

# Lawrence Berkeley National Laboratory

## Recent Work

### Title

THE IDENTIFICATION, ISOLATION, AND SEQUENCE OF THE REACTION CENTER PROTEIN GENES OF THE PHOTOSYNTHETIC PURPLE BACTERIUM RHODOPSEUDOMONAS CAPSULATA

### Permalink

<https://escholarship.org/uc/item/6mt4d4q2>

### Author

Hearst, J.E.

### Publication Date

1984-07-01

2



# Lawrence Berkeley Laboratory

UNIVERSITY OF CALIFORNIA

## CHEMICAL BIODYNAMICS DIVISION

Presented as a Plenary Lecture at the Ninth International Congress on Photobiology, Philadelphia, PA, July 3, 1984

RECEIVED  
LAWRENCE  
BERKELEY LABORATORY

JUL 20 1984

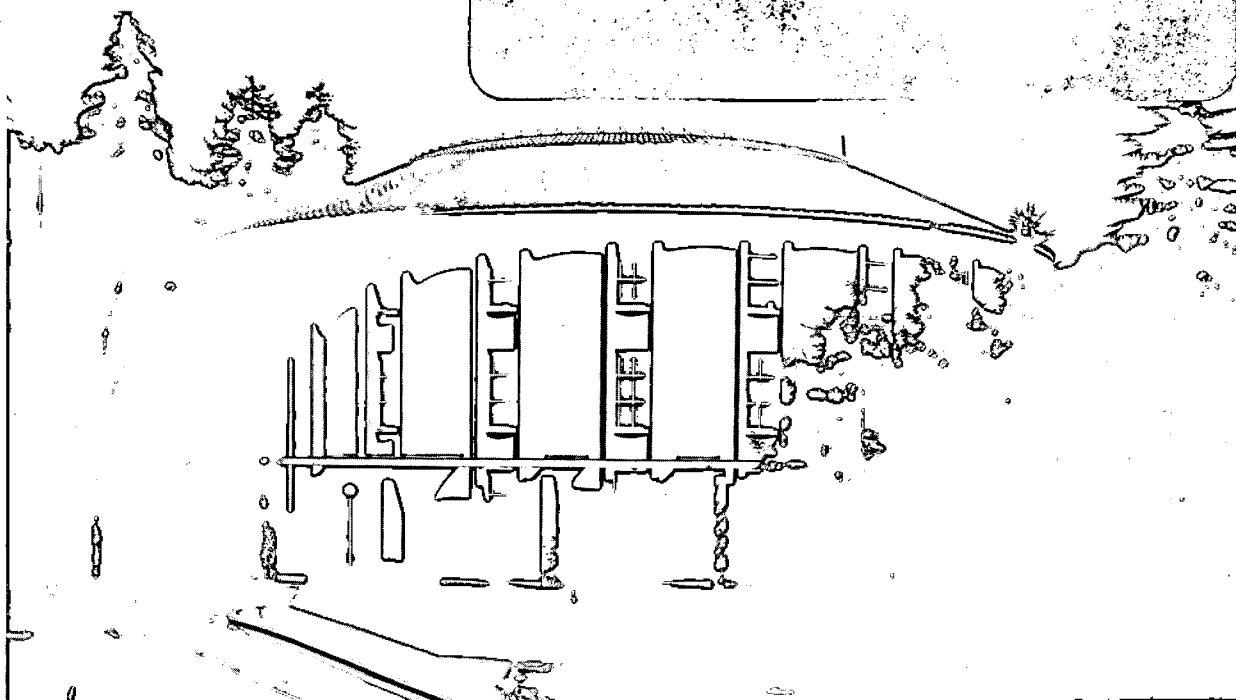
THE IDENTIFICATION, ISOLATION, AND SEQUENCE OF THE REACTION CENTER PROTEIN GENES OF THE PHOTOSYNTHETIC PURPLE BACTERIUM RHODOPSEUDOMONAS CAPSULATA

LIBRARY AND  
DOCUMENTS SECTION

J.E. Hearst

July 1984

**TWO-WEEK LOAN COPY**  
*This is a Library Circulating Copy which may be borrowed for two weeks.*



LBL-18105  
2

## **DISCLAIMER**

This document was prepared as an account of work sponsored by the United States Government. While this document is believed to contain correct information, neither the United States Government nor any agency thereof, nor the Regents of the University of California, nor any of their employees, makes any warranty, express or implied, or assumes any legal responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by its trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof, or the Regents of the University of California. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof or the Regents of the University of California.

THE IDENTIFICATION, ISOLATION, AND SEQUENCE OF THE REACTION  
CENTER PROTEIN GENES OF THE PHOTOSYNTHETIC PURPLE BACTERIUM  
RHODOPSEUDOMONAS CAPSULATA

John E. Hearst

Department of Chemistry  
University of California

and

Division of Chemical Biodynamics  
Lawrence Berkeley Laboratory

Berkeley, California 94720

Presented as a Plenary Lecture at the  
Ninth International Congress on Photobiology  
July 3, 1984 - Philadelphia

## ABSTRACT

Reaction centers in photosynthetic membranes are the centers to which electronic excitation due to light absorption is transferred. This excitation brings about a charge separation between a bacteriochlorophyll molecule and two quinone molecules which ultimately leads to the formation of a hydroquinone. The reduced hydroquinone is then utilized to produce a proton gradient across the membrane and ultimately to produce ATP. We have focused our interest on the structure of the reaction center in the photosynthetic purple bacterium, Rhodospseudomonas capsulata, with the intention of establishing a detailed understanding of these first chemical steps in the natural fixation of sunlight. In this manuscript, I outline the methods that I and my coworkers have used to identify and isolate the genes for the three reaction center subunits, L, M, and H, in Rps. capsulata.

These genes have then been sequenced, and the sequences analyzed in detail for their similarity with sequences of comparable proteins from more advanced photosynthetic bacteria such as Anabena, from algae such as Euglena and Chlamydomonas, and from higher plants such as amaranthus, soybean, tobacco and spinach. Homology was found which has been tentatively interpreted to be in the region of quinone binding in all of these reaction centers. There is growing optimism that there will be substantial structural similarity between the reaction centers of the purple bacteria and those of photosystem II in higher plants. This conclusion is important because the X-ray crystal structures of several of the purple bacteria reaction center complexes are presently being worked on and will ultimately be solved.

## TEXT

We initiated a research program in the Melvin Calvin Laboratory at the University of California on the molecular genetics of photosynthesis about three years ago. We selected Rps. capsulata as the organism of choice entirely because of the prior activity of Professor Barry Marrs whose laboratory had developed gene transfer systems for this organism and had completed the mapping of several bacteriochlorophyll biosynthesis genes and carotenoid biosynthesis genes related to photosynthesis (Yen and Marrs, 1976). Marrs (1981) had isolated a plasmid, a derivative of an R factor, RP1, which grew in E. coli, could be conjugated into Rps. capsulata, and which contained a Rps. capsulata gene cluster capable of complementing all photosynthesis mutants that had been isolated at that time. This plasmid is called pRPS404, contains 116 kilobasepairs of DNA of which the 45 kb section contained between the IS 21 elements or between map positions 23 and 68 in Figure 1 is the Rps. capsulata photosynthesis gene cluster. Figure 1 contains a restriction map worked out by the joint activities of the Marrs laboratory and this laboratory.

Next we developed procedures for the transposon mutagenesis of pRPS404 in E. coli which allow for easy mapping of the positions of the transposon insertions (Youvan et al. 1982). In order to improve the stability of the mutants generated in this way, we created a new transposon called Tn5.7 which has the transposition characteristics of Tn5 and the antibiotic resistance characteristics of Tn7 (Zsebo, Wu and Hearst (1984). Upon conjugation of the mutated pRPS404 into Rps. capsulata, site directed mutation occurs at the same map position in the Rps. capsulata genome as it occurred in the E. coli clone from which it was derived. Recombination is not accompanied by complex DNA rearrangements or deletions (Zsebo and Hearst, 1984). By screening both the

earlier mutants generated by the Marrs laboratory and the transposon mutant generated in this laboratory for fluorescence in the infra-red, we were able to identify two map positions separated by 35 kb which were prime candidates for DNA sequencing. They contained the sequences of the reaction center proteins (Youvan, Hearst and Marrs, 1983).

Youvan et al. (1984b) have just reported the primary sequence of these two 4000+ basepair fragments of Rps. capsulata which code for the light harvesting I polypeptides, B870 $\alpha$  and B870 $\beta$ , the three reaction center polypeptides H, M, and L, and for some presumed regulatory functions relating to the synthesis of these components of the photosynthetic membrane (Zsebo and Hearst, 1984). From the deduced amino acid sequences of the reaction center subunits it was concluded that the H subunit is a hydrophilic protein with a very hydrophobic amino terminus of thirty amino acids. The H subunit is a polypeptide 254 amino acids long with a molecular weight of 28,534 daltons. The L and M subunits are both very hydrophobic proteins with 282 and 307 amino acids respectively. Their molecular weights calculate to be 31,565 and 34,440 daltons, indicating that they have anomolous SDS PAGE electrophoretic mobilities because of their hydrophobic characters. Hydropathy plots suggest that the L and M subunits are transmembrane proteins which may cross the membrane five times.

Feher and Okamura (1978) have shown that the H subunit is not essential to the electron transfer activities of the reaction centers in Rps. sphaeroides. While the H, M, and L subunits always occur in equimolar quantities in reaction centers, the H subunit is thirty-five kilobasepairs away from the operon containing the LH B870 $\alpha$ , LH B870 $\beta$ , L subunit, and M subunit in Rps. capsulata (Youvan et al., 1984a) and so the H subunit gene is clearly under separate transcriptional control from the other structural proteins in the reaction center. Chory et al. (1984) have suggested that the

H subunit serves as the structural anchor in the membrane around which the reaction center forms and evidence is presented suggesting that the H subunit can be found in membrane fractions isolated even from aerobically grown cells.

We are therefore at a point where it is generally accepted that the only two polypeptides essential for reaction center activity are the L and M subunits in Rps. sphaeroides. These two subunits must contain binding sites for the strongly bound and weakly bound ubiquinones,  $Q_A$  and  $Q_B$ , for the Fe(II), and for the bacteriochlorophylls and the bacteriopheophytins. Of these interactions, I consider the sites for binding the quinones most crucial since they serve the roles of preserving or stabilizing very labile anion radicals which are one electron intermediates in the quinone to hydroquinone reduction pathway. This model supposes that pigment binding functions are more variable and therefore less tightly conserved in evolution.

Hearst and Sauer (1984a, 1984b) have discovered that there is a highly conserved pattern of sequences of amino acids which is common to the L and M subunits of Rps. capsulata and the  $Q_B$ -protein of chloroplast thylakoid membranes in spinach and tobacco. The conservation has survived an estimated three billion years of evolution (J. M. Olson, 1981). Such a striking conservation of amino acid sequence suggests that these portions of all three proteins, all of which are approximately 60% in from the amino terminus, are at the functional centers of these proteins. It is their hypothesis that these highly conserved sequences of amino acids are involved in quinone binding and function.

There is a large literature of physical measurements on thylakoid membranes from higher plants which establish the functional and compositional similarities between photosystem II reaction centers in these membranes and



the reaction centers of Rps. capsulata and Rps. sphaeroides. The sequences of the corresponding proteins in plants and purple bacteria now have proven the evolutionary relationships between these reaction center proteins. I would like to suggest here that all PS II reaction centers are likely to contain two similar protein subunits with molecular weights between 30 kd and 40 kd with roles analogous to those of the L and M subunits in the reaction centers of photosynthetic purple bacteria. I am stressing quinone binding here although common features of pigment binding might also prevail. The strongest evidence in favor of this position comes from the recent nucleic acid sequence data of J.-D. Rochaix et al. (1984) and of O.F. Rasmussen et al. (1984). Rochaix et al. (1984) present the nucleic acid sequence of the psbD locus of the Chlamydomonas reinhardtii chloroplast genome. The psbD gene codes for the D2 polypeptide which has been associated with photosystem II (Chua and Gillham, 1977) together with its partner, the D1 polypeptide, which is coded by the psbA gene and which corresponds to the 32 kd herbicide binding protein or Q<sub>B</sub> protein of higher plants. Rochaix et al. make the striking observation that while the overall homology between the Q<sub>B</sub> protein and the D2 protein of C. reinhardtii is only 27%, there are several domains where one can find homology ranging between 33% and 58%. Q<sub>B</sub> contains 352 amino acids while D1 contains 339 or possibly 350 amino acids. The putative quinone binding region is highly conserved in both proteins, with Q<sub>B</sub> corresponding best with the L subunit of Rps. capsulata and D2 corresponding best with the M subunit of Rps. capsulata.

Rasmussen et al. (1984) have sequenced the corresponding psbD gene from the higher plant chloroplast genome, Pisum sativum or pea. Their D2 protein is 353 amino acids long and shows an 84.7% homology to the D2 protein of C. reinhardtii. Again the putative quinone binding region is the most highly

conserved region of the protein relative to the M subunit of Rps. capsulata.

Figure 2 shows an alignment of the amino acid sequences of the L subunit of Rps. capsulata (Youvan et al., 1984b) with the Q<sub>B</sub> proteins of spinach and tobacco (Zurawski et al., 1981), of soybean (Spielmann and Stutz, 1983), of Amaranthus hybridus (Hirschberg and McIntosh, 1983), of C. reinhardii (Erickson et al., 1983), of Euglena (Hallick, 1984), and of Anabena I (Hazelkorn, 1984). Overall, there is 21.5% homology between the L subunit and the Q<sub>B</sub> protein of spinach, 21.0% homology between the L subunit and the Q<sub>B</sub> protein of C. reinhardii and 92.3% homology between the two Q<sub>B</sub> proteins. The most striking feature of this figure is the strong homology between all seven proteins in the hypothetical quinone binding region:

N----PFHMLG----F-----AMHG-LV-S.

Figure 3 shows an alignment of the amino acid sequences of the M subunit of Rps. capsulata (Youvan et al., 1984b), the D2 protein of pea (Rasmussen et al., 1984) and the D2 protein of C. reinhardii (Rochaix et al., 1984). Overall, there is 20.8% homology between the M subunit and the D2 protein of pea, 19.5% homology between the M subunit and the D2 protein of C. reinhardii and an 84.7% homology between the two D2 proteins. The most striking feature of this figure is the strong homology between all three proteins in the hypothetical quinone binding region: F---H-N---NPFH----A---G-ALL-A-HGAT.

In conclusion, I am suggesting that the psbA and psbD loci of chloroplast genomes encode for the photosystem II analogs of the L and M reaction center subunits of R. capsulata and other photosynthetic purple bacteria. While the corresponding chloroplast proteins, D1 and D2, appear to have molecular weights of 32 to 34 kd by SDS PAGE, they are very hydrophobic membrane bound proteins which may cross the membrane five or more times. Their true molecular weights will be between 38 and 40 kd (Zurawski et al.,

1981 and Rasmussen et al., 1984). Hearst and Sauer, 1984a and 1984b have proposed that the most highly co-conserved sequences between the L subunit, the M subunit and the  $Q_B$  protein are involved in quinone binding. While Wolber and Steinback (1984) have proven that the tryptic peptide of the  $Q_B$  protein which contains the presumptive quinone binding site also contains the major site for the covalent linkage between azidoatrazine and the herbicide binding protein in thylakoid membranes, direct physical evidence relating to quinone binding remains elusive and requires further experimentation. In this spirit, it must be remembered that identification of the  $Q_B$  protein with the weak quinone binding site, and the D2 protein with the strong binding site relates to a model and cannot yet be considered experimental fact.

#### ACKNOWLEDGEMENTS

This work was supported by the National Institutes of Health, Grant # GM30786 and by the Office of Energy Research, Office of Basic Energy Sciences, Biological Energy Research Division of the U. S. Department of Energy, Contract DE-AC03-76SF00098.

## BIBLIOGRAPHY

Chory, J., Donohue, T.J., Varga, A.R., Staehelin, L.A., Kalplan, S. (1984).

Reduction of the Photosynthetic Membranes of Rhodopseudomonas sphaeroides: Biochemical and Morphological Studies. In press.

Chua, N.H., Gillham, N.W. (1977). The sites of synthesis of the principal thylakoid membrane polypeptides in Chlamydomonas reinhardtii. J. Cell Biol. 74, 441-452.

Erickson, J.M., Schneider, M., Vallet, J.M., Dron, M., Bennoun, P., Rochaix, J.D. (1983). Chloroplast gene function: combined genetic and molecular approach in Chlamydomonas reinhardtii. Proceedings of 6th International Congress on Photosynthesis, Sybesma, C., editor, M. Nijhoff, W. Junk Publ., in press.

Feher, G., Okamura, M.Y. (1978). Chemical composition and properties of reaction centers. The Photosynthetic Bacteria (Clayton, R.K., and W.R. Sistrom, eds.), pp. 349-386.

Hallick, R. (1984). In press.

Hazelcorn, . (1984). In press.

Hearst, J.E., Sauer, K. (1984a). Protein sequence homologies between portions of the L and M subunits of reaction centers of Rhodopseudomonas capsu-

lata and the 32 kd herbicide-binding polypeptide of chloroplast thylakoid membranes and a proposed relation to quinone-binding sites. In Advances in Photosynthesis Research, Proceedings of the VIth International Congress of Photosynthesis, C. Sybesma, ed., Vol. III.5, 355-359.

Hearst, J.E., Sauer, K. (1984b). Protein sequence homologies between portions of the L and M subunits of reaction centers of Rhodospseudomonas capsulata and the Q<sub>B</sub>-protein of chloroplast thylakoid membranes: A proposed relation to quinone-binding sites. Z. Naturforsch., in press.

Hirschberg, J., McIntosh, L. (1983). Molecular basis of herbicide resistance on Amaranthus hybridus. Science 222, 1346-1349.

Marrs, B. (1981). Mobilization of the genes for photosynthesis from Rhodospseudomonas capsulata by a promiscuous plasmid. J. Bacteriol. 146, 1003-1012.

Olson, J.M. (1981). Evolution of photosynthetic and respiratory prokaryotes and organelles. In Fredrick, J.F., ed. Origins and Evolution of Eukaryotic Intracellular Organelles, pp. 8-19. New York Academy of Sciences.

Rasmussen, O.F., Bookjans, G., Stummann, B.M., Henningsen, K.W. (1984). Localization and nucleotide sequence of the gene for the membrane polypeptide from pea chloroplast DNA. Plant Mol. Biol., submitted.

Rochaix, J.-D., Drom, M., Rahire, M., Malnoe, P. (1984). Sequence homology

between the 32 kd and the D2 chloroplast membrane polypeptide. Plant Mol. Biol., submitted.

Spielmann, A., Stutz, E. (1983). Nucleotide sequence of soybean chloroplast DNA regions which contain the psb A and trn H genes and cover the ends of the large single copy region and one end of the inverted repeats. Nucleic Acids Research 11, 7157-7167.

Wolber, P.K., Steinback, K.E. (1984). Identification of the herbicide binding region of the Q<sub>B</sub>-protein by photoaffinity labeling with azidoatrazine. Z. Naturforsch., in press.

Yen, H.-C. and B. Marrs (1976). Map of genes for carotenoid and bacteriophyll biosynthesis in Rhodospseudomonas capsulata. J. Bacteriol. 126, 619-629.

Youvan, D.C., Elder, J.T., Sandlin, D.E., Zsebo, K., Alder, D.P., Panopoulos, N.J., Marrs, B.L. and Hearst, J.E. (1982). R-Prime site-directed transposon Tn7 mutagenesis of the photosynthetic apparatus in Rhodospseudomonas capsulata. J. Mol. Biol. 162, 17-41.

Youvan, D.C., Hearst, J.E. and Marrs, B.L. (1983). Isolation and characterization of enhanced fluorescence mutants of Rhodospseudomonas capsulata. J. Bacteriology 154, 748.

Youvan, D.C., Alberti, M., Begusch, H., Bylina, E.J., Hearst, J.E. (1984a). Reaction center and light harvesting I genes from Rhodospseudomonas

capsulata. Proc. Nat. Acad. Sci. USA 81, 189-192.

Youvan, D.C., Bylina, E.J., Alberti, M., Begusch, H., Hearst, J.E. (1984b).

Nucleotide and deduced polypeptide sequences of the photosynthetic reaction center, B870 antennae and flanking polypeptides from Rps. capsulata. Cell, in press.

Zsebo, K.M. and Hearst, J.E. (1984). Genetic-physical mapping of a photosynthetic gene cluster from Rhodopseudomonas capsulata. Cell, in press.

Zsebo, K.M., Wu, F. and Hearst, J.E. (1984). Tn5.7 construction and physical mapping of pRPS404 containing photosynthetic genes from Rhodopseudomonas capsulata. Plasmid 11, 182-184.

Zurawski, G., Bohnert, H.J., Whitfield, P.R., Bottomley, W. (1981).

Nucleotide sequence of the gene for the Mr 32,000 thylakoid membrane protein from Spinacia oleracea and Nicotiana debneyi predicts a totally conserved primary translation product of Mr 38,950. Proc. Nat. Acad. Sci. USA 79, 7699-7703.

## FIGURE LEGENDS

Figure 1. Physical map of pRPS404. Arrows indicate IS21 elements which bracket the Rps. capsulata DNA. The Tn5.7 insertions generated during the transposon mutagenesis studies are named for their positions in kilobases on this map.

Figure 2. An amino acid sequence alignment for the L subunit reaction center polypeptide of Rps. capsulata, the herbicide binding thylakoid membrane protein,  $Q_B$ , of spinach and tobacco, of soybean, of Amaranthus hybridus, of C. reinhardii, of Euglena, and of Anabena. This figure was created with the kind assistance of Professor Richard Hallick of the University of Colorado in Boulder.

Figure 3. An amino acid sequence alignment for the M subunit reaction center polypeptide of Rps. capsulata, the D2 protein of pea, and the D2 protein of C. reinhardii. The two D2 proteins are encoded by the psbD loci of the chloroplast genomes of their respective organisms. There are 64 matches and 20.8% homology between the M and D2 pea, 60 matches and 19.5% homology between the M and D2 chlamydomonas, and 287 matches and 84.7% homology between D2 chlamydomonas and D2 pea.



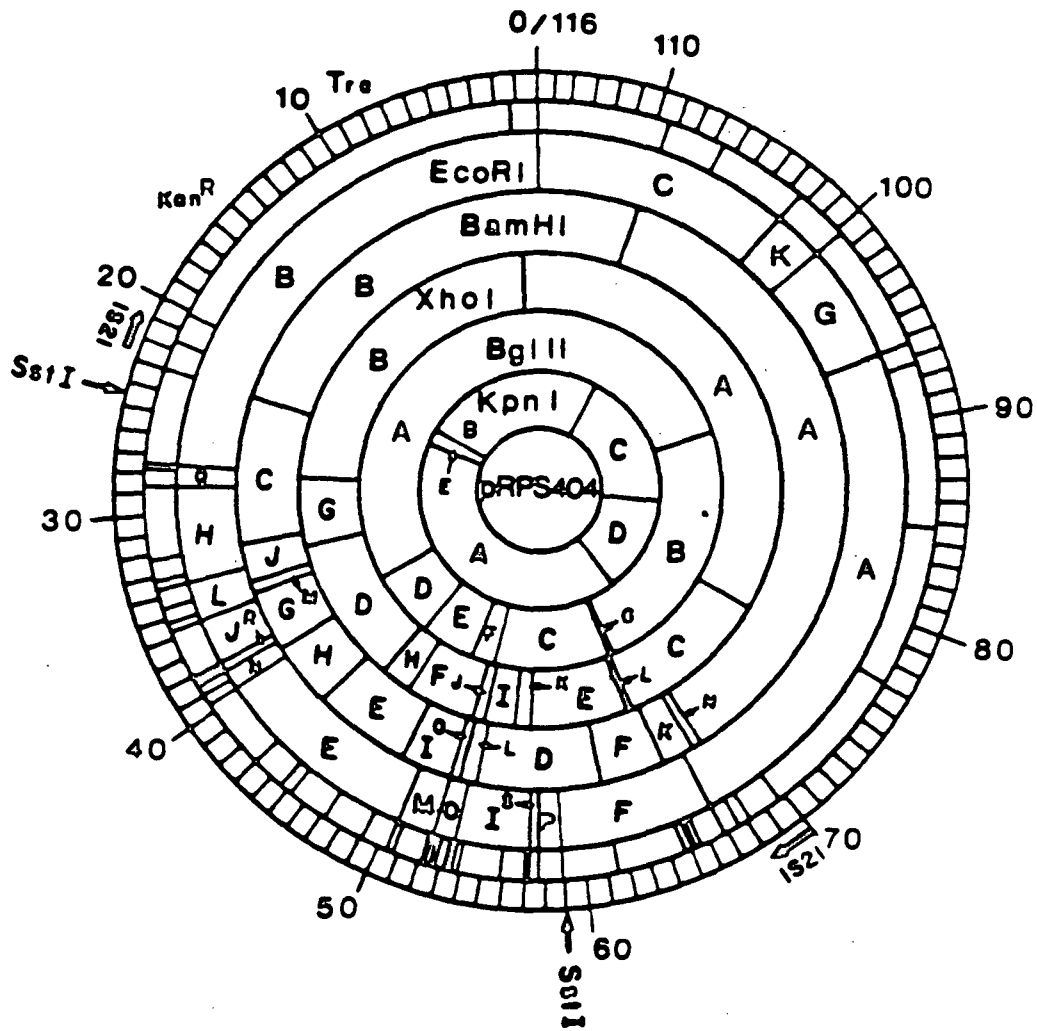


Figure 1

10 20 30 40 50 60 70 80 90  
 MT.AILQRRESESLWGRFCNWI.TSYENRLYIGWFOVLMIPTLLTATSVP.IIAFIAAPPVDIDGIREPVSGSLLYGNNIISGAIIP TSAAIG  
 MT.AILQRRESESLWGRFCNWI.TSTENRLYIGWFOVLMIPTLLTATSVP.IIAFIAAPPVDIDGIREPVSGSLLYGNNIISGAIIP TSAAIG  
 MT.AILQRRESESLWGRFCNWI.TSTENRLYIGWFOVLMIPTLLTATSVP.IIAFIAAPPVDIDGIREPVSGSLLYGNNIISGAIIP TSAAIG  
 MT.AILQRRENSLWARFCEMI.TSTENRLYIGWFOVIMIPCLLTATSVP.IIAFIAAPPVDIDGIREPVSGSLLYGNNIITGAVIPT SNAIG  
 MISPVLKKYARPSLWYHFCAWV.ASKKNRLYVQWFOVLMIPTLLTAATVP.IIAFIAAPPVDIDGIREPVSGSLLYGNNIITGAVIPT SNAIG  
 MT.TTLQQRSSANVWERFCTWI.TSTENRILVQWFOVLMIPTLLTAATVP.IIAFVAAPPVDIDGIREPVAGSLIYGNNIISGAVVPSSNAIG  
 MALLSPERKYRVPQGTLLIGQSLPQWVOPPYVOPFOVTTI...FPATLQFLLLVQAAHQGTQHP.....GLI

SPINACH, NICOTIANA  
 SOYBEAN  
 AMARANTHUS  
 CHLAMYDOMONAS  
 EUGLENA  
 ANABENA I  
 RPS. CAPSULATA

10 20 30 40 50 60  
 100 110 120 130 140 150 160 170 180  
 LHPYPIWQAASVDEWLYNOOPYELIVLHFLLGVACYMGRQWLSFRLOHRPWIAYAYDAPVAAATAVFLIYPICQOQSFSDGMPLOISGTF  
 LHPYPIWQAASVDEWLYNOOPYELIVLHFLLGVACYMGRQWLSFRLOHRPWIAYAYDAPVAAATAVFLIYPICQOQSFSDGMPLOISGTF  
 LHPYPIWQAASVDEWLYNOOPYELIVLHFLLGVACYMGRQWLSFRLOHRPWIAYAYDAPVAAATAVFLIYPICQOQSFSDGMPLOISGTF  
 LHPYPIWQAASLDEWLYNOOPYELIVCHFLLGVACYMGRQWLSFRLOHRPWIAYAYDAPVAAASAVFLIYPICQOQSFSDGMPLOISGTF  
 LHPYPIWQAATSLDEWLYNOQPNQLIVCHFFIGICGYMGRQWLSFRLOHRPWIAYAYDAPVAAASAVFLIYPICQOQSFSDGMPLOISGTF  
 LHPYPIWQAATSLDEWLYNOQPNQLIVHFFIGCACYLGRQWLSYRLOHRPWIAYAYDAPLAATAVFLIYPICQOQSFSDGMPLOISGTF  
 QIFPPPVHOLHVAALDRQGLQVITYCATOAPCQWALREVEICRKLQIGFNI PVAFQHAIPAVLTLVYIRP...HOSWYAFPPYIDTML

SPINACH, NICOTIANA  
 SOYBEAN  
 AMARANTHUS  
 CHLAMYDOMONAS  
 EUGLENA  
 ANABENA I  
 RPS. CAPSULATA

70 80 90 100 110 120 130 140 150  
 190 200 210 220 230 240 250 260 270  
 NFMIVFQAEH.HILMHPFMLOVAGVPGGSLFSAMHOSLVTOSLIRETTEHESANEGYRFGQERETYNIVAACHGYFGRLLIFQYASFNNSRS  
 NFMIVFQAEH.HILMHPFMLOVAGVPGGSLFSAMHOSLVTOSLIRETTEHESANEGYRFGQERETYNIVAACHGYFGRLLIFQYASFNNSRS  
 NFMIVFQAEH.HILMHPFMLOVAGVPGGSLFSAMHOSLVTOSLIRETTEHESANEGYRFGQERETYNIVAACHGYFGRLLIFQYASFNNSRS  
 NFMIVFQAEH.HILMHPFMLOVAGVPGGSLFSAMHOSLVTOSLIRETTEHESANEGYRFGQERETYNIVAACHGYFGRLLIFQYASFNNSRS  
 NFMIVFQAEH.HILMHPFMLOVAGVPGGSLFSAMHOSLVTOSLIRETTEHESANEGYRFGQERETYNIVAACHGYFGRLLIFQYASFNNSRS  
 NFMIVFQAEH.HILMHPFMLOVAGVPGGSLFSAMHOSLVTOSLIRETTEHESANEGYRFGQERETYNIVAACHGYFGRLLIFQYASFNNSRS  
 NFMIVFQAEH.HILMHPFMLOVAGVPGGSLFSAMHOSLVTOSLIRETTEHESANEGYRFGQERETYNIVAACHGYFGRLLIFQYASFNNSRS  
 QVVSHTQYTYGHMHPFMHLOIGLPPFTAVALAHMGAALVLO.....AANPVKKTTRTPQHEOT.....YFRDLQYGVGTLOINR

SPINACH, NICOTIANA  
 SOYBEAN  
 AMARANTHUS  
 CHLAMYDOMONAS  
 EUGLENA  
 ANABENA I  
 RPS. CAPSULATA

160 170 180 190 200 210 220 230  
 280 290 300 310 320 330 340 350  
 LHFFLAANPVVIGIWF TALGISTHAFNLNGFNFNQSVVDSQGRVINTWADIINRANLGMVMMHERNAHNFPDLAAIEA.....PSTHQ  
 LHFFLAANPVVIGIWF TALGISTHAFNLNGFNFNQSVVDSQGRVINTWADIINRANLGMVMMHERNAHNFPDLAAIDA.....PSTHQ  
 LHFFLAANPVVIGIWF TALGISTHAFNLNGFNFNQSVVDSQGRVINTWADIINRANLGMVMMHERNAHNFPDLAAIEA.....PSTHQ  
 LHFFLAANPVVIGIWF TALGLSTHAFNLNGFNFNQSVVDSQDRVLTWADIINRANLGMVMMHERNAHNFPDLASTNS.....SSNN.  
 LHFFLAANPVVIGIWF TALGVSTHAFNLNGFNFNQSVVDSQGRVINTWADIINRANLGMVMMHERNAHNFPDLAA.....  
 LHFFLAANPVVIGIWF TALGVSTHAFNLNGFNFNQSVVDSQGRVINTWADIINRANLGMVMMHERNAHNFPDLAAAGEVAPVALTAPAIHQ  
 LOLLALHAFVQACCHLVBOTIYFDLWQDQWYV.....HHPFWAD.....HA.....GGING

SPINACH, NICOTIANA  
 SOYBEAN  
 AMARANTHUS  
 CHLAMYDOMONAS  
 EUGLENA  
 ANABENA I  
 RPS. CAPSULATA

SPINACH, NICOTIANA	(353)	: RPS. CAPSULATA	76 (21.5%)	: SPINACH	353 (100%)
SOYBEAN	(353)	: RPS. CAPSULATA	75 (21.2%)	: SPINACH	352 (99.7%)
AMARANTHUS	(353)	: RPS. CAPSULATA	75 (21.2%)	: SPINACH	351 (99.4%)
CHLAMYDOMONAS, psba	(352)	: RPS. CAPSULATA	74 (21.0%)	: SPINACH	325 (92.3%)
EUGLENA	(345)	: RPS. CAPSULATA	62 (18.0%)	: SPINACH	305 (86.4%)
ANABENA I	(361)	: RPS. CAPSULATA	70 (19.4%)	: SPINACH	296 (83.9%)
RPS. CAPSULATA	(282)	: RPS. CAPSULATA	282 (100%)	: SPINACH	76 (21.5%)

41

Figure 2

MAEYQNFNQQVQVAGAPEMGLKEDVDTPERTPAGMFNILGWMGNAQIQPIYLGIAGTVSLAFQAAWFFTIQVWYWYQAGFDPPIFMRDLF  
MTIALGKFTKDQNDLFDIMDDWLRRDRFVFGWSGLLLFPCAYFAVGQWFTGTTFVTSWYTHQLASSYLEQCNFLTAAVSTP ANSLAH  
MTIAIGTYQEKRT WFDDADDWLRQDRFVFGWSGLLLFPCAYFALGQWLTGTTFVTSWYTHQLATSYLEQCNFLTAAVSTP ANSMAH

FFSLEPPPAEYQLAIAPLKQQVQIASLFMAISVIAWVVRVYTRADQLGMGKHMWAFLSAIWLWSVLGFWRPILMGSWSVAPPYQIFS  
SLLLLWGPEAQQDLTRWCQLQQLWTFVALHGAFGLIGFMLRQFELARSVQLRPYNAIAFSGPIAVFVSVFLIYPLGQSGWFFAPSFQVAA  
SLLFVWGPEAQQDFTRWCQLQQLWAFVALHGAFGLIGFMLRQFEIARSVNLRPYNAIAFSAPIAVFVSVFLIYPLGQSGWFFAPSFQVAA

HLDTNQPFLDHGMLFYHPPHGLSIAALYQSALLFAMHATILAVTRFGQERELEQIVDRGTASERAALFWRWTMGFNATMEQ IHR  
IFRFILFPQGFH NWTLNPFHMMGVAGVLQAALLCAIHQATVENTLFEDEQDGANTFRAFNPQAE ETYSMVTANRFWSQIFQVAFSNKR  
IFRFILFPQGFH NWTLNPFHMMGVAGVLQAALLCAIHQATVENTLFEDEQDGANTFRAFNPQAE ETYSMVTANRFWSQIFQVAFSNKR

QAIWMAVMVTLT QGIGILLSGTVVDNWYVWAQVHGYAPVTP  
QLHFFMLFVPVTGLWMSALGVVGLALNLRAYDFVSQEIIRAAEDPEFETFYTKNILLNEGIRAWMATODOPHENLIFPEEVLPRGNAL  
QLHFFMLLVPVTGLWMSAIGVVGLALNLRAYDFVSQEIIRAAEDPEFFFSIFIIPNHIINGSYFFNKSQKQIVYI

Amino Acid Sequence Alignments for :

M Subunit	<i>Rhodospseudomonas capsulata</i>	(307 amino acids)
D2 Protein (psbD)	<i>Pisum sativum</i>	(353 amino acids)
D2 Protein (psbD)	<i>Chlamydomonas reinhardtii</i>	(352 amino acids)

This report was done with support from the Department of Energy. Any conclusions or opinions expressed in this report represent solely those of the author(s) and not necessarily those of The Regents of the University of California, the Lawrence Berkeley Laboratory or the Department of Energy.

Reference to a company or product name does not imply approval or recommendation of the product by the University of California or the U.S. Department of Energy to the exclusion of others that may be suitable.

TECHNICAL INFORMATION DEPARTMENT  
LAWRENCE BERKELEY LABORATORY  
UNIVERSITY OF CALIFORNIA  
BERKELEY, CALIFORNIA 94720