

# Plant Gene Register PGR98-166

**Brock Weers and Robert Thornburg** (1998) Characterization of the cDNA and Gene for the *Arabidopsis thaliana* Adenylate Kinase (Accession No. AF082882). (PGR98-166) Plant Physiol. **118**: 711

---

## Characterization of the cDNA and Gene for the *Arabidopsis thaliana* Adenylate Kinase (Accession No. [AF082882](#))

Brock Weers and Robert Thornburg (\*)

Department of Biochemistry and Biophysics, Iowa State University, Ames, Iowa 50011

(\*) Corresponding author: Robert Thornburg

Fax: 1-515-294-0453

Email: thorn@iastate.edu

Journal Paper No. J-18038 of the Iowa Agriculture and Home Economics Experiment Station, Ames, Iowa  
Project No. 3202.

Keywords: adenylate kinase, AMP kinase, nucleotide metabolism, phosphoryltransferase

Nucleotide monophosphates are a common intermediate in cellular biochemistry. The re-utilization of polynucleotides within the cell produces high levels of nucleotide monophosphates. In addition, many energy-requiring biosynthetic reactions hydrolyze ATP into AMP plus pyrophosphate. For these monophosphates to be re-utilized within the cell, they must be converted back into the triphosphate form. Various nucleotide monophosphate kinases perform the first step of this phosphorylation, converting the monophosphate form into the diphosphate form at the expense of one mole of ATP. In the case of adenylate kinase, the reaction is given as  $AMP + ATP \rightarrow 2 ADP$ .

In addition to the salvage of preformed nucleotides, adenylate kinase plays an important role in the light reactions of photosynthesis in mesophyll chloroplasts. In a cyclical process ATP hydrolysis is linked to the conversion of pyruvic acid into PEP, which is the substrate for the cytoplasmic PEP carboxylase. The coproducts of this reaction are AMP and pyrophosphate. Adenylate kinase then converts AMP into ADP, which is the substrate for the light phase of photosynthesis. The result of this cycle is a net accumulation of PEP for the capture of carbon dioxide.

Adenylate kinase also plays an important role in the regulation of glycolysis and gluconeogenesis. Glycolysis and gluconeogenesis are opposing pathways shuttling carbon into net accumulation of glucose or glycolytic catabolism. The pathway is controlled by two enzymes, phosphofructokinase (PFK) and fructose-1,6-bisphosphatase (FBPase). When FBPase is active the equilibrium shifts towards gluconeogenesis, and when PFK is active, the equilibrium shifts towards glycolysis. Both of these enzymes are allosterically regulated by AMP and fructose-2,6-bisphosphate. These compounds inhibit FBPase and activate PFK. Because adenylate kinase controls the conversion of AMP into ADP and modifies the ratios of AMP to ATP, adenylate kinase activity therefore plays an important role in the regulation of glycolytic to gluconeogenic activity.

Because of our interests in the nucleotide monophosphate kinases of *Arabidopsis* (Zhou *et al.*, 1998; Zhou and Thornburg, 1998), we searched the *Arabidopsis* EST libraries for other clones with high similarity to the *Arabidopsis* UMP/CMP kinase. In this screen we identified the clone 170P10T7. This clone was obtained from the ABRC (ABRC, 1995), sequenced, characterized and deposited in GenBank as Accession Number AF092992.

## Clone Identification

The adenylate kinase cDNA was characterized in this study. It was identified as an expressed sequence tag (170P10T7) in the ABRC collection (ABRC, 1995) due to its similarity with the *Arabidopsis thaliana* UMP kinase (Zhou *et al.*, 1998). After the clone had been sequenced, Blast searches (Altschul *et al.*, 1990) revealed that the highest identity was found with a series of proteins identified as adenylate kinases. On the basis of this high degree of identity, this clone was considered to be the *Arabidopsis thaliana* adenylate kinase.

Blast searches also identified an *Arabidopsis thaliana* chromosomal P1 clone (MLE2) sequenced by the Kazusa Genome Project that had 100% identity with this cDNA. Therefore, we also included this sequence in our analysis of the *Arabidopsis* ADK1 gene.

## cDNA Sequencing

DNA sequence reactions were performed using the Applied Biosystems Prism Dye-deoxy Cycle Sequencing Kit. The reactions were run on an Applied Biosystems Prism 377 DNA sequencer (Perkin-Elmer Corp.). The sequencing was initiated from known vector sequences. On the basis of these runs, primers specific to each GLP sequence were constructed to extend the DNA sequence. DNA sequences were performed in duplicate or triplicate for each run. Each strand was completely sequenced, including sequencing through all restriction sites, and the entire sequence of the cDNA was confirmed on the opposite strand.

## Characteristics of the Adenylate kinase cDNA

The cDNA was found to be 969 nucleotides in length and encoded the full length adenylate kinase mRNA. The initiation ATG start codon is located at position [32..34]. The sequence in this region is in good agreement with the consensus eukaryotic translation start sequence (Kozak, 1986), and this mRNA is expected to be efficiently translated. The poly A site of this cloned mRNA was found at nucleotide 950. Upstream of this site, we could not identify a classical poly adenylation signal; however, one potential sequence, AATATG, was identified at 925 to 930 of this cDNA, 20 nucleotides upstream from the poly A site.

## Characteristics of the Deduced Amino Acid Sequence

The coding region of the *Arabidopsis* adenylate kinase contains an open reading frame of 246 codons. The protein sequence deduced from this open reading frame is a 26,934 Da polypeptide with a pI of 8.24. When this protein was analyzed for intracellular location using the PSORT tool (Nakai and Kanehisa, 1992), it was predicted to be cytoplasmic, as was also found for the *Arabidopsis* UMP/CMP kinase (Zhou *et al.*, 1998).

Adenylate kinases have been identified from a wide variety of sources including 27 species of archaeobacteria and eubacteria, 3 species of single cell eukaryotes, 2 species of yeast, 3 species of plants, 3 species of mammals, and birds. Thus, a wide variety of sequences are available for analysis. Comparison of these sequences with the GCG tool "PileUp" (GCG, v. 9.0) revealed that these adenylate kinase sequences consist of several distinct subfamilies. This is typical for multigene families.

The *Arabidopsis thaliana* adenylate kinase is most closely related to the adenylate kinase cDNAs (ADKa and ADKb) isolated from rice (Kawai, *et al.*, 1992), sharing 81% amino acid identity with each of these cDNAs. These three sequences form a subfamily that is distinct from other adenylate kinases. This subfamily of sequences is most similar to the yeast adenylate kinase genes and to mammalian ADK2 genes. The mammalian ADK1 genes form a separate subfamily and show less similarity with the *Arabidopsis* gene. The adenylate kinase cDNA from apricots, *Prunus armeniaca*, (Mbeguie, *et al.*, 1997) belongs to the ADK1 subfamily and shows only 21% identity with the *Arabidopsis* cDNA described in this report; however, this sequence shares 57% identity with the *Arabidopsis* UMP/CMP kinase (Zhou *et al.*, 1998).

We also compared the *Arabidopsis* adenylate kinase cDNA sequence to the *Arabidopsis* UMP/CMP kinase

sequence. This analysis showed that the Arabidopsis adenylate kinase has only 29% identity with the Arabidopsis UMP/CMP kinase. In addition, the Arabidopsis adenylate kinase has a 25 amino acid insertion that is not found in the eukaryotic UMP/CMP kinases.

### Characteristics of the Adenylate kinase Gene

The chromosomal P1 clone (MLE2) was identified in BLAST searches using the adenylate kinase cDNA sequence. This clone maps to Arabidopsis chromosome V at approximately 24.5 cM. The cDNA aligns with genomic sequences located from 4747 to 6501 on the negative strand of the MLE2 P1 clone. This alignment reveals the presence of 5 introns within the cDNA. With the exception of intron 1, which is 421 bp long, the remainder of these introns are short (95, 105, 91, and 91 bp, respectively) and polydisperse throughout the cDNA. Interestingly, the location of one of the introns is exactly conserved between the Arabidopsis and human genes. The Arabidopsis ADK1 intron I interrupts the codon for Glycine 40 (G<sup>A</sup>GA). After alignment for sequence conservation, we found that the Human ADK1 intron 2 also interrupts the conserved Glycine codon at the identical location (G<sup>A</sup>GT). The location of this intron is in the first glycine residue of the phosphate-binding loop which is conserved among known adenylate kinases (Reinstein, *et al.*, 1990). Whether the intron is necessary for some unknown physiological process is not clear; however, the conservation of this intron from plants to mammals in a biochemically important structural feature of the protein is very intriguing.

Flanking genes identified for this P1 clone include a receptor protein kinase upstream from the adenylate kinase gene and a hypothetical 69.1 kD protein 2.75 kB downstream from the adenylate kinase gene. The upstream receptor protein kinase gene is expressed from the same strand and is extremely close to the adenylate kinase gene. Only 323 nucleotides separate the stop codon of the receptor protein kinase from the start of the adenylate kinase cDNA. The receptor protein kinase encodes a 674 amino acid protein that identifies an expressed sequence tag, indicating that this sequence is expressed. The proximity of this flanking gene leaves very little room for the 3' end of the receptor protein kinase gene and for the promoter of the adenylate kinase gene.

Nevertheless, we searched for potential regulatory elements within the flanking regions of the adenylate kinase gene. Within the 5' flanking region, we identified the TATA box at [-25..-16] relative to the start of the adenylate kinase cDNA. Four putative light responsive elements (Terzaghi, *et al.*, 1995) were identified within the short promoter. The physiological importance of any light responsive elements are as yet unknown. However, because ATP biosynthesis is up-regulated in the light reactions of photosynthesis, one might expect that an adenylate kinase gene would be light regulated.

Other potential regulatory elements were not identified. We were unable to identify a consensus TATA box. The best sequence fit for the TATA box was "gaataaaac" which was located only 16 nucleotides upstream from the start of the adenylate kinase cDNA. No other sequence within the intergenic region showed a good fit for a putative TATA box. The proximal promoter is very GC rich. The proximal 100 nucleotides are 41% G+C.

Because of the short sequence between these two genes, we postulated that the adenylate kinase promoter might overlap with the receptor protein kinase structural gene. Therefore, we also analyzed sequences within the receptor protein kinase structural gene for potential regulatory sequences. For this analysis, we examined the proximal 2.9 kb of the sequence, which contains the full length of the receptor protein kinase structural gene plus the short intergenic sequence analyzed previously as the "short" promoter. The sequence of this longer fragment identified numerous additional putative light regulated elements scattered throughout this longer fragment. In addition, a series of four putative myb binding sites, CNGTTR (Urao, *et al.*, 1993) were identified within the receptor protein kinase structural gene between -1400 and -231 relative to the start of the adenylate kinase cDNA.

---

#### Table 1. Characteristics of the GMP1 cDNA and gene

Organism:

*Arabidopsis thaliana*

**Nomenclature:**

cDNA ID = 170P10T7 (GenBank Accession # [AF082882](#)) Gene Name ADK1 Genomic Clone ID = P1 clone MLE2 (GenBank Accession # [AB007649](#)) Chromosomal Location = Chromosome V near 24.5 cM

**Gene Product:**

Adenylate kinase, AMP kinase, EC [2.7.4.3](#) identification by sequence comparison to known genes

**Characteristics of the ADK1 cDNA:**

(Positions are numbered relative to the cDNA sequence) The sequenced DNA fragment of the ADK1 cDNA was a 969 bp fragment containing the entire ADK1 cDNA.

**5' untranslated region**

31 nucleotides of the 5' UTR were identified [1..31]

**Coding region**

738 nucleotides [32..769]

Terminates with an opal codon (TGA) [770..772]

**3' untranslated region**

177 nucleotides [773..950]

Poly A signal [925..930]

Poly A addition site at 950 of this sequence

**Characteristics of the ADK1 gene:**

(Positions are numbered relative to the MLE2 P1 clone and where appropriate to the location of the cDNA start site. Note: the ADK1 gene is on the negative strand of the MLE2 P1 clone, therefore the numbering of the genomic sequences proceeds from larger to smaller numbers)

**5' flanking sequences**

Two potential promoter fragments were analyzed. First the 323 bp short promoter that lies between the receptor protein kinase and the ADK1 gene. Second, 2920 nucleotides of DNA directly upstream of the ADK1 gene were analyzed as a long promoter.

**Short Promoter**

323 nt of the promoter adjacent to the cDNA were analyzed no consensus plant TATA box was identified. Four GT1 consensus light-regulated sequences were identified.

**Long Promoter**

2920 nt of the sequences adjacent to the cDNA were analyzed. No consensus plant TATA box was identified. Many GT1 consensus light-regulated sequences were identified. Four putative myb binding sites were identified.

Coding region structure (identified by comparison with the ADK1 cDNA)

5' UTR [6501..6471]  
Exon I [6501..6353]  
Intron I [6352..5932] 421 nt  
Exon 2 [5931..5785]  
Intron 2 [5784..5689] 95 nt  
Exon 3 [5688..5569]  
Intron 3 [5568..5464] 105 nt  
Exon 4 [5463..5296]  
Intron 4 [5295..5205] 91 nt  
Exon 5 [5204..5115]  
Intron 5 [5114..5024] 91 nt  
Exon 6 [5023..4928]

Stop codon [4927..4925]  
3' UTR [4924..4747]

3' flanking region  
No significant repeats, palindromes, or hairpin loops were  
detected within the 3' flanking region

---

## Literature Cited

- ABRC:** (1995) Library information for ESTs. in "Arabidopsis Biological Resource Center: Seed and DNA Stock List". The Ohio State University, 1735 Neil Avenue, 309 Botany & Zoology Bldg. Columbus, OH 43210 USA pp 239
- Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ:** (1990) Basic local alignment search tool. *J. Mol. Biol.* **215**: 403-410
- GCG: (Version 9.0) Wisconsin Package, Genetics Computer Group, Madison, WI.
- Kawai M, Kidou S, Kato A, Uchimiya H** (1992) Molecular characterization of cDNA encoding for adenylate kinase of rice (*Oryza sativa* L.). *Plant J.* **2**:845-854
- Kozak M** (1986) Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. *Cell* **44**:283-292
- Mbeguie D, Mbeguie A, Gomez R-M, Fils-Lycaon B** (1997) Molecular cloning and nucleotide sequence of an adenylate kinase from apricot. GenBank Accession U82330.
- Nakai K, Kanehisa M:** (1992) A knowledge base for predicting protein localization sites in eukaryotic cells. *Genomics* **14**: 897-911
- Reinstein J, Schlichting I, Wittinghofer A** (1990) Structurally and catalytically important residues in the phosphate-binding loop of adenylate kinase of *Escherichia coli*. *Biochemistry* **29**:7451-7459
- Terzaghi W, Cashmore A** (1995) Light-regulated transcription. *Annu Rev Plant Physiol Plant Mol Biol* **46**:445-474
- Urao T, Yamaguchi-Shinozaki K, Urao S, Shinozaki K** (1993) An Arabidopsis myb homolog is induced by dehydration stress and its gene product binds to the conserved MYB recognition sequence. *Plant Cell* **5**:1529-1539
- Zhou L, Lacroute F, Thornburg R** (1998) Cloning, expression in *Escherichia coli* and characterization of Arabidopsis thaliana UMP/CMP kinase. *Plant Physiol* **117**:245-253
- Zhou L, Thornburg RW** (1998) Site Specific mutagenesis of conserved residues in the phosphate binding loop of the Arabidopsis UMP/CMP kinase alter ATP and UMP binding. *Arch. Biochem. Biophys.* (in press)



[Return to Plant Gene Register Index](#)



[Return to Plant Physiology ONLINE](#)