 275-280.

psaE

Munch, S., U. Ljunberg, J. Steppuhn, A. Schneiderbauer, R. Nechushtai, K. Beyreuther and R.B. Herrmann. 1988. Nucleotide sequences of cDNAs encoding the entire precursor polypeptides for subunits II and III of the photosystem I reaction center from spinach. Curr. Genet. 14: 511-518.

Franzen, L.G., G. Frank, H. Zuber and J.D. Rochaix. 1989. Isolation and characterization of cDNA clones encoding the 17.9 and 8.1 kDa subunits of photosystem I from *Chlamydomonas reinhardtii*. Plant Mol. Biol. 12: 463-474. (*psaE* shown in Figure 3).

Okkels, J.S., L.B. Jepsen, L.S. Hønberg, J. Lehmbeck, H.V. Scheller, P. Brandt, G. Høyer-Hansen, B. Stummann, K.W. Henningsen, D. von Wettstein, and B.J. Møller. 1988. A cDNA clone encoding a 10.8 kDa photosystem I polypeptide of barley. FEBS Lett. 237: 108-112.

psaF

Steppuhn, J., J. Hermans, R. Nechushtai, U. Ljunberg, F. Thummler, F. Lottspeich and R.G. Herrmann. 1988. Nucleotide sequence of cDNA clones encoding the entire precursor

polypeptides for subunits IV and V of the photosystem I reaction center from spinach. FEBS Lett. 237: 218-224. (*psaF* shown in Figure 2a).

Franzen, L.G., G. Frank, H. Zuber and J.D. Rochaix. 1989. Isolation and characterization of cDNA clones encoding the 17.9 and 8.1 kDa subunits of photosystem I from *Chlamydomonas reinhardtii*. Plant Mol. Biol. 12: 463-474. (*psaF* shown in Figure 2).

psaG

Steppuhn, J., J. Hermans, R. Nechushtai, U. Ljunberg, F. Thummler, F. Lottspeich and R.G. Herrmann. 1988. Nucleotide sequence of cDNA clones encoding the entire precursor polypeptides for subunits IV and V of the photosystem I reaction center from spinach. FEBS Lett. 237: 218-224. (*psaG* shown in Figure 2b).

psaH

Okkels, J.S., H.V. Scheller, L.B. Jepsen and B.L. Møller. 1989. A cDNA clone encoding the precursor for a 10.2 kDa photosystem I polypeptide of barley. FEBS Lett. 250: 575-579.

psaI

Scheller, H.V., J.S. Okkels, Høj, I. Svendsen, P. Roepstorff, and B.L. Møller. 1989. The primary structure of a 4.0-kDa photosystem I polypeptide encoded by the chloroplast *psaI* gene. J. Biol. Chem. in press.

Photosystem II

psbH and *psbI*

Ikeuchi, M., and Y. Inoue. 1988. A new photosystem II reaction center component (4.8 kDa protein) encoded by chloroplast genome. FEBS Lett. 241: 99-104.

psbK

Murata, N., M. Miyao, N. Hayashida, T. Hidaka and M. Sugiura. 1988. Identification of a new gene in the chloroplast genome encoding a low-molecular-mass polypeptide of photosystem II complex. FEBS Lett. 235: 283-288.

Koike, H., K. Mamada, M. Ikeuchi and Y. Inoue. 1989. Low-molecular-mass proteins in cyanobacterial photosystem II: identification of *psbH* and *psbK* gene products by N-terminal sequencing. FEBS Lett. 244: 391-396.

psbL

Webber, A.N., S.M. Hird, L.C. Packman, T.A. Dyer and J.C. Gray. 1989. A photosystem II polypeptide is encoded by an open reading frame co-transcribed with genes for cytochrome b-559 in wheat chloroplast DNA. Plant Mol. Biol. 12: 141-151.

Ikeuchi, M., K. Takio and Y. Inoue. 1989. N-terminal sequencing of photosystem II low-molecular-mass proteins. 5 and 4.1 kDa components of the O_2-evolving core complex from higher plants. FEBS Lett. 242: 263-269.

psbM and *psbN*

Ikeuchi, M., H. Koike and Y. Inoue. 1989. N-terminal sequencing of low-molecular-mass components in cyanobacterial photosystem II core complex: two components corre-

spond to unidentified open reading frames of plant chloroplast DNA. FEBS Lett. in press.

psbO

Kuwabara, T., Reddy, K.J., and Sherman, L.A. 1987. Nucleotide sequence of the gene from the cyanobacterium *Anacystis nidulans R2* encoding the Mn-stabilizing protein involved in photosystem II water oxidation. Proc. Natl. Acad. Sci. USA 84: 8230-8234.

Tyagi, A., Hermans, J., Steppuhn, J., Jansson, C., Vater, F., and Herrmann, R.G. 1987. Nucleotide sequence of cDNA clones encoding the complete '33 kDa' precursor protein associated with the photosynthetic oxygen-evolving complex from spinach. Mol. Gen. Genet. 207: 288-293.

Philbrick, J.B., and Zilinskas, B.A. 1988. Cloning, nucleotide sequence and mutational analysis of the gene encoding the photosystem II maganese-stabilizing polypeptide of *Synechocystis* 6803. Mol. Gen. Genet. 212: 418-425.

Wales, R., Newman, B.J., Pappin, D., and Gray, J.C. 1989. The extrinsic 33 kDa polypeptide of the oxygen-evolving complex of photosystem II is a putative calcium-binding protein and is encoded by a multi-gene family in pea. Plant Mol. Biol. 12:439-451.

psbP and *psbQ*

Jansen, T., C. Rother, J. Steppuhn, H. Reinke, K. Beyreuther, C. Jannson, B. Andersson and R.G. Herrmann. 1987. Nucleotide sequence of cDNA clones encoding the complete '23 kDA' and '16 kDa' precursor proteins associated with the photosynthetic oxygen-evolving complex from spinach. FEBS Lett. 216: 234-240.

Rochaix, J.D. 1987. Expression of the nuclear gene encoding enhancer protein 2 is required for high levels of photosynthetic oxygen evolution in *Chlamydomonas reinhardtii*. PNAS. 84: 749-753.

psbR

Lautner, A., R. Klein, U. Ljungberg, H. Reiländer, D. Bartling, B. Andersson, H. Reinke, K. Beyreuther and R.G. Herrmann, R.G. 1988. Nucleotide sequence of cDNA clones encoding the complete precursor for the "10-kDa" polypeptide of photosystem II from spinach. J. Biol. Chem. 263: 10077-10081.

Eckes, P., S. Rosahl, J. Schell and L. Willmitzer, L. 1986. Isolation and characterization of a light-inducible, organ-specific gene from potato and analysis of its expression after tagging and transfer into tobacco and potato shoots. Mol. Gen. Genet. 205: 14-22.