Volume 13 Number 7 1985 Nucleic Acids Research

Sequence of a genomic DNA clone for the small subunit of ribulose bis-phosphate carboxylase-oxygenase from tobacco

Barbara J.Mazur and Chok-Fun Chui

E.I.du Pont de Nemours & Co., Central Research & Development Department, Experimental Station, Wilmington, DE 19898, USA

Received 21 December 1984; Revised and Accepted 8 March 1985

ABSTRACT

We have cloned and sequenced a gene for the small subunit (SS) of ribulose bis-phosphate carboxylase-oxygenase from Nicotiana tabacum. The tobacco gene is most closely related to the SS genes from the dicots soybean and pea, and less so to the monocots wheat and Lemna; the deduced amino acid sequence of the mature protein is in all cases more closely conserved than is its chloroplast transit sequence. Unlike the genomic sequences of the two monocots, which have one intron, and the two other dicots, which have two introns, the tobacco gene has three introns. The third tobacco intron lies within a highly conserved region of the protein. Its position coincides with the boundary of a 12 amino acid insertion in the SS genes of higher plants, relative to those of blue green algae. The 5' flanking end of the gene carries 67 bp inverted repeats, which flank a series of eight direct repeats; the direct repeats themselves each carry inverted repeats. The 3' untranslated end of this gene differs by only 2 bp from that of an N. sylvestris SS gene.

INTRODUCTION

Ribulose bisphosphate carboxylase-oxygenase (RUBISCO) is a plant chloroplast enzyme which catalyses the first step in the conversion of atmospheric carbon dioxide to sugar, carboxylating and then hydrolysing ribulose 1,5-bisphosphate to form two molecules of 3-phosphoglycerate.[1] RUBISCO is an abundantly synthesized protein, comprising approximately 50% of the soluble protein in green leaves.[2] It is made up of eight large subunits, which are coded for and localized in the chloroplast, and eight small subunits, which are also chloroplast localized but are nuclear encoded.[3-5] The small subunit (SS) protein is synthesized as a larger precursor carrying an amino-terminal transit sequence which directs the protein to the chloroplast; the transit sequence is post-translationally cleaved from the

precursor polypeptide.[6-8] Expression of the SS gene is light regulated.[9-11]

SS genes from the dicots pea,[12-15] petunia,[16] and soybean,[17] and from the monocots wheat[18] and Lemna[19] have been sequenced. Genomic clones were sequenced from pea, soybean and wheat; cDNA clones were sequenced from petunia and Lemna. A partial cDNA clone has been sequenced from Nicotiana sylvestris, a diploid relative of the tetraploid Nicotiana tabacum with which we work.[20] We present here the sequence of a SS gene from Nicotiana tabacum. The sequence of this gene allows us to compare its features with those of the other sequenced SS genes. It will also allow us to construct gene fusions so that foreign genes of interest may be expressed in tobacco and their products localized in chloroplasts.

MATERIALS AND METHODS

Phage and Bacterial Strains

Phage lambda EMBL4[21] was grown on E. coli strain LE392.[22] M13 phage mp8 and mp9 were grown on E. coli JM101.[23] M13 phage mWB2344 was grown on E. coli WB373.[24] E. coli DH1 was the recipient strain for plasmid transformations[25] using pBR322 vectors.[26] Plasmid pSS15 carries a cDNA insert of a pea SS of RUBISCO.[27] Plasmid pSRS2.1 carries a genomic clone of a soybean RUBISCO SS gene.[17]

Tobacco DNA Isolation

Tobacco DNA was isolated by a modification of the method of Walbot and coworkers.[28] Fifty gram portions of callus tissue were frozen in liquid nitrogen, and then lyophilized. The dried tissue was ground at room temperature in a Tekmar blender, in 15 second bursts, until powdered. Ten volumes of buffer (0.3M sucrose-0.05M Tris pH8.0-0.005M $MgCl_2$) were added, and the suspension incubated at 0°C for five minutes. The suspension was then filtered through cheesecloth, and spun at 350Xg for 10 minutes. The nuclear pellet was resuspended in lysis buffer (0.02M EDTA-0.05M Tris pH8.0-0.1% Sarkosyl), CsCl was added to 0.95 g/ml buffer,and the mixture was spun at 17000 rpm for 20 minutes at 4°C. Ethidium bromide was added to 400 µg/ml to the supernatant, the refractive index was adjusted to 1.39, and the

solution was spun at 35000 rpm in a Ti70 rotor at 20°C for 3 days. The flourescent DNA band was removed from the gradient, the ethidium bromide was extracted with isopropanol, and the DNA was then dialysed against Tris-EDTA(0.01M pH8-0.001M) and ethanol precipitated.

Construction of a tobacco DNA library

The tobacco library was constructed in the phage lambda vector EMBL4.[21] EMBL4 phage were prepared from agarose plate stocks.[29] Phage DNA was prepared by concentrating the phage with polyethylene glycol, removing the polyethylene glycol by chloroform extraction, and purifying the phage on glycerol step gradients. The phage were then treated with DNase and RNase prior to phenol extraction. Phage DNA was spooled out of ethanol. To prepare arms of the EMBL4 phage, phage DNA was sequentially digested with Sal I and Bam HI. The arms were annealed and then separated from the central fragment on a 10-40% sucrose gradient, according to the method of Maniatis et al.[30] The arms were completely denatured and reannealed before being ligated to tobacco DNA. Tobacco DNA, prepared as described above, was partially digested with the restriction enzyme Sau 3A and sedimented through a 10-40% sucrose gradient.[30] Fractions from the sucrose gradient were analysed by running aliquots on 0.5% agarose gels. Fractions containing fragments in the 20-40kb range were dialysed, precipitated, and ligated to the lambda DNA arms. The DNA was ligated at a concentration of 135 µg/ml of vector and 45 µg/ml insert DNA. The ligated concatamers were then packaged using lambda DNA packaging extracts prepared according to Protocol II of Maniatis et al.[30] The yield of phage was approximately 4.5×10^5 phage per microgram of insert DNA. A library of approximately 400,000 phage was constructed, which represents a 99% complete library for tobacco. Tobacco has a genomic content of 1.52 picograms, or 1.5×10^9 base pairs.[31]

Identification of clones containing the SS gene

Phage were plated on 90mm Petri dishes using E. coli strain LE392 as the host. Approximately two thousand phage were plated per plate, on 100 plates. Nitrocellulose blots of the plaques were made according to the procedure of Benton and Davis,[32] and the filters were hybridized against a ^{32}P-labeled

pea SS cDNA probe. Positive plaques were picked after overnight autoradiography and replated and rehybridized until a pure phage had been isolated. DNA was prepared from these phage, and restriction digests of the DNA were separated on agarose gels and blotted onto nitrocellulose filters. The filters were probed either with the pea cDNA clone, or with a soybean genomic clone carrying the 5' end of the SS gene.

Filter hybridizations

Filters were prehybridized in 6X SSPE-0.2% SDS-100 µg/ml calf thymus DNA-5X Denhardt's solution, as described in Maniatis et al.[30]; 1X SSPE is 0.15 M NaCl - 0.01 M NaH_2PO_4 - 0.001 M EDTA, pH 7.0. Filters were hybridized in the same solutions, with approximately 20 million cpm of nick-translated probe, for a minimum of 20 hours. Hybridizations were carried out at 65°C for both heterologous and homologous probes. Filters were rinsed in 6X SSPE-0.2% SDS at 65°C, and then in 2X SSPE at room temperature.

Subcloning a tobacco SS gene

A Hind III-Eco RI fragment of a SS gene, NtSS23, was subcloned into M13mp8 and M13mp9 for DNA sequencing. Additionally, a Hind III fragment spanning the gene and its 5' and 3' flanking regions was subcloned into M13 mp9 for sequencing. A Pst I-Eco RI fragment at the 5' end of the gene was subcloned into M13mWB2344 for sequencing.

Sequencing a tobacco SS gene

The NtSS23 SS gene described above was sequenced using the dideoxy chain termination procedure.[33] Oligonucleotides, usually 14 bases long, were used to prime synthesis. Over part of the DNA both strands were sequenced; all sections of the DNA were sequenced multiple times. All sequences overlapped each other. Sequences were analysed using programs developed by IntelliGenetics, Inc.

RESULTS AND DISCUSSION

Isolation of two genomic clones carrying SS genes

Two tobacco genes for the SS of ribulose bisphosphate carboxylase-oxygenase (SS RUBISCO) were identified through their homology with the RUBISCO genes from pea and soybean. Approximately 200,000 phage of a lambda-EMBL4 library which we con-

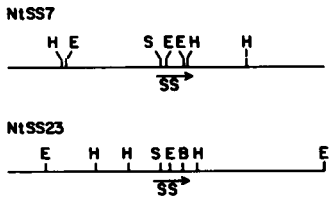


Figure 1. Restriction maps of two tobacco genomic clones carrying the RUBISCO SS gene.

structed were plated; two SS clones were isolated after probing the library with a pea cDNA clone, pSS15.[27] Restriction maps of these two clones, designated NtSS7 and NtSS23, are shown in Figure 1. The two restriction maps are different, indicating that the two SS genes come from different segments of the genome. From the hybridization patterns of these clones with the pea cDNA probe and the soybean 5' genomic probe, the 5' and 3' orientations of the tobacco genes were established. These hybridization patterns also indicated that each fragment carries only one SS gene. Since the lengths of the cloned tobacco fragments are approximately 15 kb, these two genes are not closely linked to other members of the SS gene family. Hybridization of a subclone of NtSS23 to digests of tobacco DNA indicated that there are additional SS genes in tobacco, as has been found in other plants[13-19]. Approximately six Bam HI and five Eco RI fragments hybridized to the probe (not shown), but it is difficult to estimate the number of genes that these fragments represent. The

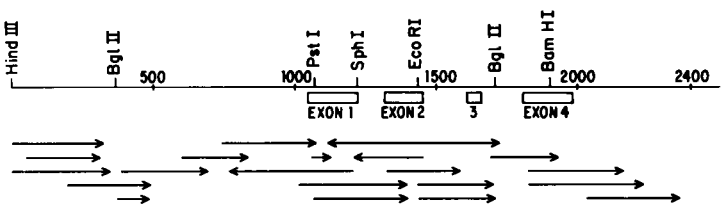


Figure 2. Strategy used to sequence the tobacco RUBISCO SS clone NtSS23. Restriction fragments were subcloned into M13 phage as described in Materials and Methods. Oligonucleotides were used to prime synthesis; the direction and extent of each reaction is denoted by the arrows.

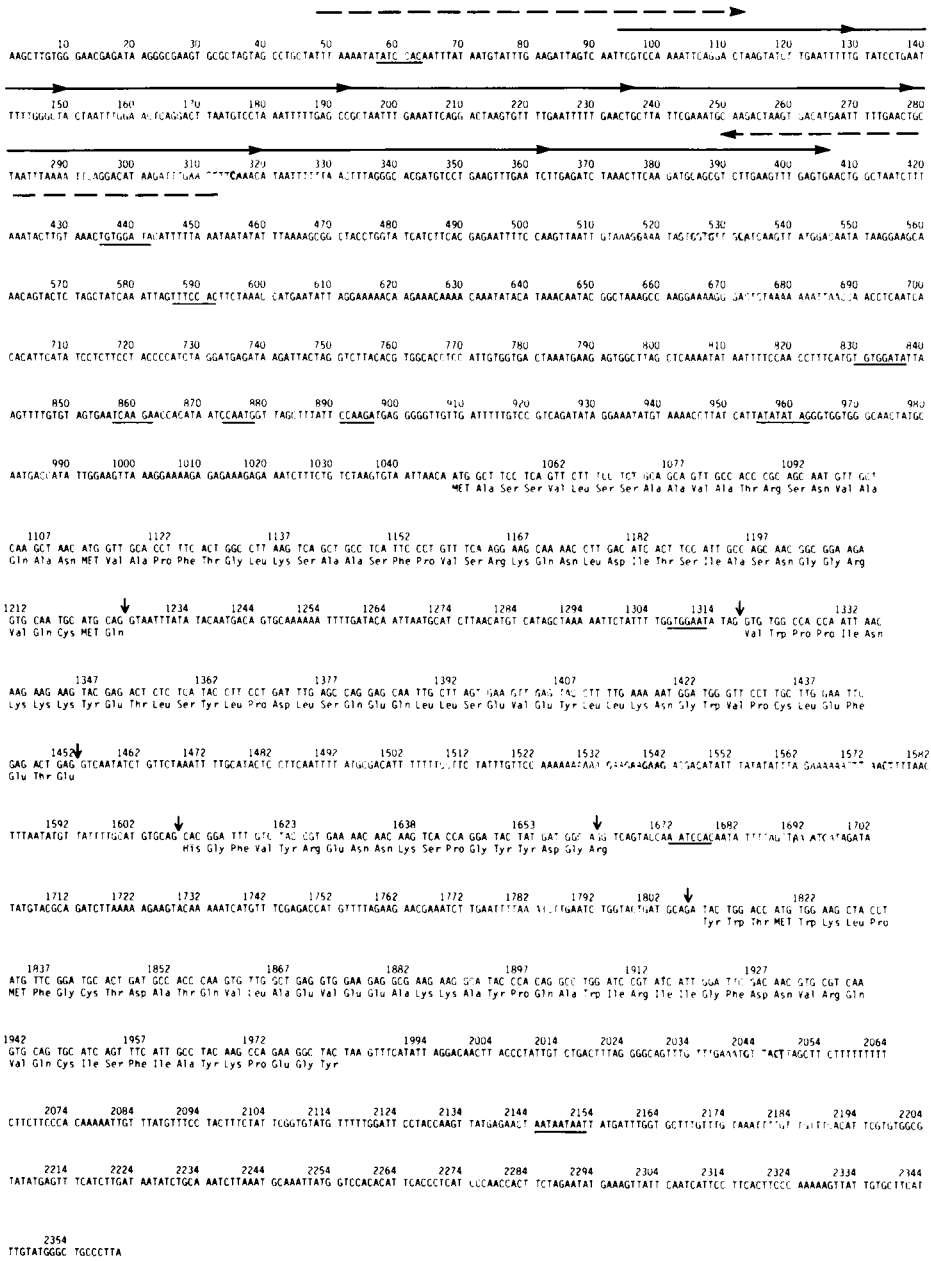


Figure 3. Sequence of the NtSS23 genomic DNA clone of a RUBISCO SS gene from Nicotiana tabacum. The deduced amino acid sequence is shown below the DNA sequence; introns were assigned using the protein sequence data of Muller et al.[34] The third intron is

assumed to have the boundaries indicated by vertical arrows, so that the arg codon at the 5' end of the intron (1660) would be AGA, rather than AGG. Underlining indicates possible transcription regulatory sequences, as described in the text. Vertical arrows show assumed intron splice sites. The upper set of two dashed horizontal arrows indicates the position of the inverted repeats shown in Figure 5, while the lower set of eight solid horizontal arrows marks the position of a series of 43 bp imperfect tandem direct repeats.

number is obscured by the imperfect homologies between small subunit genes and by the presence of introns in the genes which could contribute to the total number of restriction fragments.

Sequence of a tobacco SS gene

Hybridizing fragments of one of the genomic clones, NtSS23, were subcloned into the vectors pBR322, M13mp8, M13mp9, and M13mWB2344. The M13 clones were sequenced using the dideoxy sequencing method with synthetic oligonucleotide primers, such that the DNA sequence was extended on a cloned fragment in overlapping segments. The sequencing strategy is shown in figure 2. The sequence of the gene and of the 5' and 3' flanking regions is shown in figure 3.

The amino acid sequence of the SS protein and its transit sequence, as deduced from the nucleotide sequence, is also shown in figure 3. Intron positions were assigned by comparing the sequence to the amino acid sequence of a Nicotiana tabacum SS protein,[34] and by comparing intron boundary sequences with consensus splice site sequences.[35] The tobacco SS gene codes for a protein predicted to be 180 amino acids long, including a transit sequence of 57 amino acids. The predicted molecular weight of the precursor polypeptide is 20,400, while that of the mature protein is 14,588. Codon usage for this gene is biased; sixteen amino acid codons are never used, eleven codons are used but once, and most of the remaining codons are preferentially used.

The predicted amino acid sequence of this SS gene varies at five positions from the amino acid sequence of the Nicotiana tabacum SS protein sequenced by Muller et al.[34] We find ser rather than leu at residue 28, val rather than pro at 30, glu rather than asp at 31, glu rather than gly at 88, and gln rather than glu at 96. Studies in petunia, Lemna, and pea have shown that there are multiple SS genes and that they are

differentially expressed, for example by tissue type.[15-16,19] Thus, the differences between our DNA-based sequence and the protein sequence derived by Muller et al. may result from our having sequenced a different gene in the SS gene family, or they may result from protein sequencing errors.

A partial cDNA sequence of a Nicotiana sylvestris SS gene has been obtained by Pinck et al.[20]; the two Nicotiana DNA sequences are identical at the amino acid level as far as they can be compared. Thus the N. sylvestris cDNA sequence at amino acids 88 and 90 is identical to our N. tabacum DNA sequence, and not to the N. tabacum protein sequence determined by Muller et al. The N. sylvestris sequence extends from the poly A tail through part of the second exon, to amino acid 38. At the nucleotide level, there are but five differences in the available sequences of the two genes. Three of these changes are silent, third position changes in the last exon. There are two changes in the 3' non-coding region: there is a one-base insertion in the N. sylvestris sequence, and there is a T-C exchange just before the poly A tail begins. Given the extensive differences in the 3' non-coding sequences from two wheat genomic clones that were sequenced,[17] it was somewhat surprising that the N. sylvestris and N. tabacum genes would be so identical in both the coding and non-coding regions. N. sylvestris is one of the diploid ancestors of the tetraploid N. tabacum; presumably we have sequenced one of the N. tabacum SS genes which originated from N. sylvestris. This ancestry is also indicated by the ile-asn at residues 7-8, and by the his at residue 48. N. sylvestris has these same amino acids, whereas Nicotiana tomentosiformis, the other parent of N. tabacum, has substituted tyr-gly and arg, respectively at these positions.[34]

The virtually exact homology between our N. tabacum genomic clone and the N. sylvestris cDNA clone in the 3'non-coding regions suggests that we have cloned a functional SS gene. Divergence at the 3' non-coding ends of SS genes has been used to distinguish among individual members of SS gene families.[15,17-18] To confirm that we have sequenced an active SS gene, we probed tobacco poly(A) mRNA with two types of SS probes. We hybridized the mRNA with a clone of the SS gene, and with an oligonucleotide

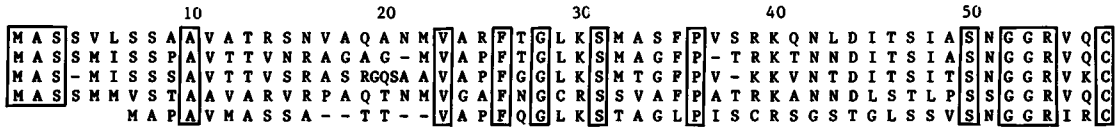
complementary to the 3' non-coding region of the NtSS23 gene. Both probes hybridized to the same species of leaf mRNA (not shown). We also found that a transcript is initiated from the 5' end of this gene when it is used in a chimeric gene construction and returned to tobacco; the chimeric nature of the construction allows us to distinguish transcripts initiated from this gene as opposed to those from other SS genes (Mazur and Chui, unpublished).

Among different plant species, the mature RUBISCO SS protein is more closely conserved than is the transit sequence. SS sequences have been obtained for the the monocots Lemna[19] and wheat,[18] and for the dicots soybean,[17] pea,[12-15] spinach[36] and petunia.[16] The mature tobacco SS sequence is 73% conserved relative to either monocot, and 82, 80, 78, and 75% conserved relative to petunia, soybean, pea, and spinach, respectively. The Nicotiana tabacum transit peptide of 57 amino acids is comparable in length to the soybean transit of 55 amino acids,[17] and to the pea[14] and Lemna[19] transit peptides, which are also 57 amino acids long. The transit peptide for wheat, in contrast, has only 46 or 47 amino acids, depending on the gene from which it is derived.[18] The amino acid sequence of the tobacco transit peptide is most closely related to the soybean transit peptide, with which it is 70% conserved. The tobacco transit peptide is 61% conserved compared to that of pea, but only 46 and 37% conserved relative to the Lemna and wheat transit peptides, respectively. A comparison of these sequences is shown in Figure 4.

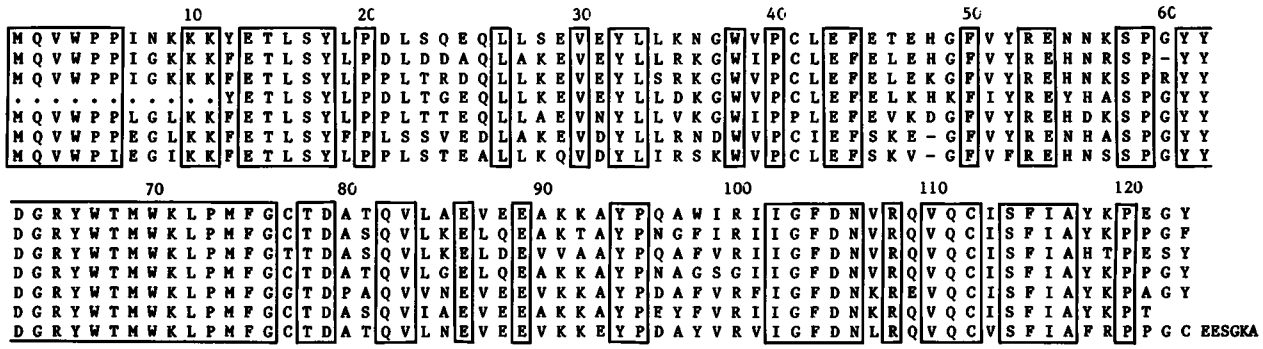
Introns in the SS genomic DNA

The tobacco SS gene has three introns. This is one more intron than is found in the soybean and pea SS genes, and two more introns than are found in the wheat SS genes.[14,15,17,18] This additional intron found in tobacco, a dicot, blurs the potential distinction between monocots and dicots as having one or two SS gene introns, respectively.[18] The third intron in tobacco is 3' to the wheat intron, and 3' to both soybean introns. It occurs between amino acids 65 and 66, in the middle of one of the most conserved regions of the protein. The position of this novel intron coincides with the boundary of a 12 amino acid insertion in the SS genes of higher plants, relative

TRANSIT PEPTIDE



MATURE PROTEIN



to those of the blue green algae Anacystis[37] and Anabaena[38]. This coincidence of intron and insertion boundaries is intriguing, particularly in light of theories that exons may encode movable functional domains[39]. The other two tobacco introns correspond in position exactly to those in soybean and pea, and are between amino acids 2 and 3, and between amino acids 47 and 48. The tobacco introns are 93 bp, 156 bp, and 147 bp long. The acceptor and donor sequences at the putative tobacco splicing sites are conserved compared to the animal consensus sequence.[35] Consensus animal branch point sequences are also present in each of the introns.

Sequence of the 5'and 3'non-coding regions

By comparing the tobacco DNA sequence with the genomic sequences of pea, soybean, and wheat, the 5' regulatory sequences can be inferred. The sequence TATATATA is believed to be important for RNA polymerase II recognition and is found preceding the sequenced SS genes.[14,15,17,18] It begins at base 954 in the tobacco sequence. The sequence CCAATG is found approximately 80 bp upstream of the TATATATA sequence in tobacco, at position 872. It is flanked nearby with the sequences TCAAGA and CCAAGA, at positions 857 and 890. These sequences have homology to regulatory sequences found upstream of other plant and animal genes.[40] A TAT sequence occurs about 30 bp downstream from the TATATATA sequence.

The 5' non-coding region may also contain sequences important in light regulation of the SS gene; such sequences have not yet been identified. We noted, however, that there are sequences in this region which are homologous to the viral and mammalian consensus enhancer sequence GTGG$^{AAA}_{TTT}$ or $^{TTT}_{AAA}$CCAC.[41-43] This same sequence occurs four times in the 5' region of the SS gene, and once each in the first and third introns of the gene. A single copy of this sequence and a variant of it have also been found upstream of the pea SS gene, and a variant of it has been been noted upstream of the wheat and soybean SS genes.[15]

Figure 4. Comparison of the sequences of the transit peptides[17] and mature SS proteins from tobacco, soybean,[14] pea,[16] spinach,[35,45] Lemna,[19] and wheat.[18] Conserved sequences are denoted by boxes.

Figure 5. Illustration of a possible hairpin structure in the 5' region of the tobacco RUBISCO SS gene. Of the 67 base pairs in the hairpin, 54 are perfect matches, 6 are G:T pairs, 5 are mismatches, and there are 2 single base loopouts. Enhancer-like sequences are underlined. The DNA from positions 93 to 406 is composed of the eight tandem direct repeats underlined in figure 3.

The 5' flanking DNA sequence in which several of the tobacco enhancer-like signals are embedded is unusual. Two of the signals are linked by inverted repeats which can potentially form a hairpin 67 base pairs long, as shown in figure 5. The loopout on such a hairpin, and part of the potential hairpin itself, is made up of eight tandem repeats of a 43 base sequence. The repeats are imperfect. Part of the DNA in each repeat is also capable of assuming hairpin structures, as the repeats themselves carry inverted repeats. It will be of interest to learn whether these sequences actually function as enhancers for the SS gene.

Like the pea, soybean and Lemna RUBISCO genes, the tobacco gene does not have a conventional polyadenylation signal. The sequence TAATAATAAT is the apparent signal in tobacco, compared to AATGAA in pea and soybean,[14,17] and AATTAA in Lemna.[19] This sequence occurs 160 bp after the TAA termination signal, at position 2144. Because of the near-identity between the 3' ends of the N. tabacum and the N. sylvestris SS genes, we expect that the poly A tail on the N. tabacum mRNA begins after position 2175.

CONCLUSION

We have cloned two RUBISCO SS genes from tobacco. These genes are not linked to each other, nor are they closely linked to other tobacco SS genes. Sequencing of one of these genes confirmed that it is closely related to the SS genes from other species, particularly those from dicots. The tobacco SS transit sequence, however, is much less conserved relative to the SS transit sequences of other species. The tobacco promoter and

transit sequences may therefore be the DNA of choice to use in gene fusions in which foreign genes are expressed in tobacco, and in fusions in which foreign proteins are directed to the chloroplast.

ACKNOWLEDGEMENTS

We would like to thank N.-H. Chua and colleagues for the pea cDNA RUBISCO clone, and R. Meagher and colleagues for the soybean RUBISCO genomic clone. We would like to thank C. J. Mauvais for doing Northern blot analyses of the SS mRNA, and S. C. Falco, N. S. Yadav, J. A. Rafalski and T. J. Andrews for helpful discussions.

REFERENCES

1. Miziorko, H.M. and Lorimer, G.H. (1983) Ann. Rev. Biochem. 52, 507-535.
2. King, S.D. (1976) Science 191, 429-434.
3. Coen, D.M., Bedbrook, J.R., Bogorad, L., and Rich, A. (1977) Proc. Natl. Acad. Sci. U.S.A. 74, 5487-5491.
4. Blair, G.E. and Ellis, R.J. (1973) Biochim. Biophys. Acta 319, 223-234.
5. Kawashima, N. and Wildman, S.G. (1972) Biochim. Biophys. Acta 262, 42-49.
6. Highfield, P.E. and Ellis, R. J. (1978) Nature 271, 420-424.
7. Chua, N.-H. and Schmidt, G.W. (1978) Proc. Natl. Acad. Sci. U.S.A. 75, 6110-6114.
8. Cashmore, A.R., Broadhurst, M.K. and Gray, R.E. (1978) Proc. Natl. Acad. Sci. U.S.A. 75, 655-659.
9. Smith, S.M. and Ellis, R.J. (1981) J. Mol. and Applied Genetics 1, 127-137.
10. Tobin, E.M. and Suttie, J.L. (1980) Plant Physiol. 65, 641-647.
11. Tobin, E.M. (1981) Plant Mol. Biol. 1, 35-51.
12. Bedbrook, J.R., Smith, S.M. and Ellis, R.J. (1980) Nature 287, 692-697.
13. Coruzzi, G., Broglie, R., Cashmore, A.R. and Chua, N.-H. (1983) J. Biol. Chem. 258, 1399-1402.
14. Cashmore, A.R. (1983) in Genetic Engineering of Plants, Kosuge, T., Meredith, C.P. and Hollaender, A. Eds., pp. 29-38, Plenum, New York.
15. Coruzzi, G., Broglie, R., Edwards, C. and Chua, N.-H. (1984) EMBO J. 3, 1671-1679.
16. Dunsmuir, P., Smith, S.M. and Bedbrook, J. (1983) Nucl. Acids Res. 11, 4177-4183.
17. Berry-Lowe, S.L., McKnight, T.D., Shah, D.M. and Meagher, R.B. (1982) J. Mol. and Appl. Gen. 1, 483-498.
18. Broglie, R., Coruzzi, G., Lamppa, G., Keith, B. and Chua, N.-H. (1983) Bio/Technology 1, 55-61.
19. Stiekema, W.J., Wimpee, C.F. and Tobin, E.M. (1983) Nucl. Acids Res. 11, 8051-8061.

20. Pinck, L., Fleck, J., Pinck, M., Hadidane, R. and Hirth, L. (1983) FEBS Lett. 154, 145-148.
21. Frischauf, A.-M., Lehrach, H., Poustka, A. and Murray, N. (1983) J. Mol. Biol. 170, 827-842.
22. Silhavy, T.J. Berman, M.L. and Enquist, L.W. (1984) Experiments with Gene Fusions, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.
23. Messing, J. (1983) in Methods in Enzymology, Wu, R., Grossman, L. and Moldave, K. Eds., pp. 20-78, Academic Press, New York.
24. Barnes, W., Bevan, M. and Son, P.H. (1983) in Methods in Enzymology, Wu, R., Grossman, L. and Moldave, K. Eds., pp. 98-122, Academic Press, New York.
25. Hanahan, D. (1983) J. Mol. Biol. 166, 557-580.
26. Bolivar, F., Rodriguez, R., Greene, P., Betlach, M., Heyneker, H., Boyer, H., Crosa, J. and Falkow, S. (1977) Gene 2, 95-113.
27. Broglie, R., Bellemaire, G., Bartlett S.G., Chua, N.-H. and Cashmore, A.R. (1981) Proc. Natl. Acad. Sci. U.S.A. 78, 7304-7308.
28. Riven, C.J., Zimmer, E.A. and Walbot, V. (1982) in Maize for Biological Research, Sheridan, W.F. Ed., pp. 161-164, Plant Mol. Biol. Assn., Charlottesville, Virginia.
29. Davis, R.W., Botstein, D. and Roth, J.R. (1980) Advanced Bacterial Genetics, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, U.S.A.
30. Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) Molecular Cloning, a Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.
31. Zimmerman, J.L. and Goldberg, R.B. (1977) Chromosoma 59, 227-252.
32. Benton, W. and Davis, R. (1977) Science 196, 180-181.
33. Sanger, F., Nicklen, S. and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. U.S.A. 74, 5463-5467.
34. Muller, K.-D., Salnikow, J. and Vater, J. (1983) Biochim. Biophys. Acta 742, 78-83.
35. Mount, S.M. (1982) Nucl. Acids Res. 10, 459-472.
36. Martin, P.G. (1979) Aust. J. Plant Physiol. 6, 401-408.
37. Shinozaki, K. and Sugiura, M. (1983) Nucl. Acids Res. 11, 6957-6964.
38. Nierzwicki-Bauer, S., Curtis, S.E. and Haselkorn, R. (1984) Proc. Natl. Acad. Sci. U.S.A. 81, 5961-5965.
39. Gilbert, W. (1978) Nature 271, 501.
40. Messing, J., Geraghty, D., Heidecker, G., Hu, N.-T., Kridl, J. and Rubenstein, I. (1983) in Genetic Engineering of Plants, Kosuge, T., Meredith, C.P. and Hollaender, A. Eds., pp. 211-227, Plenum, New York.
41. Weiher, H., Konig, M. and Gruss, P. (1983) Science 219, 626-631.
42. Gillies, S.D., Morrison, S.L., Oi, V.T. and Tonegawa, S. (1983) Cell 33, 717-728.
43. Weiher, H. and Botchan, R. (1984) Nucl. Acids Res. 12, 2901-2916.
44. Takruri, I.A.H., Boulter, D. and Ellis, R.J. (1981) Phytochemistry 20, 413-415.