Cloning and expression of diadenosine 5′,5‴-*P*¹,*P*⁴-tetraphosphate hydrolase from *Lupinus angustifolius* L

Danuta MAKSEL*, Andrzej GURANOWSKI†, Steven C. ILGOUTZ*, Arthur MOIR‡, Michael G. BLACKBURN‡ and Kenwyn R. GAYLER*[1]

*Department of Biochemistry and Molecular Biology, University of Melbourne, Parkville, Victoria 3052, Australia, †Department of Biochemistry and Biotechnology, Agricultural University, Poznan, Poland, and ‡Department of Chemistry, University of Sheffield, Sheffield S3 7HF, U.K.

The first isolation, cloning and expression of cDNA encoding an asymmetric diadenosine 5′,5‴P^1,P^4-tetraphosphate pyrophosphohydrolase (Ap_4A hydrolase) from a higher plant is described. Ap_4A hydrolase protein was purified from seeds of both *Lupinus luteus* and *Lupinus angustifolius* and partially sequenced. The Ap_4A hydrolase cDNA was cloned from *L. angustifolius* cotyledonary polyadenylated RNA using reverse transcription and PCR with primers based on the amino acid sequence. The cDNA encoded a protein of 199 amino acids, molecular mass 22982 Da. When expressed in *Escherichia coli* fused to a maltose-binding protein, the enzyme catalysed asymmetric cleavage of Ap_4A to AMP and ATP which was inhibited at concentrations of F^- as low as 3 μM. These are properties characteristic of Ap_4A hydrolase (*asymmetrical*) (EC 3.6.1.17). Comparison of the Ap_4A hydrolase sequences derived from the four known cDNAs from pig, human, lupin and fission yeast showed that, like the mammalian hydrolase, the lupin enzyme possesses a Mut T motif but no other significant similarities. No sequence similarity to the human fragile histidine triad protein, as found in the Ap_4A hydrolase from *Schizosaccharomyces pombe*, was detected in the Ap_4A hydrolase from lupin.

INTRODUCTION

Diadenosine 5′,5‴P^1,P^4-tetraphosphate (Ap_4A) is found in virtually all cells at concentrations ranging from 0.05 to 5 μM [1,2]. Roles proposed for intracellular Ap_4A include involvement in cellular responses to metabolic stress in micro-organisms [3] and in DNA replication in eukaryotic cells [4]. In higher organisms, Ap_4A is involved in several other physiological responses. Ap_4A and related nucleotides have been implicated in neurotransmission through both A_1 and P_2 receptors in the nervous system [5–7]. Ap_3A, Ap_4A and ATP also have the capacity to interact with granulocyte–macrophage colony-stimulating factor during inflammation to delay neutrophil apoptosis [8]. In addition, Ap_4A is a well-established competitive inhibitor of the ADP-induced platelet aggregation associated with blood clotting [9]. As a consequence, analogues of Ap_4A are under investigation as potential therapeutic agents [10] and this has led in turn to a great deal of interest in accurately describing enzymes capable of the catabolism of Ap_4A and its analogues. This paper is part of an investigation of the properties of enzymes capable of the specific hydrolysis of Ap_4A aimed ultimately at clarification of the structures of their catalytic sites.

Two major types of enzyme capable of the specific hydrolysis of Ap_4A have been described. The Ap_4A hydrolase that has been described in the slime mould *Physarum polycephalum* [11] and in *Escherichia coli* [12,13] is a (*symmetrical*) Ap_4A hydrolase (EC 3.6.1.41) capable of splitting Ap_4A into two ADPs. By contrast, most of the Ap_4A hydrolases of eukaryotes characteristically catalyse the asymmetric cleavage of Ap_4A to AMP and ATP. Ap_4A hydrolase (*asymmetrical*) (EC 3.6.1.17) has been purified from a range of eukaryotes including the fission yeast *Schizosaccharomyces pombe* [14], the green alga *Scenedesmus obliquus* [15], humans and other mammals [16–19], and plants [20,21]. Seeds of the grain legumes *Lupinus luteus* and *Lupinus angustifolius* are a rich source of this enzyme [20], although, apart from suggestions that link Ap_4A hydrolase to stress responses in tomatoes [21], there is no other information on the role that it might play in plants. The mechanism of action of the lupin enzyme is, however, reasonably well understood. Detailed models of the mechanism of cleavage of Ap_4A and other nucleotides by the plant enzyme have been developed on the basis of studies including ^{18}O MS [22]. Clear distinction has been drawn between the mode of action of this eukaryote asymmetrical Ap_4A hydrolase and that of the symmetrical Ap_4A hydrolase from prokaryotes. Access to cloned and expressed protein for the eukaryote enzyme is now sought as the basis for future structural studies to confirm or modify the mechanistic proposals that have been made.

The cDNA sequences described so far for asymmetric Ap_4A hydrolases, those from *S. pombe* [23], from humans [24] and from pigs [25], show a clear distinction between these enzymes. Whereas the two mammalian cDNAs contain sequences corresponding to the Mut T motif of the prokaryote Mut T protein [24,25], the fission-yeast cDNA contains sequence with substantial similarity to that encoding the human fragile histidine triad (FHIT) protein [26]. Human Ap_4A hydrolase cDNA, however, shows no similarity to cDNA encoding the FHIT protein [26]. Rather, the FHIT protein, which is implicated in tumour suppression in humans [26], has recently been shown to be an Ap_3A hydrolase [27]. The kinetic properties of the asymmetrical Ap_4A hydrolases from plants [20,21] suggest that they are more closely related to the mammalian Ap_4A hydrolases

Abbreviations used: Ap_5A, diadenosine 5′,5‴P^1,P^5-pentaphosphate; Ap_4A, diadenosine 5′,5‴P^1,P^4-tetraphosphate; Ap_3A, diadenosine 5′,5‴P^1,P^3-triphosphate; p_4A, adenosine 5′-tetraphosphate; FHIT, fragile histidine triad; RACE, rapid amplification of cDNA ends; IPTG, isopropyl β-D-thiogalactopyranoside; MBP–Ap_4AH fusion protein, maltose-binding protein–Ap_4A hydrolase fusion protein; RT, reverse transcriptase; ORF, open reading frame; EST, expressed sequence tag; poly(A)$^+$, polyadenylated.

[1] To whom correspondence should be addressed.

than to the hydrolase from *S. pombe*. In this paper we describe the first cloning of cDNA encoding an asymmetric Ap_4A hydrolase from a higher plant. We show that the cDNA encodes a protein belonging to the group of proteins characterized by the presence of a Mut T motif in the sequence and demonstrate that an active asymmetric Ap_4A hydrolase can be produced from the cDNA using a bacterial expression system.

EXPERIMENTAL

Materials

Ap_4A hydrolase was purified from whole dried seeds harvested at maturity from field-grown *L. angustifolius* L. cv Unicrop and *L. luteus* cv Topaz. Cotyledons from *L. angustifolius* harvested 30–40 days after flowering and frozen at −70 °C after removal of the seed coats were used for RNA isolation.

[α-^{32}P]ATP (4000 Ci/mmol) was from Bresatec. Ap_4A and p_4A were from Sigma. All other chemicals used were either of analytical or electrophoresis grade.

Purification of Ap_4A hydrolase from *Lupinus* spp.

Ap_4A hydrolase was purified from the mature lupin seeds by a modification of the method of Jakubowski and Guranowski [20], which included as the final purification step affinity chromatography on a Remazol-Red–Sepharose column equilibrated in 50 mM potassium phosphate buffer, pH 6.8, containing 5% (v/v) glycerol and 1 mM 2-mercaptoethanol, and eluted with the same buffer containing 5 mM $MgCl_2$ and 50 μM p_4A. The enzyme purified in this way was found to be blocked at the N-terminus and hence could not be sequenced directly. Instead, pure protein was digested with Pronase from *Streptomyces*, and peptides of 16 and 13 kDa were separated by PAGE and collected with a Bio-Rad Transblot system. N-Terminal sequence (1) [G]SDWKG-QAQKWFLFK[F]T was obtained from the internal 13 kDa peptide by Edman degradation on an ABI476A protein sequencer. Internal sequence (2) PRDAAIRELRE[R]T was similarly obtained on material obtained by separate treatment of the pure protein with formic acid.

cDNA synthesis

Polyadenylated [poly(A)$^+$]RNA was isolated from 50 mg of frozen lupin cotyledons on 0.5 mg of Dynabeads Oligo(dT)$_{25}$ (Dynal, Oslo, Norway). An immobilized unamplified single-stranded cDNA library was constructed by incubating beads containing the poly(A)$^+$ RNA with 50 μl of 10 mM Tris/HCl buffer, pH 8.3, containing 50 mM KCl, 5 mM $MgCl_2$, 1 mM each dNTP and 125 units of murine leukaemia virus reverse transcriptase (RT; Perkin–Elmer) at 42 °C for 1 h. RNA was removed from the beads by heating at 95 °C and the Dynabead-bound cDNA washed with 100 μl of 20 mM Tris/HCl buffer, pH 8.4, containing 50 mM KCl, 1.5 mM $MgCl_2$ and 0.1 mg/ml BSA and suspended in 100 μl of the same buffer. Dynabead-bound cDNA (25 μg) was used as template in a PCR with degenerate primers, Ap_4A-1 (5′-GCT/CTGA/G/T/CCCT/CTTCCAA/GTC-3′) and Ap_4A-4 (5′-CCIA/CGIGAT/CGCIGCIATT/C/AA/CG-3′), which were based on the protein sequences from Ap_4A hydrolase fragments (1) and (2) respectively. Sequence from the 155 bp-long partial cDNA fragment (AH-2) obtained in this way was used to design primers for subsequent rapid amplification of cDNA ends (RACE) reactions (Figure 1). A library of double-stranded cDNAs was also synthesized from poly(A)$^+$ RNA, after its elution from the beads, using a combination of RT and DNA polymerase from the Marathon cDNA Amplification Kit (Clontech, Palo Alto, CA, U.S.A.). An adaptor which contained the binding site for the 3′ and 5′ RACE adaptor primer, Ap1, was ligated to each end of the double-stranded cDNA made in this way. In a 3′ RACE reaction using double-stranded cDNA as template, the gene-specific forward primer Ap_4A-9 (bases 52–74 of fragment AH-2) and adaptor-specific primer Ap1 produced fragment AH-3 spanning 489 bp at the 3′ end of Ap_4A hydrolase cDNA. Reverse primer Ap_4A-12 based on the sequence of fragment AH-3 was used with the adaptor primer Ap1 in a 5′ RACE reaction to generate fragment AH-5 which spanned 581 bp at the 5′ end of the Ap_4A hydrolase cDNA. The 3′ and 5′ fragments AH-3 and AH-5 contained 251 bp of overlapping and identical sequence and together were equivalent to a full-length clone of 819 bp (AH-6). The entire 603 bp coding region of the putative Ap_4A hydrolase was also obtained as a continuous cDNA fragment (AH-7) by PCR from both single-stranded cDNA on Dynabeads and double-stranded cDNA using primers Ap_4A-17 and Ap_4A-18 (5′-GC**TCTAGA**ATGGCTTTATGCCGATTGG-3′) and Ap_4A-18 (5′-GC**TCTAGA**TTATTAGAGATGTGGAGCAAAC-3′) based on the putative sequence of AH-6 and which also introduced *Xba*I-recognition sites at both ends of the clone. PCR fragment AH-7–*Xba* was cloned into pCR Script Cloning Vector from pCR Script SK(+) Cloning Kit (Stratagene) for maintenance (pAH-7) and into the *Xba*I site of vector pMAL-c2 (New England Biolabs) for expression (pMALAH-7).

Expression and purification of maltose-binding protein–Ap_4A hydrolase (MBP–Ap_4AH) fusion protein

Luria broth (100 ml) containing 0.2% (w/v) glucose and 100 μg/ml ampicillin was inoculated with 1 ml of an overnight culture of Epicurian Coli® XL1-Blue cells of *E. coli* containing the fusion plasmid. Cells were grown with shaking at 37 °C to $A_{600} \cong 0.5$, then isopropyl β-D-thiogalactopyranoside (IPTG) was added to a concentration of 0.3 mM and incubation continued for 2–3 h. Cells were harvested by centrifugation for 10 min at 4000 ***g***, resuspended in 5 ml of 20 mM Tris/HCl buffer pH 7.4, containing 200 mM NaCl, 1 mM EDTA, 1 mM dithiothreitol and 5% (v/v) glycerol and frozen at −20 °C. Thawed cells were sonicated (5 × 12 s), debris was removed by centrifugation, and soluble fusion protein bound to amylose resin (New England Biolabs) equilibrated in the above buffer. After the column had been washed with 12 vol. of column buffer, the fusion protein was eluted with the same buffer containing 10 mM maltose.

Assays for Ap_4A hydrolase (*asymmetrical*) activity

Ap_4A hydrolase activity was assayed in 50–200 μl of 50 mM Tris/HCl buffer, pH 8, containing 5 mM $MgCl_2$ 0.1 mM dithiothreitol and 0.5 mM Ap_4A at room temperature for up to 30 min. For HPLC analysis, reactions were stopped with 0.7% (w/v) $HClO_4$, the mixture centrifuged and the supernatant filtered through a 0.45 μm filter (Millipore) before 20 μl aliquots were analysed on a Partisil 10 SAX column (Whatman) equilibrated with 1.8 M ammonium formate and eluted at 1 ml/min with a linear gradient of 1.8–3 M ammonium formate, for 20 min. Alternatively the reaction was stopped by spotting a 2 μl aliquot on to a TLC plate coated with silica gel with fluorescent indicator (Merck 5554). Plates were developed in dioxan/ammonia/water (6:1:4, by vol.), and the products of Ap_4A hydrolysis visualized under UV light. To test the inhibition of Ap_4A hydrolase activity by F^-, limiting amounts of MBP–Ap_4AH fusion protein were preincubated for 5 min in the assay mixture containing NaF at concentrations ranging from 3 to 100 μM before substrate was added and incubation continued. Reaction rates, which were

linear throughout the period of the experiment, were determined by analysis of Ap_4A, AMP and ATP levels by HPLC in aliquots removed at time intervals up to 10 min.

Protein analysis

SDS/PAGE analysis of proteins was performed using the discontinuous buffer system of Laemmli [28] and gels were silver stained (Bio-Rad). To determine the purity of expressed protein fractions, gels were scanned in a Molecular Dynamics Densitometer after staining with 0.1 % (w/v) Coomassie Brilliant Blue R250 in methanol/acetic acid/water (5:1:5, by vol.). Protein concentrations were estimated using the bicinchoninic acid assay as described by Redinbaugh and Turley [29].

DNA manipulation

DNA sequence analysis was performed on an ABI model 373A DNA Sequencer using the PRISM Ready Reaction Dye-Deoxy Terminator Cycle Sequencing Kit from Applied Biosystems. The DNA sequence in Figure 1 was fully determined on both strands and on several independently generated clones. Analysis of the DNA sequence was performed using the Australian National Genomic Information Service and the programs of Staden, and at the National Centre for Biotechnology Information using the BLAST network service. Similarities between sequences were calculated using the GAP program.

RESULTS AND DISCUSSION

Cloning of lupin Ap_4A hydrolase

Asymmetric Ap_4A hydrolase was purified from mature seeds of both *L. angustifolius* and *L. luteus* and partial amino acid sequences determined on internal fragments obtained by cleavage of each of the purified enzymes with either *Streptomyces* Pronase or formic acid. Whereas the intact Ap_4A hydrolase from both sources was N-terminally blocked, internal sequences of [G]SDWKGQAQKWFLFK[F]T (1) and PRDAAIRELRE[R]T (2) were obtained from Pronase and formic acid fragments respectively of the *L. luteus* Ap_4A hydrolase. The latter sequence was also obtained from *L. angustifolius* Ap_4A hydrolase. Degenerate inosine-containing oligonucleotide primers corresponding to the sequences encoding each of these fragments were used in RT-PCR from *L. angustifolius* cotyledonary poly(A)$^+$ RNA. As the order of fragments (1) and (2) within the protein was unknown, primers allowing for both possible combinations were tested. As shown in Figure 1(A), PCRs using primers that assumed that sequence 2 was positioned nearer to the N-terminus produced a partial cDNA segment (AH-2) 155 bp long. Ap_4A hydrolase-specific primer Ap_4A-9 designed from the sequence of partial cDNA AH-2 was used in combination with the adaptor primer Ap1 in a 3′ RACE reaction to obtain cDNA AH-3 which spanned the 3′ end of Ap_4A hydrolase through to the poly(A)$^+$ tail. Ap_4A hydrolase-specific primer Ap_4A-12 designed from the sequence of partial cDNA AH-3 was then used in combination with the adaptor primer Ap1 in a 5′ RACE reaction to obtain cDNA AH-5 which spanned the 5′ end of Ap_4A hydrolase. Fragments AH-3 and AH-5 produced in this way together contained 819 bp cDNA spanning the entire coding region of a putative Ap_4A hydrolase. As shown in Figure 1(B), the complete open reading frame (ORF) had two potential start sites for translation, one at base 22 leading to a protein of apparent molecular mass 22982 Da and one at base 139 leading to a protein of molecular mass 18404 Da. Of the two ATG start codons in the ORF, the first ATG was assigned as the likely predominant start site on the basis that it is (i) both in the context of T**G**GA<u>ATG</u>**G**CT which includes both the G at +4 and the purine at −3 relative to the ATG codon which are the preferred conserved bases in eukaryote initiation sites GCC**A/G**CC<u>ATG</u>**G** [30] and (ii) it is the furthest 5′. Bases $A_{22}TGGCT_{27}$ also encode Met-Ala as the first residues in the protein, a feature widespread in plant proteins [31]. The enzyme initiated at this ATG includes a putative signal sequence at the N terminus. The mature protein derived from such a pre-protein would have Ser^{39} as its N-terminus after cleavage between Ser^{38} and Ser^{39}, the most likely signal-peptide-cleavage site predicted within the ORF by the method of von Heijne (Australian National Genomic Information Service). Such a protein would have an apparent molecular mass of 18491 Da. The protein produced if the alternative ATG at base 139 were used for initiation would commence at Met^{40} and have an apparent molecular mass of 18404 Da. It is possible therefore that the protein is produced both as a secretory protein initiated at Met^1 and subsequently cleaved at Ser^{39} and also as a non-secreted protein initiated at Met^{40}. The mature products would contain either Ser^{39} or Met^{40} as the respective N-terminal residues of the mature Ap_4A hydrolase *in vivo*. Verification of this by direct protein sequencing was not possible as the N-terminus of the enzyme purified from lupin seeds was blocked. However, the apparent molecular mass of about 19 kDa estimated by SDS/PAGE for the native enzyme purified from *L. angustifolius* (Figure 2) and the estimates of 18–18.5 kDa for the Ap_4A hydrolase from *L. luteus* [20] are both consistent with the size predicted for the truncated protein.

Most asymmetric Ap_4A hydrolases so far characterized have similar molecular masses of between 17500 [32] and 19800 [33] Da when assayed by SDS/PAGE. However, estimates of 16824 and 16743 Da obtained for the human Ap_4A hydrolase by electrospray MS, which matched the apparent value of 16829 Da for the peptide encoded by the ORF of cloned human Ap_4A hydrolase [24], suggested that the value of 19200 Da estimated by SDS/PAGE [18] was anomalously high. On the basis of the cDNA sequence, we conclude that the lupin enzyme is nevertheless significantly larger than the human enzyme.

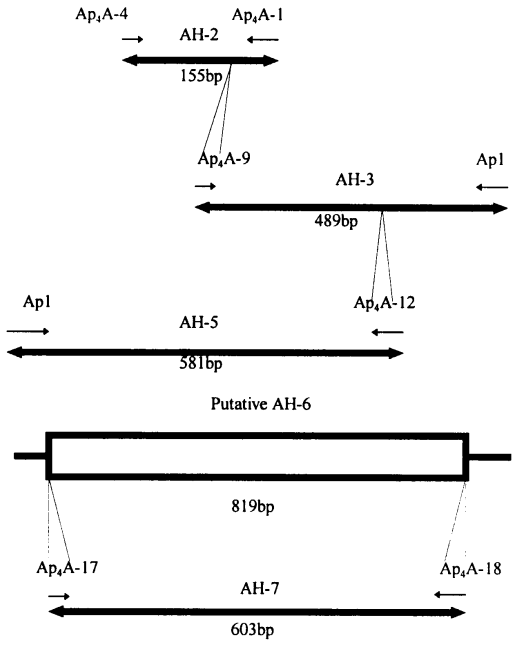
Expression of lupin Ap_4A hydrolase in *E. coli*

The region of plasmid pAH-7 corresponding to the entire coding region of the putative Ap_4A hydrolase was obtained as a single fragment AH-7 by PCR using primers Ap_4A-17 and Ap_4A-18 based on the sequences of AH-5 and AH-3 respectively (Figure 1). The 603 bp fragment AH-7 was cloned into the expression vector pMAL-c2 creating plasmid pMALAH-7 and expressed in *E. coli* fused with MBP. Fusion protein of the expected size 67200 Da, was released by sonication from IPTG-induced *E. coli* cells. The yield of fusion protein was estimated by laser densitometry of proteins separated by SDS/PAGE to be 10 % of the total soluble protein released.

Expressed fusion protein (MBP–Ap_4A) was purified by affinity chromatography on amylose resin, and its ability to cleave Ap_4A was tested. As shown in Figure 3, the purified fusion protein catalysed asymmetric cleavage of Ap_4A. ATP and AMP alone were produced in stoichiometric amounts from Ap_4A as detected by HPLC. Products were also separately identified by TLC as described in the Experimental section (results not shown). These expression studies demonstrated that the sequence encoded by the lupin clone AH-7 was that of an asymmetrical Ap_4A hydrolase.

There was no evidence that the fusion protein could catalyse symmetrical hydrolysis of Ap_4A to ADP, the activity previously characterized in *E. coli* itself [12,13]. The fact that Ap_4A hydrolysis catalysed by the fusion protein occurred in the presence

A



B

```
1   CTCTCCTTTTCATCGAGTGGAATGGCTTTATGCCGATTGGTACAATCTCCTCCTACCAATTTTCATTTCC   70
                    M  A  L  C  R  L  V  Q  S  P  P  T  N  F  H  F  R

71  GAAAATACCCTTCTAAATTCCTTAAATTCTCTTCTCTGTCACTCGCATTCCGTTATTGCCATTCTTCAAT  140
      K  Y  P  S  K  F  L  K  F  S  S  L  S  L  A  F  R  Y  C  H  S  S  M

141 GGATTCTCCTCCCGAAGGTTATCGCAGAAACGTTGGAATCTGTCTCATGAATAATGATAAAAAGATTTTT  210
      D  S  P  P  E  G  Y  R  R  N  V  G  I  C  L  M  N  N  D  K  K  I  F

211 GCAGCTTCACGCCTCGACATACCCGATGCTTGGCAAATGCCGCAGGGTGGTATTGATGAAGGCGAGGATC  280
    A  A  S  R  L  D  I  P  D  A  W  Q  M  P  Q  G  G  I  D  E  G  E  D  P

281 CAAGAAATGCAGCTATCAGGGAATTAAGAGAAGAGACAGGAGTTACTTCTGCAGAAGTGATTGCAGAGGT  350
      R  N  A  A  I  R  E  L  R  E  E  T  G  V  T  S  A  E  V  I  A  E  V

351 GCCTTATTGGTTAACATATGACTTCCCACCGAAAGTCAGAGAGAAATTAAATATTCAATGGGGATCTGAT  420
      P  Y  W  L  T  Y  D  F  P  P  K  V  R  E  K  L  N  I  Q  W  G  S  D

421 TGGAAGGGCCAAGCACAGAAATGGTTTCTTTTTAAGTTCACTGGGCAGGACCAAGAAATCAATCTTTTAG  490
    W  K  G  Q  A  Q  K  W  F  L  F  K  F  T  G  Q  D  Q  E  I  N  L  L  G

491 GTGATGGTTCTGAGAAACCTGAGTTTGGGGAATGGTCATGGGTCACACCAGAACAATTAATTGACCTTAC  560
      D  G  S  E  K  P  E  F  G  E  W  S  W  V  T  P  E  Q  L  I  D  L  T

561 AGTGGAATTCAAGAAGCCTGTCTACAAGGAAGTCCTATCAGTGTTTGCTCCACATCTCTAATAATCTTAT  630
      V  E  F  K  K  P  V  Y  K  E  V  L  S  V  F  A  P  H  L  *  *

631 CAGTTGCCTTGTATCTTATCTGTTGCAACATTTTCATTTAAAGTGCCCATAAAATGAGATTTATTTCGGT  700

701 TTTTTTTGTAATACTCAACCTATGCCGTTCTTAATCTTACTTTTCTGAATTAATAGAGCTTCATCTCCAA  770

771 ATTACGGTCCTGATTTCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 819
```

Figure 1 Ap_4A hydrolase cloned by RT-PCR and 3′,5′ RACE

(**A**) RT and PCR products produced from poly(A)$^+$ RNA. Primers Ap_4A-1 and Ap_4A-4 were derived from Ap_4A hydrolase protein sequence. Primers Ap_4A-9 to 18 corresponded to the sequences of AH-3,5 and 6 as indicated and Ap1 to the sequence of the cDNA adaptor primer. (**B**) cDNA and deduced amino acid sequence of lupin Ap_4A hydrolase. Nucleotides are numbered from the 5′ end of the cDNA sequence. Amino acid sequences established by protein sequencing of purified lupin Ap_4A hydrolase are underlined. Bases 22–625 (AH-7) were obtained by independent RT-PCR and cloned into pMAL-c2 for expression.

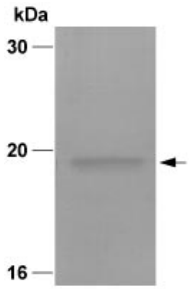


Figure 2 Ap_4A hydrolase from *L. angustifolius*

SDS/PAGE of Ap_4A hydrolase purified from mature seeds of *L. angustifolius* and silver stained. Molecular masses assigned from standard proteins are marked.

of Mg^{2+} and in the absence of Co^{2+}, an essential cofactor for the symmetrical *E. coli* Ap_4A hydrolase, is further evidence that the clone AH-7 encodes a eukaryotic/lupin asymmetrical Ap_4A hydrolase.

In view of the presence in the coding region of a putative cleavage site as $Ser^{38/39}$ (Figure 1), a truncated version of the Ap_4A hydrolase starting from Met^{40} was also expressed as an MBP-fusion protein. This also showed asymmetrical Ap_4A hydrolase activity equivalent to that obtained with the full-length MBP–Ap_4AH. Estimates by SDS/PAGE and laser densitometry of the purity of the fusion proteins showed that, after affinity chromatography, the fusion proteins were only approx. 85% of the protein present. Nevertheless, when expressed and purified in the same way, MBP alone did not hydrolyse Ap_4A and we conclude that in each case the expressed lupin protein is responsible for the observed Ap_4A hydrolase activity.

Gene number and complexity

Southern blotting of genomic DNA from lupin (Figure 4) showed that only one or two bands were detected by the Ap_4A hydrolase-specific probe within DNA cleaved with a range of endonucleases. *Eco*RI, which cleaves once within the coding region of the AH-7 cDNA sequence, produced two main bands on cleavage of lupin genomic DNA. Likewise by using *Bgl*II in combination with *Pst*I it was possible to detect the two bands expected from the *Pst*I cleavage site within the coding region. It is likely that the two bands generated by *Dra*I were due to cleavage within an intron in a single gene rather than to two variants of the gene. We therefore obtained no evidence from Southern blotting of more than one version of the Ap_4A hydrolase gene in the *L. angustifolius* genome.

Homology between mammalian and plant Ap_4A hydrolases

Lupin Ap_4A hydrolase protein sequence was used to screen the GenBank databases for related sequences. Although no sequences known to encode plant Ap_4A hydrolases were detected, putative Ap_4A hydrolase sequences were detected among the GenBank expressed sequence tags (ESTs) from *Arabidopsis thaliana* (GenBank EST accession number Z34022) and *Oryza sativa* (GenBank EST accession number C20013). The deduced protein sequence of the lupin Ap_4A hydrolase was compared with both the sequences deduced for these putative plant Ap_4A hydrolases and those of all other known asymmetric Ap_4A hydrolases (Figure 5). There was high similarity between the lupin sequence and that deduced for each of the putative enzymes from *Arabidopsis* and rice. The sequences shown in Figure 5 were also analysed by the GAP program. There was 87% similarity and 72% identity across a 99-amino acid overlap between lupin and rice and 83% similarity and 72% identity across a 99-amino

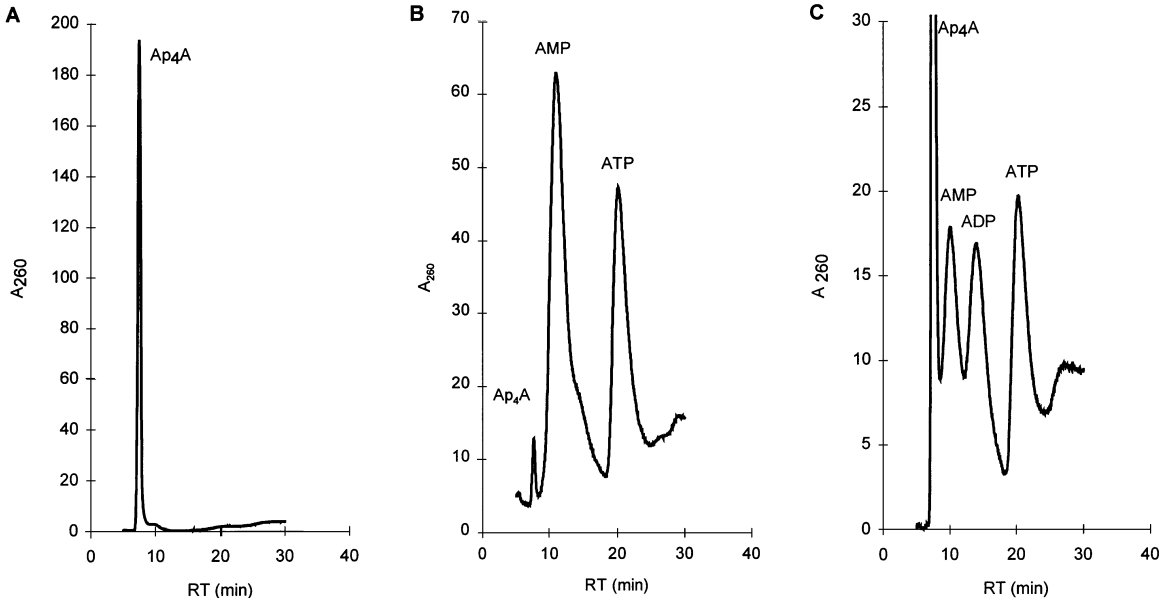


Figure 3 Ap_4A hydrolase activity of expressed MBP–Ap_4AH fusion protein

Expressed protein from cells containing pMALAH-7 was purified by affinity chromatography on amylose resin and incubated with Ap_4A as described in the Experimental section. Reactions were terminated at zero time (**A**) or 10 min (**B**) and the products separated by HPLC. Retention times (RT) of standards are given in (**C**).

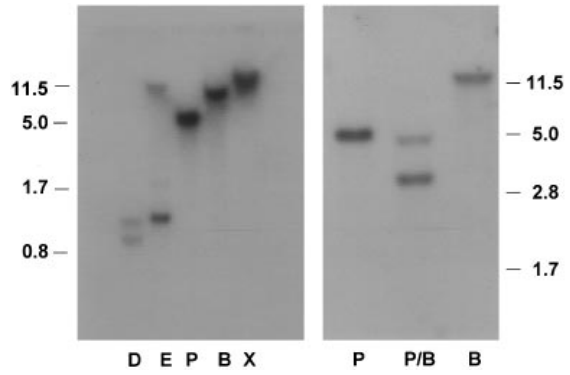


Figure 4 Southern-blot analysis of Ap_4A hydrolase genes in the *L. angustifolius* genome

Genomic DNA from *L. angustifolius* cv. Unicrop cut with *Eco*RI (E), *Dra*I (D), *Pst*I (P), *Bgl*II (B) and *Xba*I (X) and probed with radiolabelled insert from pAH-7 DNA. Sizes in kb are indicated on both sides.

acid overlap between lupin and *Arabidopsis*. The high degree of identity between the amino acid sequences of these and the lupin Ap_4A hydrolase implies the existence in *Arabidopsis* and rice of equivalent Ap_4A hydrolases. Whereas proteins with Ap_4A hydrolase activity are yet to be identified in either *Arabidopsis* or rice, asymmetric Ap_4A hydrolase has been purified from tomato [21] and it too aligns well within the limited region of protein sequence known for the tomato enzyme (Figure 5).

In contrast with the putative enzymes, identity between lupin and the known mammalian Ap_4A hydrolases was much more limited. Even though the sequences deduced from the cDNAs encoding Ap_4A hydrolases from pigs and from humans have a high degree of identity (88%) with each other [24,25], identity between the lupin, pig and human sequences was quite low (< 25%) (Figure 5). Identity across all the deduced sequences shown in Figure 5 was restricted very much to the region corresponding to the motif conserved in the Mut T family of proteins [34] and already known to be conserved in pig and human Ap_4A hydrolases [24,25]. Between residues 76 and 99 of the lupin protein, we detected conserved amino acids equivalent to six of the seven amino acids that have previously been shown by NMR structure studies to form a glutamate-rich motif in the Mut T protein from *E. coli* [35–37] and this conservation was detected even within the short section of peptide sequence available for the tomato enzyme. Outside of this conserved region, lupin and the putative *Arabidopsis* and rice sequences remained closely related further towards the C-terminus of the proteins. Human and pig sequences were similarly closely related to each other in the regions towards the C-terminus, but the similarity between the plant and mammalian sequences in all but the Mut T region was quite low (Figure 5). The differences in sequence outside of the Mut T region are surprising in view of the similarity in catalytic characteristics of these plant and mammalian enzymes. The lupin, human and pig hydrolases all show a narrow substrate range, preferentially hydrolysing Ap_4A as the major substrate and showing low rates of hydrolysis with Ap_3A and Ap_5A. By contrast, other Mut T-containing proteins show quite different substrate specificities. The original bacterial Mut T protein for example is a pyrophosphatase which has 8-oxo-dGTP as the preferred substrate [38]. In spite of the above sequence differences, mammalian and lupin enzymes also show similar sensitivities to F^- ions. Both the mammalian and plant enzymes are inhibited by F^- anion, a common inhibitor of ATPases and phosphatases [17,21]. Studies on the lupin enzyme showed this to be non-competitive inhibition and to be achieved at relatively low concentrations of F^- [17]. Whereas IC_{50} values for F^- reported for some enzymes metabolizing nucleotides are in the millimolar range [39], the IC_{50} values for the lupin and tomato Ap_4A hydrolases are 2–3 μM and 6 μM respectively [17,21] and that for the mammalian hydrolase 20 μM. As shown in Table 1 the expressed fusion protein produced from clone AH-7 was also inhibited by 3 μM F^- with an IC_{50} value of 6 μM,

```
                1            15 16           30 31           45 46           59 60           74
1 lupin         MALCRLVQSPPTNFH FRKYPSKFLKFSSLS LAFRYCHSSMDSPPE GYRRNVGICLMNN-D KKIFAASRLDIPDAW  74
2 rice          --------------- --------------- --------------- ----------LAXPSL TKIFTASRIDIANTW  21
3 arabidopsis   --------------- --------------- --------------- ------------P-C RKIFTASKIHLPDTW  17
4 tomato        --------------- --------------- --------------- ---------------- ----------------
5 human         --------------- --------------- MALRACG------LI IFRRCLIPKVDNN-A IEFLLLQASDGIHHW  38
6 pig           --------------- --------------- MALRACG------LI IFRRRLIPKVDNT-A IEFLLLQASNGIHHW  38

                75           89 90          104 105         119 120         133 134         148
1 lupin         QMPQGGIDEGEDPRN AAIRELREETGVTSA EVIAEVPYWLTYDFP PKVREKLNIQWG-SD WKGQAQKWFLFKFTG 148
2 rice          QMPQGGIDAGEDPRE AAFRELREETGVTSA EMVAEVPVWLTYDFP VDVKEKLNARWGGTN WKGQAQKWFLFRFTG  96
3 arabidopsis   QMPQGGADEGEDLRN AAFRELREETGVTSA EFIAEIPNWLTYDFP REVKDKLNRKLR-TS YKGQAQKWFLFKFTG  91
4 tomato        -MPQGGVDDNEDPTN AAKR------------ ---------------- ---------------- ----------  18
5 human         TPPKGHVEPGEDDLE TALRETQEEAGIEAG QLTIIEGFKRELNYV ARNKPKIVIYWL-AE VKDYDVEIRLSHEHQ 112
6 pig           TPPKGHVEPGESDLQ TALRETQEEAGIDAG QLTIIEGFRKELNYV AWEKPKIVIYWL-AE VKDYDVEVRLSREHQ 112

 MutT  Consensus  G.....E.... ...RE..EE

                149         163 164         178 179         193 194         210
1 lupin         QDQEINLLGDGSEKP EFGEWSWVTPEQLID LIVEFKKPVYKEVLS VFAPHL    199
2 rice          EGR-------------- ---------------- ---------------- ------     99
3 arabidopsis   KEEEINLL------- ---------------- ---------------- ------     99
4 tomato        ---------------- ---------------- ---------------- ------     18
5 human         AYRWLG-LEEACQLA QFKEMKAALQEGHQF LCSIEA---------- ------    147
6 pig           AYRWLG-LDEACQLA QFKDMKAVLQEGHQF LCSMAA---------- ------    147
```

Figure 5 Asymmetric Ap_4A hydrolase: derived amino acid sequences

Sequences for putative Ap_4A hydrolases derived from cDNAs from human [24], porcine [25], lupin (Figure 1), *Arabidopsis* (GenBank EST accession number Z34022) and rice (GenBank EST accession number C20013) and from peptide sequence from tomato [21]. The Mut T consensus shows the residues conserved in Mut T-like proteins [34,35]. Boxed areas indicate conservation across at least three sequences.

Table 1. Inhibition of Ap_4A hydrolase activity by F^- ions

Rate-limiting amounts of MBP–Ap_4AH fusion protein were preincubated for 5 min in a standard assay mixture containing NaF at the concentrations shown before substrate was added and incubation continued. Reaction rates, which were linear throughout the period of the experiment, were determined by analysis of Ap_4A, AMP and ATP levels by HPLC in aliquots removed at time intervals up to 10 min.

F^- concentration (μM)	AMP or ATP production (nmol/min per μg of fusion protein)	Inhibition (%)
0	3.3	–
3	2.5	25
6	1.6	51
100	0.1	97

further evidence that the cloned and expressed enzyme had the expected properties. Such sensitivity contrasts with that for the alternative types of enzymes. F^- does not inhibit the *S. pombe* asymmetric Ap_4A hydrolase, the Ap_3A hydrolase or the phosphodiesterases from lupins, and the symmetrical Ap_4A hydrolase from *E. coli* is unaffected by even 20 mM F^- [17]. The grouping by F^- sensitivity is consistent with the favoured substrate profiles of these enzymes both with naturally occurring Ap_nA substrates and with phosphonate derivatives [21].

Contrasting with the close relationship between the enzymes from the higher eukaryotes was the observation that identity between the Ap_4A hydrolase from *S. pombe* and any of the above plant or mammalian enzymes was low throughout the whole length of the proteins (results not shown). Whereas *S. pombe* Ap_4A hydrolase has been shown to have a very high similarity ($> 60\,\%$) to the human FHIT protein sequence [26], we were unable to detect significant sequence similarity between FHIT protein and either the lupin or the human Ap_4A hydrolase sequences, and suggest therefore that the FHIT-like domains are limited to the *S. pombe*-type of hydrolase. In this context it has recently been shown that the human FHIT protein has Ap_3A hydrolase activity [27], a property it shares with the *S. pombe* type of Ap_4A hydrolase. Specific Ap_3A hydrolase distinct from the Ap_4A hydrolase has also been purified from lupin [40].

We conclude that the similarity of the catalytic properties of the human and lupin enzymes and the fact that the Mut T motifs of these enzymes are virtually the only regions conserved between the two proteins implies that the Mut T motif is likely to be closely associated with the catalytic site of these enzymes. Site-directed mutagenesis of this region of the lupin gene is in progress to define more clearly its role in the catalytic activity of this enzyme as a hydrolase of Ap_4A rather than a pyrophosphatase of 8-oxo-dGTP. We also conclude that, in contrast, the domains that provide the capacity for the Ap_3A hydrolase activities of both the tumour-suppressing human FHIT protein and the Ap_4A hydrolase from *S. pombe* are not present in this higher-plant enzyme.

This work was supported by grants from the Australian Research Council (to K.R.G.), the Underwood Fund of the Biotechnology and Biological Science Research Council, U.K. (to M.G.B., K.R.G., A.G.) and University of Melbourne Postgraduate Award (to D.M.). In part, the work was also financially supported by the State Committee for Scientific Research, KBN, within project 6 P04A 060 08 (to A.G.).

Received 29 May 1997/4 September 1997; accepted 11 September 1997

REFERENCES

1 Brevet, A., Chen, J., Lévêque, F., Plateau, P. and Blanquet, S. (1989) Proc. Natl. Acad. Sci. U.S.A. **86**, 8275–8279
2 Garrison, P. N. and Barnes, L. D. (1992) in Ap_4A and Other Dinucleoside Polyphosphates (McLennan, A. G., ed.), pp. 29–61, CRC Press, Boca Raton, FL
3 Kitzler, J. W., Farr, S. W. and Ames, B. N. (1992) in Ap_4A and Other Dinucleoside Polyphosphates (McLennan, A. G., ed.), pp. 135–149, CRC Press, Boca Raton, FL
4 Weinmann-Dorsch, C., Hedl, A., Grummt, I., Albert, W., Ferdinand, F. J., Friss, R. R., Pierron, G., Moll, W. and Grummt, F. (1984) Eur. J. Biochem. **138**, 179–185
5 Klishnin, A., Lozovzyz, N., Pintor, J., Miras-Portugal, M. T. and Krishtal, O. (1994) Neuroscience **58**, 235–236
6 Pintor, J. and Portugal, T. M. (1995) Br. J. Pharmacol. **115**, 895–902
7 Lazarowski, E. R., Watt, W. C., Stutts, M. J., Boucher, R. C. and Harden, T. K. (1995) Br. J. Pharmacol. **116**, 1619–1627
8 Gasmi, L., McLennan, A. G. and Edwards, S. W. (1996) Blood **87**, 3442–3449
9 Zamecnik, P. C., Kim, B., Gao, M. J., Taylor, G. and Blackburn, G. M. (1992) Proc. Natl. Acad. Sci. U.S.A. **89**, 2370–2373
10 Baxi, M. D. and Vishwanatha, J. K. (1995) J. Pharmacol. Toxicol. Methods **33**, 121–128
11 Barnes, L. D. and Culver, C. A. (1982) Biochemistry **21**, 6123–6128
12 Guranowski, A., Jakubowski, H. and Holler, E. (1983) J. Biol. Chem. **258**, 14784–14789
13 Plateau, P., Fromant, M, Brevet, A., Gesquière, A. and Blanquet, S. (1985) Biochemistry **24**, 914–922
14 Robinson, A. K., de la Peña, C. E. and Barnes, L. D. (1993) Biochim. Biophys. Acta **1161**, 139–148
15 McLennan, A. G., Mayers, E., Gunaratne, R. S., Prescott, M. and Powls, R. (1994) Biochem. Soc. Trans. **22**, 231S
16 Ogilvie, A. and Antl, W. (1983) J. Biol. Chem. **258**, 4105–4109
17 Guranowski, A. (1990) FEBS Lett. **262**, 205–208
18 Lazewska, D., Starzynska, E. and Guranowski, A. (1993) Protein Expression Purif. **4**, 45–51
19 Hankin, S., Matthew, N., Thorne, H. and McLennan, A. G. (1995) Int. J. Biochem. Cell. Biol. **27**, 201–206
20 Jakubowski, H. and Guranowksi, A. (1983) J. Biol. Chem. **258**, 9982–9989
21 Feussner, K., Guranowski, A., Kostka, S. and Wasternack, C. (2996) Z. Naturforsch. **51c**, 477–486
22 Guranowski, A., Brown, P., Ashton, P. A. and Blackburn, G. M. (1994) Biochemistry **33**, 235–240
23 Huang, Y., Garrison, P. N. and Barnes, L. D. (1995) Biochem. J. **312**, 925–932
24 Thorne, N. M. H., Hankin, S., Wilkinson, M. C., Nunez, C., Barraclough, R. and McLennan, A. G. (1995) Biochem. J. **311**, 717–721
25 Hankin, S., Wintero, A. K. and McLennan, A. G. (1996) Biochem. Soc. Trans. **24**, 418S
26 Ohta, M., Hiroshi, I., Cotticelli, M. G., Kastury, K., Baffa, R., Palazzo, J., Siprashvili, Z., Mori, M., McCue, P., Druck, T., Croce, C. M. and Huebner, K. (1996) Cell **84**, 587–597
27 Barnes, L. D., Garrison, P. N., Siprashvili, Z., Guranowksi, A., Robinson, A. K., Ingram, S. W., Croce, C. M., Ohta, M. and Heubner, K. (1996) Biochemistry **35**, 11529–11535
28 Laemmli, U. K. (1970) Nature (London) **227**, 680–685
29 Redinbaugh, M. and Turley, R. (1986) Anal. Biochem. **153**, 267–271
30 Kozak, M. (1991) J. Cell Biol. **115**, 887–903
31 Joshi, C. P. (1987) Nucleic Acids Res. **15**, 6643–6653
32 Prescott, M., Milne, A. D. and McLennan, A. G. (1989) Biochem. J. **259**, 832–838
33 Moreno, A., Lobaton, C. D., Gunther Sillero, M. A. and Sillero, A. (1982) Int. J. Biochem. **14**, 629–634
34 Koonin, E. V. (1993) Nucleic Acids Res. **21**, 4847
35 Frick, D. N., Weber, D. J., Abeygunawardana, C., Gittis, A. G., Bessman, M. J. and Mildvan, A. S. (1995) Biochemistry **34**, 5577–5586
36 Abeygunawardana, C., Weber, D. J., Gittis, A. G., Frick, D. N., Lin, J., Miller, A.-F., Bessman, M. J. and Mildvan, A. S. (1995) Biochemistry **34**, 14997–15005
37 Lin, J., Abeygunawardana, C., Frick, D. N., Bessman, M. J. and Mildvan, A. S. (1995) Biochemistry **35**, 6715–6726
38 Maki, H. and Sekiguchi, M. (1992) Nature (London) **355**, 273–275
39 Kornberg, A. (1955) Methods Enzymol. **2**, 655–659
40 Guranowski, A., Starzynska, E., Bojarska, E., Stepinski, J. and Darzynkiewicz, E. (1996) Protein Expr. Purif. **8**, 416–422

The nucleotide sequence of the cDNA of lupin Ap_4A hydrolase (Figure 1) has been deposited with the EMBL/GenBank Nucleotide Sequence DataBases under the accession number U89841.