Plant Molecular Biology **20**: 333–336, 1992.
© 1992 *Kluwer Academic Publishers. Printed in Belgium.*

Update section

Short communication

Molecular cloning of the gene encoding developing seed L-asparaginase from *Lupinus angustifolius*

James M.J.J. Dickson, Eva Vincze, Murray R. Grant, Laura A. Smith, Karen A. Rodber, Kevin J.F. Farnden [1] and Paul H.S. Reynolds*

Immunology and Plant Molecular Biology Unit, Department of Scientific and Industrial Research-Fruit and Trees, Palmerston North, New Zealand (author for correspondence); [1] Biochemistry Department, Otago University, Dunedin, New Zealand*

Received 26 August 1991; accepted in revised form 28 April 1992

Key words: L-asparaginase, DNA sequence, developing seed, *Lupinus angustifolius* (lupin)

Abstract

A genomic sequence encoding *Lupinus angustifolius* L-asparaginase has been obtained, and is the first report of this gene from a plant source. The 3.2 kb of DNA sequenced contains a 1136 bp 5′ flanking sequence, four exons and three introns. Intron-exon borders were mapped by comparing the genomic sequence with that of a *L. arboreus* cDNA. Primer extension analysis revealed transcription start sites 16 bp and 13 bp 5′ of the initiating ATG for *L. angustifolius* and *L. arboreus*, respectively. The 5′ flanking region contained sequences associated with seed-specific expression.

In temperate legumes, of which lupin is an example, asparagine is the major compound formed from the assimilation of ammonia and plays a principal role in the transport and metabolism of nitrogen [11]. Asparagine is the major nitrogen compound translocated from the nodules in lupin and accounts for 50 to 70% of nitrogen supplied to lupin seeds [1]. During fruit maturation greater than 80% of asparagine entering the fruit is metabolised. The enzyme L-asparaginase hydrolyses the amide group of asparagine to produce aspartate and ammonia and thus provides a route for the utilisation of asparagine for the synthesis of amino acids and proteins. Indeed a range of studies have supported the view that asparaginase is a key enzyme involved in asparagine utilisation in lupin and plants in general [2, 18]. In previous work the developing seed L-asparaginase was purified and polyclonal antibodies raised against the protein. These antibodies precipitated asparaginase activity from a partially purified lupin seed extract. Partial protein sequence was obtained from polypeptides identified as asparaginase [12]. A *Lupinus arboreus* developing cDNA

The nucleotide sequence data reported will appear in the EMBL, GenBank and DDBJ Nucleotide Sequence Databases under the accession number X60691.

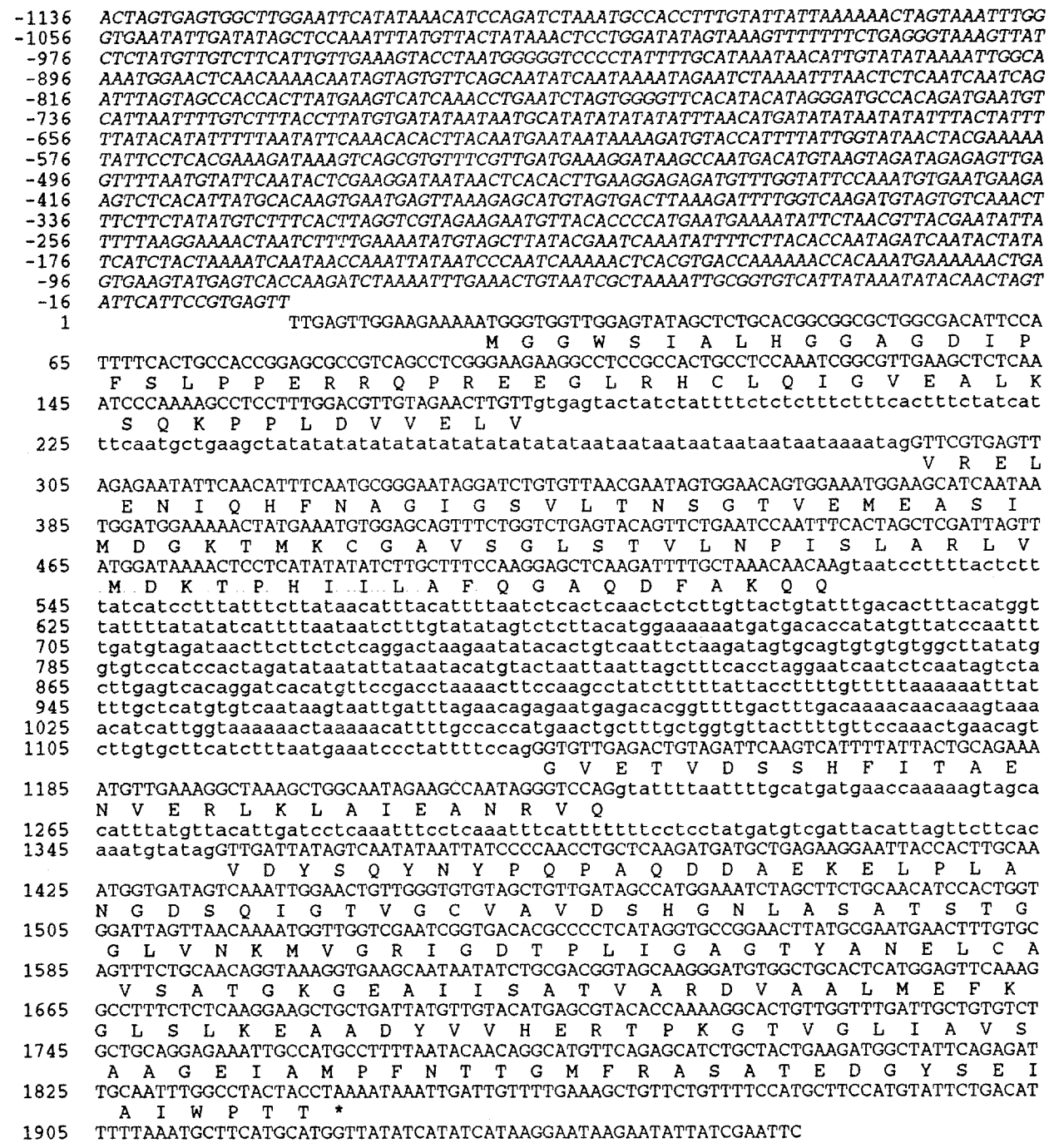

Fig. 1. Nucleotide sequence of the gene encoding L-asparaginase from *Lupinus angustifolius*. The 5′ flanking sequence is italicised. Coding region from the cap site is in upper case and introns are in lower case.

library was screened with polyclonal antibodies and a positive clone sequenced which, when back-translated, showed homology with sequenced asparaginase polypeptides [13]. This paper is the first report of the cloning of a plant gene encoding a L-asparaginase enzyme activity.

Genes encoding L-asparaginase have, however, been cloned from bacterial [6] and yeast [10] sources. The gene from *Lupinus arboreus* is known to be expressed in the developing seed [12] and northern hybridisation experiments have shown maximum mRNA levels 30 days after anthesis [13]. Other seed-specific and embryo-specific genes have been cloned including B-conglycinin [3], rice lectin [21] and pea convicilin [14]. These are all seed storage proteins, whereas L-asparaginase is an enzyme activity implicated in the amino acid metabolism of the developing seed. It has been suggested [15] that embryo development relies on the seed coat for the supply of amino acids.

DNA was prepared from nuclei obtained from the radicles of germinating seeds of *Lupinus angustifolius* [L] var. Uniharvest [5], the DNA partially digested with *Sau* 3aI and fragments of mean size 15 kb to 20 kb ligated into the lambda vector EMBL3 and packaged [8]. This library was screened by plaque hybridisation using a *Lupinus arboreus* developing seed L-asparaginase cDNA probe [13]. Four positive clones were identified from a screen of 250000 and one of these, pPMB100, was used in the sequencing experiments. DNA was sequenced by the dideoxy chain termination method [17] and the sequence assembled using the computer programs of Stockwell [20]. Intron-exon borders were mapped by comparison with the *L. arboreus* cDNAsequence[13].A39bpoligonucleotide,complementary to the RNA, was synthesized and this primer was annealed to and extended against RNA templates [9] from *L. arboreus* and *L. angustifolius* developing seed, and from *L. angustifolius* leaf, nodule and root. The primer corresponds to a position between 86 and 125 bp from the initiating ATG, and was chosen as there was absolute homology between the *L. angustifolius* genomic sequence and that of the *L. arboreus* cDNA clone [13] at this position. *L. arboreus* RNA gave a product 142 bp long, consistent with a 5′ untranslated region of 13 bp and *L. angustifolius* RNA gave a major product at 145 bp, consistent with a 5′ untranslated region of 16 bp (data not shown). An S1 nuclease protection [14] experiment confirmed 5′ untranslated regions of 13 bp and 16 bp, respectively for *L. arboreus* and *L. angustifolius* transcripts. Figure 1 shows the DNA sequence of a 3.1 kb *Eco* RI fragment obtained from pPMB 100. There is 94% homology between the *L. arboreus* cDNA sequence and the *L. angustifolius* genomic sequence. The gene consists of a 1136 bp 5′ flanking sequence, exon 1 of 180 bp, intron 1 of 112 bp, exon 2 of 234 bp, intron 2 of 615 bp, exon 3 of 87 bp, intron 3 of 126 bp and exon 4 of 492 bp (to the TAA stop codon). Intron-exon boundaries followed the general exon/GT..intron..AG/exon rule. Two consensus poly(A) addition signals were observed in the 3′ untranslated region.

Peptide sequence data obtained from purified *L. arboreus* asparaginase protein [12] was aligned against the genomic data. The first peptide bridged the gap between the 5′ terminus of the cDNA and the presumed initiating methionine residue of the full-length coding region. The other seven peptides were aligned with other regions of the gene, except for exon 3. Several elements previously reported to be associated with seed- or embryo-specific expression are present in the 5′ flanking region of the L-asparaginase gene:

– A6 bp repeat, A(AGC)CCCA, implicated in embryo-specific expression has been reported [3]; this 'ACCCCA' motif is present at position −292;

– Sequences called 'RY repeats' have been implicated in the regulation of seed lectin, legumin, vicilin and seed albumin genes [4], and such an element – AATGCATA – is present at −700;

– An (ATC)AACACA(AC)(ATC) consensus sequence is present in the 5′ region of many seed protein genes [7], and this motif is present at −632.

These three motifs are present in the L-asparaginase 5′ flanking region at positions considerably more distant to the cap site than in previous reports and, furthermore, they occur at only one position in the sequence.

As well as containing sequences associated with seed- and/or embryo-specific expression an AAAGAT motif associated with legume nodule/root expression [19] was seen at −610, −563 and −365. The significance of this observation is

not known. A number of other repeats of unknown function were also observed. Of particular interest is the sequence CTAAAATT which is repeated three times at −49, −77 and −166. The significance of this motif remains to be elucidated and is the subject of ongoing investigation, particularly as it is the only significant feature of the sequence identified (aside from the TATA and CAAT motifs) in the first 200 nucleotides 5′ from the cap site.

References

1. Atkins CA, Pate JS, Sharkey PJ: Asparagine metabolism – Key to the nitrogen nutrition of developing legume seeds. Plant Physiol 56: 807–812 (1975).
2. Chang KS, Farnden KJF: Purification and properties of asparaginase from *Lupinus arboreus* and *Lupinus angustifolius*. Arch Biochem Biophys 208: 49–58 (1981).
3. Chen Z-L, Schuler MA, Beachy RN: Functional analysis of regulatory elements in a plant embryo-specific gene. Proc Natl Acad Sci USA 83: 8560–8564 (1986).
4. Dickinson CD, Evans RP, Nielsen NC: RY repeats are conserved in the 5′-flanking regions of legume seed-protein genes. Nucl Acids Res 16: 371 (1988).
5. Evans IJ, James AM, Barnes SR: Organisation and evolution of repeated DNA sequences in closely related plant genomes. J Mol Biol 10: 803–826 (1983).
6. Gilbert HJ, Blazek R, Bullman HMS, Minton NP: Cloning and expression of the *Erwinia chrysanthemi* asparaginase gene in *Escherichia coli* and *Escherichia carotovora*. J Gen Microbiol 132: 151–160 (1986).
7. Goldberg RB, Barker SJ, Perez-Grau L: Regulation of gene expression during plant embryogenesis. Cell 56: 149–160 (1989).
8. Kaiser K, Murray NE: The use of phage lambda replacement vectors in the construction of representative genomic DNA libraries. In: Glover DM (ed) DNA Cloning: A Practical Approach, vol. 1, pp. 1–48. IRL Press, Oxford/Washington DC (1985).
9. Kingston RE: Primer extension. In: Ausubel F (ed) Current Protocols in Molecular Biology, pp. 4.7.1–3. Wiley/Greene (1987).
10. Kim K-Y, Kamerud JQ, Livinston DM, Roon RJ: Asparaginase II of *Saccharomyces cerevisiae*. J Biol Chem 263: 11948–11953 (1988).
11. Lea PJ, Miflin BJ: Transport and metabolism of asparagine and other nitrogen compounds within the plant. In: Miflin BJ (ed) The Biochemistry of Plants: A Comprehensive Treatise, vol. 5: Amino Acids and Derivatives, pp. 569–608. Academic Press, New York (1980).
12. Lough TJ, Chang KS, Carne A, Monk BC, Reynolds PHS, Farnden KJF: Developmental regulation, purification and peptide sequences of L-asparaginase from developing seeds of lupin (*Lupinus arboreus*). Phytochemistry, in press (1992).
13. Lough TJ, Reddington B, Grant MR, Hill DF, Reynolds PHS, Farnden KJF: The molecular cloning of L-asparaginase from developing seeds of lupin (*Lupinus arboreus*). Plant Mol Biol, in press (1992).
14. Newbigin EJ, deLumen BO, Chandler PM, Gould A, Blagrove RJ, March JF, Korrt AA, Higgins TJV: Pea convicilin: structure and primary sequence of the protein and expression of a gene in the seeds of transgenic tobacco. Planta 180: 461–470 (1990).
15. Ratajczak W: Asparagine metabolism in developing seeds of *Lupinus luteus* L. Biochem Physiol Pflanzen 181: 17–22 (1986).
16. Sambrook J, Fritsch EF, Maniatis T: Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY (1989).
17. Sanger F, Micklen S, Coulson AR: DNA sequencing with chain terminating inhibitors. Proc Natl Acad Sci USA 74: 5463–5467 (1977).
18. Sieciecchowicz KA, Joy KW, Ireland RJ: The metabolism of asparagine in plants. Phytochemistry 27: 663–671 (1988).
19. Stougaard J, Sandal NN, Gron A, Kuhle A, Marcker KA: 5′ analysis of the soybean leghaemoglobin *lbc3* gene: regulatory elements required for promoter activity and organ specificity. EMBO J 6: 3565–3569 (1987).
20. Stockwell PA: DNA sequence analysis software. In: Bishop MJ, Rawlings CJ (eds) Nucleic Acid and Protein Sequence Analysis: A Practical Approach, pp. 19–45. IRL Press, Oxford (1987).
21. Wilkins TA, Raikhel NV: Expression of rice lectin is governed by two temporally and spatially regulated mRNAs in developing embryos. Plant Cell 1: 541–549 (1989).