Volume 11 Number 20 1983 Nucleic Acids Research

The gene for the small subunit of ribulose-1,5-bisphosphate carboxylase/oxygenase is located close to the gene for the large subunit in the cyanobacterium *Anacystis nidulans* 6301

Kazuo Shinozaki and Masahiro Sugiura

Department of Biology, Faculty of Science, Nagoya University, Chikusa, Nagoya 464, Japan

Received 26 August 1983; Accepted 26 September 1983

ABSTRACT

The gene for the small subunit (SS) of ribulose-1,5-bisphosphate carboxylase/oxygenase from a cyanobacterium, Anacystis nidulans 6301, has been cloned and subjected to sequence analysis. The SS coding region is located close to and downstream from the large subunit (LS) coding region on the same DNA strand. The spacer region between the LS and the SS coding regions contains 93 base pairs (bp), and has no promoter-like sequences. The coding region of A. nidulans SS gene contains 333 bp (111 codons). The deduced amino acid sequence of the A. nidulans SS protein shows 40% homology with those of higher plants.

INTRODUCTION

Cyanobacteria (blue-green algae) are autotrophic prokaryotes which perform plant-type photosynthesis. As in the case of plants, ribulose-1,5-bisphosphate carboxylase/oxygenase (RuBisCO, EC4.1.1.39) is the key enzyme in the Calvin-Benson cycle in most cyanobacteria and is composed of eight identical large subunit (LS) of MW 53,000 and eight identical small subunit (SS) of MW 12,000-14,000 (1). The two subunits of RuBisCO in plants are encoded separately in chloroplast and nuclear DNA; the LS protein in chloroplast DNA (2) and the SS protein in nuclear DNA (3). Since cyanobacteria are prokaryotes and ancient organisms related to cyanobacteria are thought to be the origin of chloroplasts in the endosymbiotic theory , it is of interest to determine whether the LS and the SS proteins of cyanobacteria are encoded in chromosomal DNA as a single operon, or in separate genetic entities.

We have cloned the LS gene of a cyanobacterium (Anacystis nidulans 6301) and determined its nucleotide sequence (4). We

© IRL Press Limited, Oxford, England.

found that the pea SS probe hybridized to specific restriction fragments of total DNA isolated from A. nidulans 6301 and that these restriction fragments were also hybridized with the LS probe. We then determined the nucleotide sequence around the LS coding region and found the SS gene downstream from the LS gene on the same DNA strand. The spacer region between the LS and the SS genes contains 93 base pairs (bp), and has no promoter-like sequences. Thus, the A. nidulans SS gene seems to be co-transcribed with the LS gene.

MATERIALS AND METHODS

Probes and Southern hybridization

The cDNA clone for pea RuBisCO SS (pGR407) was kindly provided by Dr. S. M. Smith. The 0.75 kilobase pairs (kb) EcoRI-BamHI DNA fragment, which contains the pea SS cDNA (5), was used as the SS probe. The 1.25-kb BamHI DNA fragment containing a part of tobacco LS gene was used as the LS probe (6). A. nidulans 6301 DNA was prepared as described (4, 7). Nick translation and Southern blotting were performed as described (4). Hybridization was performed in 30% formamide, 1M NaCl, 10 mM Tris-HCl (pH 7.5), 0.2% each of Ficoll, polyvinylpyrorydone and bovine serum albumin at 37 °C for 24 hrs, washed in 6 x SSC at 49°C for 2 hrs with several changes and then autoradiographed.

DNA sequence analysis

A. nidulans DNA fragments for sequence analysis were prepared from plasmid pANP1155, which contains the 2.3-kb PstI fragment (4). Base-specific chemical cleavages (G, A>C, T+C, C) were perfomed according to Maxam and Gilbert (8). Limited cleavage products were analyzed by electrophoresis in 12% polyacrylamide gels containing 7M urea.

RESULTS AND DISCUSSION

To determine which restriction fragments contained the SS gene, aliquotes of A. nidulans DNA digested with several restriction endonucleases were fractionated by agarose gel

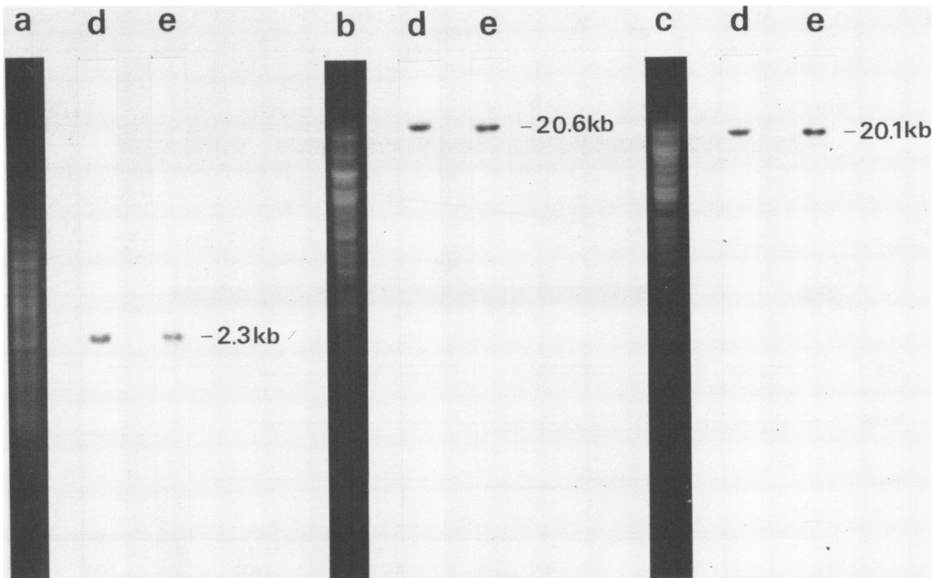

Fig. 1. Location of the DNA fragments containing the _A. nidulans_ SS and LS genes. Eight μg of _A. nidulans_ DNA was digested with PstI, BamHI and SalI, and electrophoresed in 1% agarose gels containing 40mM Tris, 20mM sodium acetate, 2mM EDTA(pH7.8), 0.5 μg/ml ethidium bromide. The DNA fragments were transferred to nitrocellulose filter sheets and hybridized with nick-translated DNA fragments containing the pea SS cDNA or the tobacco LS gene. Lane;a, PstI digest; b, BamHI digest; c, SalI digest; d, autoradiographs of the filter hybridized with the SS probe; e, autoradiographs of the filter hybridized with the LS probe.

electrophoresis, blotted to nitrocellulose filter sheets, and hybridized with the ^{32}P-labeled DNA fragment containing the cloned cDNA of pea SS mRNA (5). A 2.3-kb PstI fragment, a 20.6-kb BamHI fragment and a 20.1-kb SalI fragment were found to hybridize to the SS probe (Fig. 1). The ^{32}P-labeled DNA fragment containing the tobacco LS gene also hybridized to the same DNA fragments (Fig. 1). We had cloned the 2.3-kb PstI fragment in pBR322 and sequenced the gene for the LS protein (4). A physical map of the 2.3-kb PstI fragment is shown in Fig. 2a. DNA fragments containing the SS region (Fig. 2b) were sequenced according to the strategy shown in Fig. 2c. The nucleotide sequence of the SS noncoding (RNA-like) strand and its flanking regions is presented in Fig. 3.

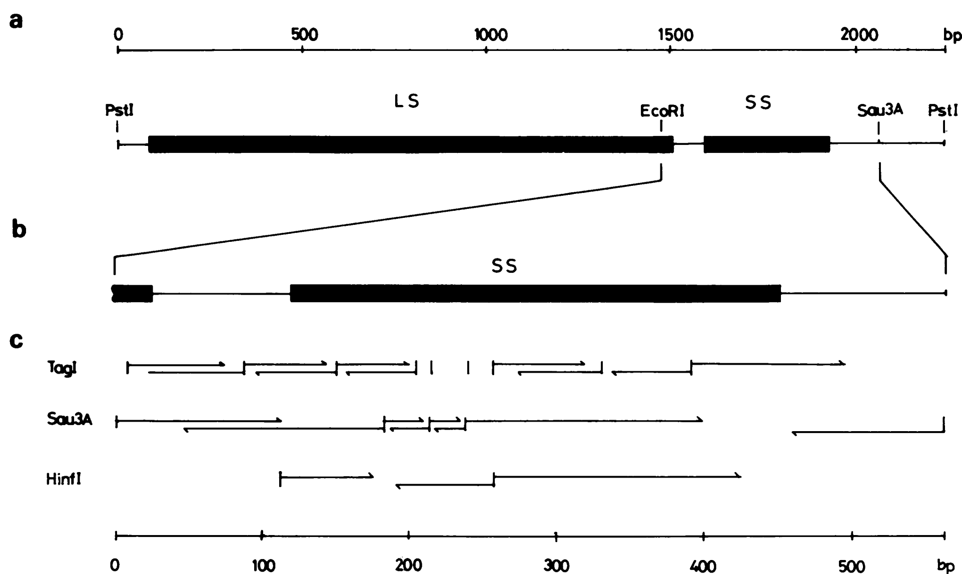

Fig. 2. Physical maps and sequence strategy for the A. nidulans SS gene. A physical map of the cloned 2.3-kb PstI DNA fragment containing both the SS and the LS genes (a), and an expanded physical map of the 564-bp EcoRI-Sau3A sub-fragment (b) are shown. Thickened lines show coding regions. (c) The sequence strategy in which horizontal arrows indicate the direction and extent of DNA regions analyzed.

The coding region of the A. nidulans SS gene contains 333 bp. The amino acid sequence deduced from the nucleotide sequence is also shown in Fig. 3. The A. nidulans SS protein contains 111 amino acid residues and has MW 13,335. For comparative purposes, the amino acid sequences of SS proteins of pea (5), soybean (9), spinach (10), wheat (11) and tobacco (12) are shown in Fig. 4. The deduced amino acid sequence of the A. nidulans SS protein shows 40% homology with that of the pea SS protein, while the sequence of the A. nidulans LS protein shows 80% homology with those of higher plants (4). Thus, the SS proteins are less conservative than the LS proteins. The A. nidulans SS protein is shorter than the pea SS protein by 12 amino acid residues. As shown in Fig. 4, the 12 consecutive amino acid residues (from Glu at position 54 to Arg at position 65 of pea SS) are deleted in the A. nidulans SS. We found two highly conserved regions in SS proteins between A. nidulans and

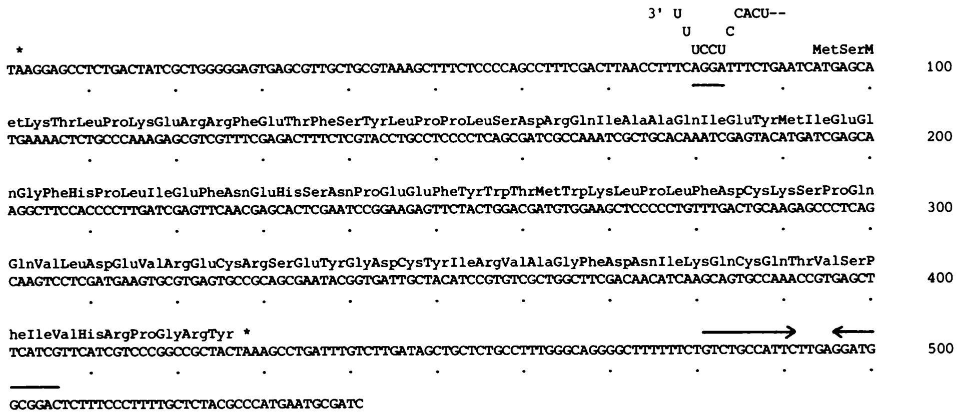

Fig. 3. Nucleotide sequence of the A. nidulans SS coding region and its flanking regions. Numbering starts at the TAA termination codon of the LS gene. The deduced amino acid sequence is indicated above the nucleotide sequence. Horizontal arrows indicate palindrome structure. An underline indicates nucleotide sequence complimentary to the 3'end region of A. nidulans 16S rRNA (14).

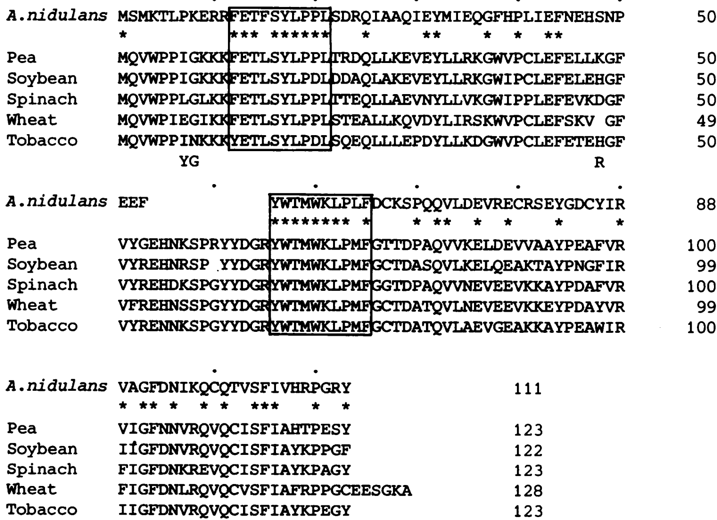

Fig. 4. Comparison of the amino acid sequences of the A. nidulans, pea, soybean, spinach, wheat and tobacco SS proteins. Asterisks indicate homologous amino acid residues of the A. nidulans SS protein with that of pea. Boxed regions indicate conservative sequences among SS proteins compared.

plants (boxed in Fig. 4). One region (from Phe at position 12 to Leu at position 21 of A. nidulans SS) is quite hydrophobic, and the other (from Tyr at position 54 to Phe at position 63 of A. nidulans SS) has α-helix structure (12). These regions may play an important role in binding SS to LS and/or in catalytic function. The A. nidulans SS protein has no transit polypeptide which functions in post-translational transport of the precursors of MW 20,000 of plant SS proteins (3).

The SS gene of A. nidulans is located downstream from the LS gene. The spacer sequence between the LS and the SS genes is 93 bp long. It is of interest whether the LS and the SS genes are transcribed as a single mRNA or not. We could not find typical promoter-like sequences which contain so-called "Pribnow box" and "-35 region" structures (13), while we found an AGGA sequence complementary to the 3' end of A. nidulans 16S rRNA 11 bp upstream from the ATG initiation codon (14, see Fig 3). We have mapped the transcriptional initiation site of the LS gene at position 158-160 bp upstream from the ATG initiation codon (unpublished observation). We found a palindrome structure 52-78 bp downstream from the TAA stop codon of the SS gene , which is likely to be a termination signal of the SS gene (Fig. 3). From these observations, we think that the LS and the SS genes of A. nidulans construct a single operon.

In eukaryotes such as plants and green algae, the LS is encoded in chloroplast genome and synthesized on chloroplast ribosomes (2), while the SS is encoded in nuclear genome and synthesized on cytoplasmic ribosomes as a precursor protein of MW 20,000 which is transported into chloroplast, processed to its mature size and then assembled with the LS protein (3). Recently the SS genes of wheat and soybean are shown to have one or two introns (9, 11). In contrast, the cyanobacterial LS and and the SS genes which contain no introns are located closely each other and are probably transcribed as a polycistronic mRNA. The endosymbiotic theory that chloroplasts in plants and green algae were derived from ancestral photosynthetic prokaryotes related cyanobacteria was proposed by Margulis (15). From this point of view, it can be imagined that the LS

and the SS genes of ancestral cyanobacteria were divided and the SS genes were inserted into nuclear genomes and the LS genes remained in chloroplast genome during evolution. Many other chloroplast proteins consist of multi-subunits which are encoded in nuclear and chloroplast genomes separately. Chloroplast H^+-ATPases have 8 distinct subunits. Five subunits are encoded in chloroplast genomes and three subunits are encoded in nuclear genomes (16, 17). In contrast, all 8 subunits of Esherichia coli H^+-ATPase are located closely each other and probably construct a single operon (18). Chloroplast ribosomal proteins are also encoded in nuclear and chloroplast genomes separately (19). Recently transposition of yeast mitochondrial genes to nuclear genome has been reported (20), which supports a possibility of transposition of chloroplast genes to nuclear genomes during evolution.

ACKNOWLEDGEMENTS

we thank Drs. S. M. Smith for providing us with the pea SS cDNA clone, Dr. Y. Sasaki for suggestions and Mrs. C. Sugita-Yamada for technical assistance. This work was supported in part by a Grant-in-Aid from the Ministry of Education, Science and Culture (Japan).

REFERENCES

1. Jensen, R.H. & Bahr, J.T. (1977) Annu. Rev. Plant Physiol. 28. 329-400
2. Poulsen, C. (1981) Carlsberg Res. Commun. 46, 259-278
3. Ellis, R.J. (1981) Annu. Rev. Plant Physiol. 32, 111-137
4. Shinozaki, K., Yamada, C., Takahata, N. & Sugiura, M. (1983) Proc. Natl. Acad. Sci. USA 80, 4050-4054
5. Bedbrook, J.R., Smith, S.M. & Ellis, R.J. (1980) Nature 287, 692-697
6. Shinozaki, K. & Sugiura, M. (1982) Gene 20, 91-102
7. Tomioka, N., Shinozaki, K., & Sugiura, M. (1981) Mol. Gen. Genet. 184, 359-363
8. Maxam, A.M. & Gilbert, W. (1977) Proc. Natl. Acad. Sci. USA 74, 560-564
9. Berry-Lowe, S.L., Mcknight, T.D., Shah, D.M. & Meagher, R.B (1982) J. Mol. Appl. Genet. 1, 483-498
10. Martin, P.G. (1979) Aust. J. Plant Physiol. 6, 401-408
11. Broglie, R., Coruzgi, G., Lamppa, G., Keith, B. & Chua, N-H. (1983) Bio/Technology 1, in press
12. Miller, K-D., Salnikow, J. & Vater, J. (1983) Biochim. Biophys. Acta. 742, 78-83

13. Pribnow, D. (1975) Proc. Nalt. Acad. Sci. USA 72, 784-788
14. Tomioka, N. & Sugiura, M. (1983) Mol. Gen. Genet. 191, 46-50
15. Margulis, L. (1970) Origin of Eukaryotic Cells. Yale University Press, New Haven.
16. Westhoff, P., Nelson, N., Buneman, H. & Herrmann, R. G. (1981) Curr. Genet. 4, 109-120
17. Deheij, J.T. and Goot, G.S.P. (1981) FEBS Lett. 134, 6-10
18. Kanazawa, H. & Futai, M. (1982) Ann. N. Y. Acad. Sci. 402, 45-64
19. Schmidt, R.J., Richardson, C.B., Gillham, N. & Boynton, J.E. (1983) J. Cell Biol. 96, 145-1463
20. Farrelly, F. & Butow, R.A. (1983) Nature 301, 296-301