Sequence of a ribulose 1,5-bisphosphate carboxylase/oxygenase cDNA from the C_4 dicot *Flaveria trinervia*

C.A.Adams, M.Babcock, F.Leung and S.M.Sun*

ARCO Plant Cell Research Institute, 6560 Trinity Court, Dublin, CA 94568, USA

Submitted January 12, 1987

We have cloned a full length cDNA for the small subunit of ribulose 1,5-bisphosphate carboxylase/oxygenase from the C_4 dicot Flaveria trinervia, deduced the complete nucleotide sequence of this cDNA and derived the amino acid sequence of the corresponding polypeptide. The cDNA includes nine nucleotides of 5' untranslated region just after the C-tail, the initiation codon ATG is nucleotides 26-28 and the stop codon is at nucleotides 545-547 (TAA) giving an open reading frame that could encode 173 amino acids. A putative polyadenylation signal is identified at nucleotides 732-737 (ATTAAA). This C_4 sequence shows the same level of homology to known C_3 sequences (1) as these C_3 sequences show among themselves. The area around the mature processing site (arrow) differs in the C_4 clone by having a lysine (broken line box) instead of a glutamine just after the methionine. The proposed intermediate processing site (2) also differs in the C_4 sequence in that the last amino acid is alanine (box) not serine. These differences raise the possibility of different processing determinants in C_4 versus C_3 chloroplasts.

```
                          M  A  S  I  P  A  T  V  A  A  V  P  Q  T  N  M  V  A  P  F  T  G  L  K  A  N  A  A
CCCCCCCCCCCCCCCCCTTATCATAATGGCTTCGATCCCCGCTACTGTCGCCGCCGTCCCTCAGACCAACATGGTGGCTCCGTTCACCGGACTTAAGGCCAACGCCGCCT110

F  P  V  T  K  K  V  N  G  F  S  T  L  P  S  N  G  G  R  V  Q  C↓M  K  V  W  P  P  V  G  K  K  K  Y  E  T  L
TCCCCGTCACCAAGAAGGTTAACGGCTTCTCCACCCTTCCCAGCAACGGTGGAAGAGTGCAATGCATGAAGGTGTGGCCCACCAGTTGGAAAGAAGAAGTACGAGACTCTT220

 S  Y  L  P  E  L  T  E  A  Q  L  A  K  E  V  D  Y  L  L  R  N  K  W  V  P  C  L  E  F  E  L  E  H  G  F  V  Y
TCATATCTTCCAGAGCTAACGGAGGCCCAGTTGGCTAAGGAAGTCGACTACCTCCTCCGCAACAAATGGGTTCCTTGTTTGGAATTCGAGTTGGAGCATGGTTTTGTGTA330

  R  E  N  A  S  S  P  G  Y  Y  D  G  R  Y  W  T  M  W  K  L  P  M  F  G  C  T  D  S  A  Q  V  M  K  E  L  Q
CCGTGAGAACGCAAGTTCTCCCGGATACTATGACGGAAGATACTGGACGATGTGGAAGTTGCCTATGTTCGGATGCACCGACTCAGCCCAAGTGATGAAGGAGCTTCAAG440

E  C  K  K  E  Y  P  Q  A  W  I  R  I  I  G  F  D  N  V  R  Q  V  Q  C  V  S  F  I  A  S  K  P  T  G  F
AGTGCAAGAAGGAGTACCCTCAGGCATGGATCCGTATCATCGGATTTGACAACGTCCGTCAAGTGCAATGCGTCAGTTTCATTGCCTCCAAACCAACTGGCTTCTAAGCA550

ACTTCTTGAACTATTCGATCTAACCTCTATATGATAGGTCGAGTTTGTTTGAATCTTTACAGTTGTTCATCAATTTGTTTTTCGAAATGGGAATTTCTCAATTTCATGTA660

TTTTGTTGTTTATTTCGCATTTGTTGTTTATCATCGGATTAGGCTCTTTGTGTTCCTTACTAACAACATATTATTAAAATGCAATTGTGTTTTAAAAAAAAAAA- to  850
```

*To whom correspondence should be addressed.

1) Stiekema, W.J. et al. Nucl. Acids Res. 22, 8051-8061 (1983).
2) Mishkind, M.L. et al. J. Cell Biol. 100, 226-234 (1985).