Structure of the rubisco operon from the multicellular red alga *Antithamnion spec.*

Markus Kostrzewa, Klaus Valentin, Udo Maid, Rudolph Radetzky*, and Klaus Zetsche

Institut ffir Pflanzenphysiologie, Justus-Liebig-Universitfit, W-6300 Giessen, Federal Republic of Germany

Received July 27/September 4, 1990

Summary. In the multicellular red alga *Antithamnion spec.* both rubisco genes *(rbcL* and *rbcS)* are encoded on the plastid DNA (ptDNA). Both genes are separated by a short A/T-rich spacer of 100 bp and are cotranscribed into an mRNA of approximately 2.7 kb. These findings are in extensive agreement with those obtained from two unicellular red algae *(Porphyridium aerugineum* and *Cyanidium caldarium).* The large subunit (LSU) of rubisco shows an amino acid homology of 82-87% with the LSUs from the two unicellular red algae and only about 55% to LSUs from green algae, higher plants and two cyanobacteria. The small subunit (SSU) of rubisco is more similar to those from the unicellular red algae and two algae which are members of the Chromophyta (about 60% homology) than to cyanobacterial and higher plant proteins (27-36% homology). These data indicate that rhodoplasts originated independently from the chloroplast line. The plastids of chromophytes and rhodophytes appear to be closely related.

Key words: Red algae **-** Plastid evolution **- Ribulose-i ,5 bisphosphate** carboxylase/oxygenase - Plastid DNA

Introduction

The concept of an endosymbiotic origin of plastids is now generally accepted (Margulis 1981; Cavalier-Smith 1987; Giovannoni et al. 1988). As the plastids of red algae (rhodoplasts) share some characteristic features with cyanobacteria [e.g., phycobiliproteins, which are organized in complex structures (phycobilisomes), and thylakoid arrangement] it appears likely that they originate from an endosymbiotic cyanobacteria-like ancestor. Chloroplasts (plastids of green algae and higher plants containing chlorophyll a and b) might have evolved from a chlorophyll *a/b-containing* procaryote, such as *Prochlorothrix hollandica,* which lacks phycobiliproteins (Morden and Golden 1989; cf. Turner et al. 1989).

We have previously analyzed the genes for the Ribulose-l,5-bisphosphate carboxylase/oxygenase (rubisco) from two unicelluar red algae (Valentin and Zetsche 1989, 1990a). In contrast to the situation in chlorophyll *a/b* plants, both rubisco genes are encoded on the ptDNA and are cotranscribed in these algae. The organization of the rubisco genes in unicellular red algae is, therefore, similar to that of the cyanobacteria (Shinozaki and Sugiura 1983; Nierzwicki-Bauer etal. 1984), *Cyanophora paradoxa* (Starnes et al. 1985) and some groups of Chromophyta (Reith and Cattolico 1986; Hwang and Tabita 1989; Douglas and Durnford 1989; Valentin and Zetsche 1990b). Sequence homologies between rubisco genes from red algae, two recent cyanobacteria, and chlorophyll *a/b-plants,* clearly indicate a polyphyletic origin of chloroplasts and rhodoplasts (Valentin and Zetsche 1990a). These findings are supported by a phylogenetic tree based on sequences for the plastid-encoded 5 S rRNA (van den Eynde et al. 1988). In addition, similarities in the structure and the sequence of the small subunit (SSU) suggested a close phylogenetic relationship between rhodoplasts and the plastids of the Chromophyta and the Cryptophyta (Valentin and Zetsche 1990a, b).

To confirm the findings mentioned above it is necessary to analyze the organization and nucleotide sequence of the rubisco genes from a more highly evolved, multicellular, red alga. Moreover, a comparison of rubisco sequence-homologies between unicellular red algae and higher evolved forms with those between green algae and higher plants should allow an estimate as to whether the endosymbiosis which has led to rhodoplasts, or that which has led to chloroplasts, occurred first during the course of evolution.

In this study we present the organization and nucleotide sequence analysis of the rubisco genes from the multicellular filamentous red alga *Antithamnion spec.* (Florideophyceae). The possible phylogenetic implications of the results obtained, such as the phyletic age of

^{} Present address:* Physiologische Chemie der Pflanze, Bergische Universität, FB9, Gaußstrasse 20, W-5600 Wuppertal, Federal Republic of Germany

the red algae and the relationships between rhodoplasts and other plastid types, are discussed.

Materials and methods

Molecular cloning and sequencing of the rubisco operon from Antithamnion spec. PtDNA from *Antithamnion* was isolated as described (Valentin and Zetsche 1989). A 7.5 kb *EeoRI* fragment containing the complete operon was isolated from a ptDNA gene library by colony-hybridization (Walker 1984) with an *rbeL-specific* gene probe from the unicellular red alga *Porphyridium aerugineum* (Valentin and Zetsche 1989). This fragment was mapped with several restriction enzymes and the location of the Rubisco genes was determined by Southern hybridization (Maniatis et al. 1982) using *rbcL-* and *rbcS-specific* gene probes from *Porphyridium.* Rubiscoencoding fragments were subcloned in pUC 18 according to Maniatis et at. (1982). The nucleotide sequence of the complete operon was determined with a "Kilobase Sequencing System" (from BRL) or a "Sequenase Sequencing System'" (from Renner/USB). Overlapping fragments were used in order to determine the sequence of both strands and to cross all restriction sites. Computer analysis was performed with the "Kröger Menue" (Kröger and Kröger-Block 1984).

Nothern analysis. Total cellular RNA from *Antithamnion* was isolated as described (Steinmüller et al. 1983). About 5 µg RNA per lane was electrophoresed on formaldehyd-agarose-gels and blotted to nitrocellulose according to Davis et al. (1986). Blots were hybridized with ³²P-labelled *rbcL*- and *rbcS*-specific gene probes from *Antithamnion spec.* as described (Maniatis et al. 1982).

Construction of dendrograms. Dendrograms of various large subunits (LSUs) and sequence alignements were constructed with the "CLUSTAL" program (Higgins and Sharp 1988). Program parameters, such as "K-TUPLE SIZE" and "GAP PENALTY", were changed in order to differentially evaluate the occurrence of insertions/deletions according to the instruction manual (see legend Fig. 3 and discussion). The program calculates dendrograms on the basis of "similarity scores" which give the number of matching residues between two sequences. From this score a fixed value is subtracted for every gap ϵ = gap penalty). The LSUs from chloroplasts and cyanobacteria are about 13 amino acids shorter than those from rhodoplasts. Thus, assuming an average amino acid homology between different rhodoplast LSUs, or between different chloroplasts LSUs, of more than 80% (see Table 1), the similarity score between chloroplast LSUs will be about 10 points (80% from 13) lower than between rhodoplast LSUs. Therefore a score of 10

1 GAAGTTTTATATTGTATTTTTGTAGACTAATTTTATTAGCATCAGCTACAATTATATAGCTAACTTAACAACTAATGATACTCCCTTAAATTCAAGGAGGAATAAATGTCTAACTCTGTA MetSerAsnSerVal

2159 **TTAATTATAAACTTGTAAATAATATAAACAATAAACAATACAA•ATGAATACAACATTTCCTGATGATACTCTAGTTAATCTACAACAAGAATATGATAAAACTCAAATTCAGAAATAAT** 2279 **AGATGAATTAGAAAAAGAACTTAT~GATTAAAACCTGTTAAAACAGAATTAAAGAGATAGCA~CTTTATTATTAATTGATAAATTAA~AAATAACTTAGGCTTAATATCTGGTAGTCCT**

Fig. 1. Nucleotide and deduced amino acid sequence of the Rubisco operon from *Antithamnion spec.* Putative Shine-Dalgarno sequences and a proposed transcription-terminator (16 bp inverted repeat) are *underlined.* The sequences have been assigned the accession No. X54532 A. SPEC: RHODOPLAST RBCL & RBC S GENES, at the EMBL data bank

was added to every similarity score between chloroplast and cyanobacterial LSUs before calculating dendrograms.

Results

Both Rubisco genes are encoded on the ptDNA of *Antithamnion spec.* as shown by heterologous hybridization using *rbcL-* and *rbcS-specific* gene probes from the unicellular red alga *Porphyridium aerugineum* (Valentin and Zetsche 1989). A 7.5 kb *EcoRI* fragment of the ptDNA hybridizing with both gene probes was cloned and the nucleotide sequence of the region coding for Rubisco was determined (Fig. 1). The *rbcS* gene (417 bp) follows at a distance of t00 bp from the stop codon of the *rbcL* gene (1467 bp). Putative Shine-Delgarno sequences are observed 6 and 9 bp upstream of the *rbcL* and *rbcS* genes, respectively (Fig. 1 underlined). Thirty-five bp downstream of the stop codon of the *rbcS* gene a 16 bp inverted repeat is found which may serve as a transcription terminator (Fig. 1 underlined). When *RbcL-* and *rbcS-specific* gene probes are used to detect Rubisco mRNAs in a Northern blot experiment the gene probes hybridize to

Fig. 2. Cotranscription of both Rubisco genes from *Antithamnion spec. RbeL-(1)* and *rbcS-(2)* specific gene probes were used to identify specific mRNAs. Both gene probes hybridize to a mRNA approximately 2.7 kb in size

Table 1. Sequence homologies (in %, deletions excluded) between Rubisco protein from different sources

SSU	C.cal.	P. aer.	Ant.	Cryp.	$0.$ lut.	E. sil.	Anab.	A.nid.	$C.$ par.	E.gra.	C.rei.	N tab.
P. aer	62.3											
Ant.	60.1	69.6										
Cryp.	61.6	72.5	67.4									
$O.$ <i>lut.</i>	53.6	64.5	57.2	62.6								
E. sil.	52.9	64.7	62.3	64.7	67.4							
Anab.	33.0	34.8	35.7	34.8	35.7	31.3						
A . nid.	36.6	30.4	33.9	33.9	33.9	34.8	58.9					
$C.$ par.	35.5	36.4	35.5	36.4	32.7	32.7	46.2	50.9				
E. gra.	32.1	29.9	27.6	31.3	31.3	29.9	48.2	39.3	43.6			
C.rei.	33.6	31.1	29.9	33.6	30.6	30.6	46.4	42.0	49.1	59.7		
N . tab.	27.6	27.6	29.3	26.8	25.2	26.8	40.2	41.1	39.1	52.8	50.6	
Maize	23.6	26.0	26.8	24.4	25.2	27.7	42.0	42.0	40.0	51.2	51.2	68.3
LSU												
P. aer	83.3											
Ant.	81.7	87.1										
E. sil.	79.3	83.9	82.3									
Anab.	56.9	55.5	54.5			54.3						
A.nid.	53.3	55.5	55.1			54.7	82.5					
$C.$ par.	55.1	53.9	64.1			53.3	82.3	82.5				
E. gra.	53.3	52.5	52.5			52.9	80.3	79.9	83.9			
C.rei.	53.1	52.1	51.5			51.5	81.5	80.9	83.9	90.9		
$N.$ tab.	54.3	54.1	54.7			52.9	82.1	81.3	83.5	84.7	86.5	
Maize	53.9	52.7	53.3			52.5	78.1	76.9	81.7	82.1	83.9	88.7

In the upper section various small subunits (SSUs) are compared. In the lower section sequence homologies between different large subunits (LSUs) of Rubisco are given. Abbreviations are: *C.cal. = Cyanidium caldarium* (Rhodophyta; Valentin and Zetsche 1990a); *P. aer. = Porphyridium aerugineum* (Rhodophyta; Valentin and Zetsche 1989); Anti. =Antithamnion spec (Rhodophyta); Cryp. =Cryptomonas Φ (Cryptophyta; Douglas and Durnford 1989); *O. lut.=Olisthodiscus luteus* (Chromophyta, Chrysophyceae; Boczar et al. 1989); *E.sil.= Ectocarpus siliculosus* (Chromophyta, Phaeophyceae; Valentin and Zetsche 1990 b); Anab. =Anabaena 7•20 (Cyanobacteria; Nierzwicki-Bauer et al. 1984, Curtis and Haselkorn 1983); *A. nid. =Anacystis nidulans* (Cyanobacteria; Shinozaki and Sugiura 1983, Shinozaki et al. 1983); *C.par. = Cyanophora paradoxa* (Starnes et al. 1985; Valentin and Zetsche 1990 c); *E. gra. = Euglena gracilis* (Gingrich and Hallik 1985; Chan et al. 1990); *N. tab. = Nicotiana tabacum* (Mazur and Chui 1985; Shinozaki and Sugiura 1982); *C.rei. = Chlamydomonas reinhardii* (Dron et al. 1982; Goldschmidt-Clermont and Rahire 1986); Maize (Mazur and Chui 1985; Matsuoka et al. 1987)

mRNAs of similar size (approximately 2.7 kb, Fig. 2) indicating that the two genes are cotranscribed.

Both Rubisco genes from *Antithamnion* are colinear with corresponding genes from the unicellular red alga *Porphyridium aerugineum* (data not shown). The *Antithamnion* SSU contains the large insertion at the carboxy terminus which was found to be typical of SSUs from red algae, chromophytes and cryptophytes (Valentin and Zetsche 1990 a, b; data not shown).

Sequence homologies (at the amino acid level) of both subunits of Rubisco from *Antithamnion* with corresponding proteins from chloroplasts and cyanobacteria are given in Table 1.

Discussion

Our results establish that both Rubisco genes are encoded on the ptDNA of *Antithamnion spec.* within a single operon. As we have previously found a similar location and organization of the Rubisco genes in two unicellular red algae *(Porphyridium aerugineum* and *Cyanidium caldarium;* Valentin and Zetsche 1989, 1990 a) this feature appears be typical of the red algae as a whole.

A similar organization of rubisco genes has been established in some unicellular chromophytes (Reith and Cattolico 1986; Hwang and Tabita 1989) and in the cryptophyte *Cryptomonas* Φ (Douglas and Durnford 1989). Therefore, the location of the *rbcS* genes in the nuclear genome of chlorophyll *a/b-plants* appears to be an exception in terms of the whole plant kingdom. In addition, the different coding sites of the *rbcS* gene in red algae, chromophytes and chlorophyll *a/b-plants* may be interpreted as an indication of a common phylogenetic origin of rhodoplasts and phaeoplasts and a polyphyletic origin of these two plastid types and chloroplasts. This hypothesis is strongly confirmed by a sequence comparison of Rubiscos from different sources. Red algal and chromophyte LSUs are well conserved throughout these taxa (79-87% homology, Table 1) but are clearly distinct from LSUs of chloroplasts and two recent cyanobacteria

(about 55% homology, Table 1). LSU homologies between chloroplasts and cyanobacteria are much higher (approximately 83%, Table 1) than between LSUs from chloroplasts and rhodoplasts or phaeoplasts. Thus, chloroplasts are obviously more closely related to cyanobacteria than to rhodoplasts or phaeoplasts. A polyphyletic origin of chloroplasts, rhodoplasts and phaeoplasts, therefore, appears to be very likely (see also Fig. 3). This hypothesis is confirmed by sequence homologies between SSUs from red algae, chromophytes, cyanobacteria, and chlorophyll *a/b-plants* (Table 1, cf. Valentin and Zetsche 1990 a, b).

A comparison of sequence homologies between primitive and advanced forms of both plastid types should help to answer the question as to whether rhodoplasts or chloroplasts arose first in the course of evolution. By sequence comparisons of Rubisco genes we have postulated that the plastids of Chromophyta and Cryptophyta arose from endosymbiotic unicellular red algae (Valentin and Zetsche 1990 a, b). If this hypothesis is correct homologies of red algal LSUs to those from the brown alga *Ectocarpus siliculosus* have to be taken into account when evaluating the phylogenetic age of rhodoplasts. LSU homologies between an advanced red alga *(Antithamnion)* and two primitive forms *(Cyanidium caldarium* and *Porphyridium aerugineum)* are about 83- 87%; that between the LSUs from *Ectocarpus* (advanced chromophyte) and *Cyanidium* (primitive red alga) is 79%. Corresponding homologies between advanced chlorophyll *a/b-pants* (tobacco and maize) and primitive members of this group *(Euglena graeilis* and *Chlamydomonas reinhardtii)* are about 82-86%. The occurrence of insertions/deletions may also be a good criterion for evaluating phylogenetic distances. Considering this, it is noteworthy that the red algal and chromophyte LSUs are colinear except for an insertion of five amino acids (aa) at the amino-terminus of the *Cyanidium* LSU (data not shown). Most LSUs from chlorophyll *a/b-plants* are colinear (e.g., in spinach, *Chlamydomonas* and *Euglena).* The maize LSU shows four insertions/deletions of one aa when compared to the former chloroplasts' LSUs (data not shown). Therefore,

> Fig. 3. Dendrogram of different LSUs constructed using the "CLUSTAL" program. Numbers represent the similarity score which indicates the number of identical residues minus a fixed score ("gap penalty") for every gap introduced to maximize homology. A gap penalty of 1 was used to construct this dendrogram. Gap penalties of up to 3 did not affect the topology of the dendrogram; gap penalties of 4 or more resulted in branching out of the maize LSU before the gree algal LSUs, which appears to be very unlikely. Abbreviations are as in Table 1; *Spin. = Spinacia oleracea* (Zurawski et al. 1981); A/f. = *Alfalfa (Medicago sativa,* Aldrich et al. 1986)

several dendrograms were constructed by different valuations of insertions/deletions (gap penalty between 1 and 3, see Material and methods). Figure 3 represents a dendrogram constructed using different LSUs (gap pen $alty = 1$) which shows rhodoplasts branching out slightly earlier than chloroplasts. By adjusting the gap penalty up to three, we achieved a dendrogram which shows chloroplasts and rhodoplasts branching out simultaneously (data not shown). Analysis of 5 S rRNA sequences from rhodoplasts and chloroplasts revealed a higher phyletic age for the rhodoplasts (van den Eynde et al. 1988). Therefore, it seems likely that rhodoplasts arose slightly earlier than chloroplasts during the course of evolution.

On the other hand, the LSU homologies between higher plants and recent cyanobacteria are in the same range as they are between green algae and higher plants, while the LSU homologies between red algae and recent cyanobacteria are remarkable low. This can be explained in two different ways: either the rhodoplasts originated from a group of cyanobacteria which have not as yet been found or else investigated, or rhodoplasts arose much earlier in the course of evolution than chloroplasts. When discussing the phyletic age of plastids one should consider the necessity for the evolution of an adequate eucaryotic host cell which was capable of integrating a procaryotic cell.

A dendrogram constructed with different LSUs shows rhodoplasts and the plastids of the brown alga *Ectocarpus siliculosus* (Chromophyta, Phaeophyceae) on the same branch (Fig. 3). This result is in agreement with the hypothesis that the plastids of Chromophyta and Cryptophyta originated from a unicellular red alga, like *Porphyridium* species (Gibbs 1981; Valentin and Zetsche 1990 a, b).

It is noteworthy that the *Cyanidium* LSU is found on a branch of the dendrogram (Fig. 3) separated from the red algal and the chromophyte LSUs. *Cyanidium* was originally isolated from hot and acidic springs and tolerates extreme habitats (temperatures up to 58 °C, pH below 1.0; Allen 1959). It may, therefore, represent a very ancient eucaryote or, in contrast, could be an organism showing secondary adaptions. Rubisco homologies seem to support the first hypothesis.

Acknowledgement. This work was supported by the Deutsche Forschungsgemeinschaft (Schwerpunktprogramm Endocytosymbiose, Ze 71/21-1).

References

Aldrich J, Cherney B, Palmer J (1986) Nucleic Acids Res 14:9535 Allen MB (1959) Arch Microbiol 32:270-277 Communicated by H. Kössel

- Cavalier-Smith T (1987) Ann New York Acad Sci 503:17-54
- Chan RL, Keller M, Canaday J, Weil J-H, Imbault P (1990) EMBO J 9:333-338
- Curtis SE, Haselkorn R (1983) Proc Natl Acad Sci USA 80: 1835- 1839
- Davis LC, Dibner MD, Battey JF (1986) Basic methods in molecular biology. Elsevier, Amsterdam
- Douglas SE, Durnford DG (1989) Plant Mol Biol 13:13-20
- Dron M, Rahire M, Rochaix J-D (1982) J Mol Biol 162:775-793
- Gibbs SP (1981) Ann New York Acad Sci 361:193-207
- Gingrich JC, Hallick RB (1985) J Biol Chem 260: 30, 16162-16168 Giovannoni SJ, Turner S, Olsen GJ, Barns S, Lane DJ, Pace NR
- (1988) J Bacteriol 170:3548-3592
- Goldschmidt-Clermont M, Rahire M (1986) J Mol Biol 191: 421- 432
- Higgins DG, Sharp PM (1988) Gene 73:237-244
- Hwang S-H, Tabita FR (1989) Plant Mol Biol 13:69-79
- Kröger M, Kröger-Block A (1984) Nucleic Acids Res 12:113-123
- Maniatis T, Fritsch EF, Sambrook J (1982) Molecular cloning: A laboratory manual. Cold Spring Harbour Laboratory, Cold Spring Harbour, New York
- Margulis L (1981) Symbiosis in cell evolution. W. H. Freeman, San Francisco
- Matsuoka M, Kano-Murakami Y, Tanaka Y, Ozeki Y, Yamamoto N (1987) J Biochem 102: 673-676
- Mazur BJ, Chui CF (1985) Nucleic Acids Res 13:2373-2386
- Morden CW, Golden SS (1989) Nature 337:382-385
- Nierzwicki-Bauer SA, Curtis SE, Haselkorn R (1984) Proc Natl Acad Sci USA 81:5961-5965
- Reith M, Cattolico RA (1986) Proc Natl Acad Sci USA 83: 8599- 8603
- Shinozaki K, Sugiura M (1982) Gene 20:91-102
- Shinozaki K, Sugiura M (1983) Nucleic Acids Res 11:6957-6964
- Shinozaki K, Yamada C, Takahata N, Sugiura M (1983) Proc Natl Acad Sci USA 80:4050-4054
- Starnes SM, Lambert DH, Maxwell ES, Stevens Jr SE, Porter RD, Shively JM (1985) FEMS Microbiol Lett 28:165-169
- Steinmüller K, Kaling M, Zetsche K (1983) Planta 159:698-702
- Turner S, Burger-Wiersma T, Giovannoni SJ, Mur LR, Pace NR (1989) Nature 337:380-382
- Valentin K, Zetsche K (1989) Curt Genet 16:203-209
- Valentin K, Zetsche K (1990a) Mol Gen Genet 222:425-430
- Valentin K, Zetsche K (1990b) Plant Mol Biol (in press)
- Valentin K, Zetsche K (1990c) Curr Genet (in press)
- van den Eynde H, De Baere R, De Rock E, Van de Peer Y, Vandenberghe A, Willckens P, De Wachter R (1988) J Mol Evol 27: 126-132
- Walker JM (1984) Methods in molecular biology, vol. 2: Nucleic acids Humana Press Clifton, New Jersey
- Zurawski G, Perrot B, Bottomley W, Whitfeld PR (1981) Nucleic Acids Res 9:3251-3269

469