Sex Plant Reprod (1997) 10:181–183

SEQUENCE UPDATE

Li-Jia Qu · Yi Zhang · Ming Xie
Hongya Gu · Zhang-Liang Chen

A chalcone synthase-like cDNA from rice anther

Revision accepted: 4 April 1997

Key words *Oryza sativa* · cDNA · Chalcone synthase (CHS) · Anther development

Chalcone synthase (CHS) catalyses the first dedicated reaction in the flavanoid biosynthesis pathway, involving three rounds of condensation of malonyl-CoA onto 4-coumaroyl-CoA. Flavanoids are known to be involved in several important physiological processes in plants, such as flower pigmentation (Koes et al. 1990), UV stress protection (Wienand et al. 1986), nodulation (Recourt et el. 1992) and pollen development (Taylor and Jorgensen 1992; van der Meer et al. 1992). Chalcone synthase genes have been cloned from a number of species and are highly conserved among angiosperms and conifers (Fliegmann et al. 1992; Martin 1993).

To isolate genes expressed predominantly during anther development in rice (*Oryza sativa* var. *japonica*), a rice cDNA library (Lambda gt11, Promega) was constructed from mRNA prepared from anthers about 1.0 mm in length. Approximately 10^5 recombinant plaques were differentially screened with first-strand cDNA probes from developing anthers, leaves and roots, respectively. Ten clones were found to be anther probe-specific (data not shown). One of these clones, designated D5, was selected for further analysis. Northern blot analysis of RNA from anthers, roots, stems, leaves and seedlings using the D5 cDNA as probe showed a hybridisation band of approximately 1.6 kb in the anthers but none in the other tissues (data not shown). Southern blot analysis indicated that the D5 cDNA hybridised to single fragments from rice genomic DNA digested with several restriction enzymes (data not shown), showing that the D5 cDNA was derived from a single copy gene or low-copy number genes.

DNA sequence analysis revealed that the D5 cDNA is 1577 bp in length with a predicted open reading frame of 389 amino acid residues starting at the first ATG codon at position 53 and ending at position 1222 (Fig. 1). The 3′ end of the cDNA terminated at a poly(A) tail, preceded by a 278-bp untranslated AT-rich region containing a potential polyadenylation signal (AATACA) (Fig. 1). A database search using BLAST (Altschul et al. 1994) showed that the encoded polypeptide of the D5 cDNA is homologous to chalcone synthase sequences from plants with a minimal identity of 39% among 62 CHS sequences (including the sequence of the *Whp* gene, a maize CHS gene which plays a critical role in pollen viability) compared (data not shown). A short leucine zipper motif of four leucine residues has been found in most CHS and CHS-related proteins and is thought to play a role in protein-protein interaction (Kouzarides and Ziff 1988; Fliegmann et al. 1992). A similar motif has also been detected in the corresponding region of the predicted D5 polypeptide, although the second leucine was replaced by a phenylalanine through substitution of CTT with TTT (Fig. 1). Despite the overall similarity, the D5 polypeptide showed some interesting differences from other CHS sequences. For example, at the position 368–379 of the D5 polypeptide, a highly conserved region in other CHS proteins (WGVLFGFGPGLT), known as the CHS signature (Martin 1993), was replaced by the sequence WGLILAFGPGIT (Fig. 1). Furthermore, the D5 polypeptide diverges from another highly-conserved region around cysteine 170 thought to be an active site for substrate binding although the cysteine 170 itself is conserved (Martin 1993). These differences suggest that the D5 cDNA represents a novel member of the CHS gene family.

The nucleotide sequence data reported will appear in the EMBL database under accession number X91811

L.-J. Qu · Y. Zhang · M. Xie · H. Gu (✉) · Z.-L. Chen
The National Laboratory of Protein Engineering
and Plant Genetic Engineering, Peking University,
Beijing 100871, P.R. China
Tel. +86-10-6275-1847; Fax +86-10-6275-1841
e-mail: guhy@plum.lsc.pku.edu.cn

Z.-L. Chen
De Montfort University, The Gateway, Leicester,
LE1 9BH, UK

Fig. 1 DNA and polypeptide sequences of the D5 gene. The start and stop codons together with the cysteine 170 and the residues forming a short leucine zipper motif are shown in bold. A possible polyadenylation signal sequence is underlined. Computer analyses of DNA and amino acids were performed using the GCG package

```
  A GGA CTA GAA TAC TTG CTG ACT ACT GAG GCG AAA GGA TAC CGG GTG GCG AAG   52

  M   A   D   L   G   F   G   D   A   R   S   G   N   G   S   R   S   Q   18
ATG GCT GAC CTT GGA TTC GGC GAT GCC AGG AGT GGC AAT GGC AGC AGG AGC CAA  106
  C   S   R   G   K   A   M   L   L   A   L   G   K   G   L   P   E   Q   36
TGC TCC AGG GGG AAG GCG ATG CTG CTC GCC CTC GGC AAG GGC CTC CCT GAG CAA  160
  V   L   P   Q   E   K   V   V   E   T   Y   L   Q   D   T   I   C   N   54
GTT CTT CCC CAG GAG AAG GTC GTC GAG ACC TAC CTC CAG GAC ACC ATC TGC AAC  214
  D   P   A   T   R   A   K   L   E   R   L   C   K   T   T   T   V   R   72
GAT CCT GCA ACA AGG GCA AAG CTG GAA AGA CTT TGC AAG ACC ACC ACA GTG AGG  268
  T   R   Y   T   V   M   S   K   E   L   L   D   E   H   P   E   L   R   90
ACA AGG TAC ACT GTC ATG TCA AAG GAG CTC CTA GAC GAG CAC CCA GAG CTC AGG  322
  T   E   G   T   P   T   L   T   P   R   L   D   I   C   N   A   A   V  108
ACT GAG GGA ACT CCA ACA CTG ACG CCA CGG CTT GAC ATC TGC AAT GCT GCA GTG  376
  L   E   L   G   A   T   A   A   R   A   A   L   G   E   W   G   R   P  126
CTT GAG CTT GGT GCT ACT GCA GCC CGT GCC GCC CTT GGT GAA TGG GGG CGT CCA  430
  A   V   D   I   T   H   L   V   Y   I   S   S   S   E   L   R   L   P  144
GCA GTT GAC ATT ACC CAC CTT GTC TAC ATC TCG TCC AGT GAG CTT CGC CTC CCA  484
  G   G   D   L   F   L   A   T   R   L   G   L   H   P   N   T   V   R  162
GGG GGT GAC CTT TTC CTG GCA ACT CGC CTT GGC CTC CAT CCA AAC ACC GTC CGC  538
  T   S   L   L   F   L   G   C   S   G   G   A   A   A   F   R   T   A  180
ACT TCC CTT CTC TTC CTT GGC TGC TCC GGT GGC GCT GCC GCG TTC CGC ACC GCC  592
  K   D   I   A   E   N   N   P   G   S   R   V   L   V   V   A   A   E  198
AAG GAC ATT GCT GAG AAC AAC CCA GGG AGC CGC GTC CTT GTA GTA GCC GCG GAG  646
  T   T   V   L   G   F   R   P   P   S   P   D   R   P   Y   D   L   V  216
ACG ACG GTG CTG GGA TTC CGG CCA CCA AGT CCT GAC CGT CCT TAC GAT CTT GTT  700
  G   A   A   L   F   G   D   G   A   S   A   A   I   I   G   A   G   P  234
GGT GCT GCC CTG TTT GGT GAC GGC GCA TCA GCT GCG ATC ATT GGA GCA GGC CCC  754
  I   A   A   E   E   S   P   F   L   E   L   Q   F   S   T   Q   E   F  252
ATT GCT GCT GAG GAG AGT CCC TTC CTA GAG CTT CAG TTC TCA ACA CAG GAG TTC  808
  L   P   G   T   D   K   V   I   D   G   K   I   T   E   E   G   I   N  270
CTA CCA GGG ACG GAC AAG GTA ATT GAT GGC AAG ATC ACT GAG GAA GGG ATT AAT  862
  F   K   L   G   R   D   L   P   E   K   I   E   S   R   I   E   G   F  288
TTC AAA CTG GGG CGT GAT TTG CCC GAA AAG ATT GAA AGC CGT ATA GAA GGG TTC  916
  C   R   T   L   M   D   R   V   G   I   K   E   F   N   D   V   F   W  306
TGC AGG ACA CTC ATG GAT CGG GTT GGG ATA AAG GAG TTC AAT GAT GTA TTC TGG  970
  A   V   H   P   G   G   P   A   I   L   N   R   L   E   V   C   F   E  324
GCT GTG CAT CCT GGT GGT CCA GCA ATA CTG AAC AGG CTA GAG GTT TGC TTT GAA 1024
  L   Q   R   E   K   L   K   I   S   R   K   A   L   M   N   Y   G   N  342
CTC CAG CGA GAG AAG CTC AAG ATC AGT AGA AAG GCC CTG ATG AAC TAT GGT AAT 1078
  V   R   S   N   T   V   F   Y   V   L   E   Y   L   R   A   E   L   K  360
GTA CGC AGC AAC ACC GTC TTC TAT GTG TTG GAG TAT TTA AGG GCT GAG TTG AAG 1132
  K   G   M   I   R   E   E   W   G   L   I   L   A   F   G   P   G   I  378
AAA GGG ATG ATA AGG GAA GAA TGG GGA CTG ATC TTG GCT TTT GGC CCA GGC ATC 1186
  T   F   E   G   M   L   V   R   G   I   N   *                       390
ACA TTT GAA GGA ATG CTA GTT CGA GGC ATT AAC TGA GACTGAAGGGTCCAAGAAGACTT 1245
TCAGCTAGATGGAAGAATCACAAACATACCATTTCCAGGTTAAGTATATAGTATGACCTTGAACATCTCAT 1316
TAGGAGCATAGTTGTCTTTACAGCATCTGTGCACTCCATAGTTATTGTTTTACTGTAATCTCTGTTTATGT 1387
TTCAAACAATGTGAGGTACCCTAGTTTGAATGCCAGATACTTGGGAAAATTTCGGGCAATACATTGCATCC 1458
TGTTACATTGTCATGTATAGTAATTTTCCAGAGTATCTATTT(A)77                         1577
```

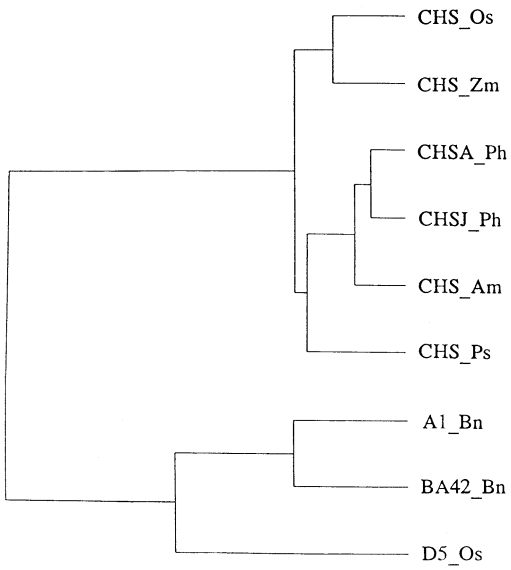

Fig. 2 Dendrogram based on amino acid sequence similarity by using the PILEUP program. CHS_Os, rice chalcone synthase (Reddy et al. 1996); CHS-Zm, chalcone synthase *Whp* in flowers of *Zea mays* (Franken et al. 1991); CHSA_Ph and CHS_Ph, two chalcone synthases in flowers of *Petunia hybrida* (Koes et al. 1989); CHS_Am, chalcone synthase in flowers of *Antirrhinum majus* (Sommer and Saedler 1986); CHS_Ps, chalcone synthase in *Pinus sylvestris* (Fliegmann et al. 1992); A1_Bn, peptide from *Brassica napus* anther-specific mRNA A1 (Scott et al. 1991); BA42_Bn, peptide from *B. napus* anther-specific mRNA BA42 (Shen and Hsu 1992); D5_Os, peptide from rice anther-specific mRNA D5 by the authors

The D5 polypeptide showed highest homology (68% and 65% identity, respectively) to peptides encoded by two anther-specific genes in *Brassica napus*, A1 and BA42 (Scott et al. 1991; Shen and Hsu 1992). The A1 gene was found to be expressed during the tetrad and microspore release stages of pollen development (Scott et al. 1991), whereas BA42 transcripts were detected only in the tapetum, periphery cells of the vascular bundle and developing microspores (Shen and Hsu 1992). However, the spatial distribution of the D5 transcript within anthers needs further analysis. The fact that D5, A1 and BA42 share the highest homology suggests that there may be a subgroup of CHS genes which are expressed specifically in anthers; this is also supported by a dendrogram (Fig. 2) derived from a phylogenetic analysis using the PILEUP program in the GCG package (Madison, Wis.).

Although the function of proteins encoded by this class of anther-expressed genes remains unclear, their homology to CHS and their expression patterns indicate that they play an important role in anther development. Further work is required to dissect their functions during anther development.

Acknowledgements This research was supported by the Chinese High-Tech Project and National Natural Sciences Foundation of China (Project GN 39670169 to L.-J. Qu). The authors wish to thank Drs. Y.-B. Xue, R. Dunford, C. Chiony and G. Moore (John Innes Centre, UK), Dr. R. Scott (University of Leicester, UK) and Dr. C. An (NLPE and PGE, China) for their valuable discussions and help with the manuscript.

References

Altschul SF, Boguski MS, Gish W, Wooton JC (1994) Issues in searching molecular sequence databases. Nat Genet 16: 119–129

Fliegmann J, Schroeder G, Schanz S, Britisch L, Schroeder J (1992) Molecular analysis of chalcone and dihydropinosylvin synthase from Scots pine (*Pinus sylvestris*), and different regulation of these and related enzyme activities in stressed plants. Plant Mol Biol 18: 489–503

Franken P, Niesbach-Klosgen U, Weydemann U, Marechal-Drouard L, Saedler H, Wienand U (1991) The duplicated chalcone synthase genes *C2* and *Whp* (white pollen) of *Zea mays* are independently regulated; evidence for translational control of *Whp* expression by the anthocyanin intensifying gene *in*. EMBO J 10: 2605–2612

Koes RE, Spelt CE, Elzen PJM van den, Mol JNM (1989) Cloning and molecular characterisation of the chalcone synthase multigene family of *Petunia hybrida*. Gene 81: 245–257

Koes RE, Blokland R van, Quattrocchio F, Tunen AJ van, Mol JNM (1990) Chalcone synthase promoters in petunia are active in pigmented and unpigmented cell types. Plant Cell 2: 379–392

Kouzarides T, Ziff E (1988) The role of the leucine zipper in the fos-jun interaction. Nature 336: 646–651

Martin CR (1993) Structure, function and regulation of the chalcone synthase. Int Rev Cytol 147: 233–284

Meer IM van der, Stam ME, Tunen AJ van, Mol JNM, Stuitze AR (1992) Antisense inhibition of flavonoid biosynthesis in petunia anthers results in male sterility. Plant Cell 4: 253–262

Recourt K, Tunen AJ van, Mur L, Brussel AAN van, Lugtenberg BJJ, Kijne JW (1992) Activation of flavonoid biosynthesis in roots of *Vicia sativa* subsp. *nigra* plants by inoculation with *Rhizobium leguminosarum* biovar *viciae*. Plant Mol Biol 19: 411–420

Reddy AR, Scheffler B, Madhuri G, Srivastava MN, Kumar A, Sathyanarayanan PV, Mair S, Mohan M (1996) Chalcone synthase in rice (*Oryza sativa* L.): detection of the CHS protein in seedlings and molecular mapping of the *chs* locus. Plant Mol Biol 32: 735–743

Scott R, Dagless E, Hodge R, Paul W, Soufleri I, Draper J (1991) Patterns of gene expression in developing anthers of *Brassica napus*. Plant Mol Biol 17: 195–207

Shen JB, Hsu FC (1992) Brassica anther-specific genes: characterisation and in situ localisation of expression. Mol Gen Genet 234: 379–389

Sommer H, Saedler H (1986) Structure of the chalcone synthase gene of *Antirrhinum majus*. Mol Gen Genet 202: 429–434

Taylor LP, Jorgensen R (1992) Conditional male fertility in chalcone synthase-deficient petunia. J Hered 83: 11–17

Wienand U, Weydemann U, Niesbach-Klosgen U, Peterson PA, Saedler H (1986) Molecular cloning of the *c2* locus of *Zea mays* the gene coding for chalcone synthase. Mol Gen Genet 203: 202–207