Plant Physiol. (1991) 97, 1599–1601
0032-0889/91/97/1599/03/$01.00/0

Received for publication July 12, 1991
Accepted August 22, 1991

Plant Gene Register

Nucleotide Sequence of a cDNA Clone Encoding a β-Amylase from *Arabidopsis thaliana*[1]

Jonathan D. Monroe, Mark D. Salminen[2], and Jack Preiss*

Biochemistry Department, Michigan State University, East Lansing, Michigan 48824

β-Amylase (EC 3.2.1.2) hydrolyzes α-1,4-glucosidic linkages from the nonreducing ends of starch or glycogen, releasing maltose with the β-anameric configuration. This enzyme occurs in plants and in certain bacteria; however, its physiological function in some plants is not clear because the enzyme and its substrates are spatially separated. In the leaves of a variety of plants including *Arabidopsis thaliana* (L.) Heynh. (3), β-amylase is located outside the chloroplast (5, 7). In pea and wheat, β-amylase appears to be confined to the vacuole (11). The *A. thaliana* β-amylase is also located in the vacuole as determined by nonaqueous fractionation (our unpublished results). However, the only known substrates for β-amylase, soluble and insoluble starch, are confined to the chloroplast (8).

The *A. thaliana* β-amylase is the most abundant of the three major leaf amylases, making up about 80% of the total crude amylolytic activity (3). In the course of studying starch metabolism in *A. thaliana*, leaves of several starchless and starch-overproducing mutants were found to contain 10- to 40-fold more β-amylase activity than wild-type leaves when the plants were grown under a 12-h/12-h light/dark cycle (1). Under continuous light, mutant and wild-type plants were indistinguishable. In addition, β-amylase activity was increased in both the wild type and mutants by growing plants under a higher light intensity. Using polyclonal antibodies raised to the purified *A. thaliana* β-amylase, we determined that the β-amylase protein accumulated in the mutants and in plants grown under high light intensity (5).

To investigate how β-amylase expression is regulated in *A. thaliana*, we isolated five cDNA clones from a λ-Zap expression library using immunoaffinity-purified anti-β-amylase IgG. Restriction mapping and partial sequence analysis indicated that the coding regions of all five cDNAs were identical (data not shown). Here, we present the sequence of the longest complete cDNA which contains 2469 bases and a single open reading frame encoding a protein of 498 amino acids (Table I, Fig. 1). The mol wt of the protein from the deduced amino acid sequence is 56,069, and the apparent molecular mass of the purified protein as determined from SDS-PAGE was 55,000 D (5).

β-Amylase sequences are known from three other plants: soybean (4), sweet potato (9), and barley (2). The *A. thaliana* β-amylase sequence is 69% identical with the soybean sequence at the amino acid level. When amino acids with similar side groups are included in the comparison, the *A. thaliana* and soybean sequences are 79% similar. The *A. thaliana* sequence is 64 and 63% identical with, and 75 and 73% similar to, the sweet potato and barley sequences, respectively. The active sites of the soybean (6) and sweet potato (9) β-amylases were identified by modification of the proteins

Table I. *Characteristics of a β-Amylase cDNA from A. thaliana*

Organism:
Arabidopsis thaliana (L.) Heynh., Columbia ecotype, Brassicaceae.
Enzyme, Function:
β-(Exo)amylase EC 3.2.1.2; physiological function not known.
Source:
cDNA library constructed from polyadenylated mRNA isolated from entire plants.
Techniques:
cDNA library in λ-Zap screened with immunoaffinity-purified anti-β-amylase IgG; restriction fragment subcloning; single- and double-stranded plasmid sequencing in either pT7T3-18, 19 (Pharmacia) or pBluescript (Stratagene) and dideoxy sequencing of both strands.
Method of Identification:
Sequence identity with four peptides derived from an endoproteinase Lys-C digest of purified β-amylase and sequenced by automated Edman degradation (Applied Biosystems).
Features of cDNA Structure:
Contains 2469 nucleotides consisting of 89 nucleotides 5′-untranslated (74% AT), 1494 nucleotides open reading frame, and 886 nucleotides 3′-untranslated; no polyadenylated addition signal or tail were observed at the 3′-end of the cDNA.
Features of Protein Structure:
Open reading frame of 498 amino acids; M_r 56,069; no amino acid accounts for more than 10% of the total amino acids; no N-terminal signal sequence.
Antibodies:
Polyclonal antibodies to native β-amylase are available.
Subcellular Location:
Vacuolar (our unpublished data).
EMBL Accession No.:
M73467

[1] Supported in part by National Science Foundation grant DMB86-10319 and by U.S. Department of Agriculture/Department of Energy/National Science Foundation Plant Science Center Program 88-37271-3964.

[2] Present address: Pinckney High School, P.O. Box 9, Pinckney MI 48169.

```
   1 CTCATATTGCTTATATTTTCCTACAACACTCTCTCAACAAAAAAACAAAAAAAAAAGACGTTCTTTCTTTGTTTGTATTTAGAAGAAAAA

  91 TGGCTACCAATTACAACGAGAAGCTTCTTCTTAATTATGTTCCCGTTTACGTTATGCTTCCGTTGGGAGTTGTGAATGTGGAAAATGTAT
     M  A  T  N  Y  N  E  K  L  L  L  N  Y  V  P  V  Y  V  M  L  P  L  G  V  V  N  V  E  N  V   30

 181 TTGCGGACCCAGAAACGCTTGAAACGCAGCTTAAACGTCTCAAAGAAGAAGCTGGCGTTGATGGCGTTATGGTCGATGTTTGGTGGGGAA
     F  A  D  P  E  T  L  E  T  Q  L  K  R  L  K  E  E  A  G  V  D  G  V  M  V  D  V  W  W  G   60

 271 TCATAGAATCCAAAGGTCCCAAACAATATGATTGGACGGCCTACAAAACGCTGTTCCAGCTGATCGCACGTTTGGGACTCAAAATCCAAG
     I  I  E  S  K  G  P  K  Q  Y  D  W  T  A  Y  K  T  L  F  Q  L  I  A  R  L  G  L  K  I  Q   90

 361 CAATCATGTCTTTTCACCAATGTGGTGGAAACGTTGGCGACATCGTTACTATCCCGATCCCACAATGGGTTCGCGATGTCGGTGACAATG
     A  I  M  S  F  H  Q  C  G  G  N  V  G  D  I  V  T  I  P  I  P  Q  W  V  R  D  V  G  D  N  120

 451 ATCCCGATATCTACTACACTAACCGTAAAGGAACTAGAGACATCGAGTATCTCTCAATCGGTGTTGATAATCTTCCCCTATTTGCCGGAA
     D  P  D  I  Y  Y  T  N  R  K  G  T  R  D  I  E  Y  L  S  I  G  V  D  N  L  P  L  F  A  G  150

 541 GAACCGCTGTTCAGTTGTACAGTGATTACATGAGTAGCTTCAAAGAAAACATGGCGGATTTGATAGAAGCTGGGGTGATTGTTGACATCG
     R  T  A  V  Q  L  Y  S  D  Y  M  S  S  F  K  E  N  M  A  D  L  I  E  A  G  V  I  V  D  I  180

 630 AAGTCGGACTTGGCCCGGCCGGTGAACTACGTTATCCTTCTTACCCCCAAAGCCAAGGTTGGGTGTTTC.CGGCATCGGAGAATTCCAGT
     E  V  G  L  G  P  A  G  E  L  R  Y  P  S  Y  P  Q  S  Q  G  W  V  F  P  G  I  G  E  F  Q  210

 721 GTTATGACAAGTACTTGAAGAAAGATTTCAAGGAAGCGGCGGCGAAAGCAGGGCACCCTGAGTGGGACTTGCCAGAGGACGCCGGAGAAT
     C  Y  D  K  Y  L  K  K  D  F  K  E  A  A  A  K  A  G  H  P  E  W  D  L  P  E  D  A  G  E  240

 811 ACAATGACAAGCCGGAGGAGACTGGATTTTTCAAGAAGGATGGGACTTATGTCTCGGAGAAGGGGAAGTTTTTCATGACATGGTACTCGA
     Y  N  D  K  P  E  E  T  G  F  F  K  K  D  G  T  Y  V  S  E  K  G  K  F  F  M  T  W  Y  S  270

 901 ACAAACTAATTTTTCATGGAGATCAGATCTTAGGAGAAGCCAACAAGATCTTTGCTGGACTTAAAGTTAACTTGGCTGCCAAGGTTTCTG
     N  K  L  I  F  H  G  D  Q  I  L  G  E  A  N  K  I  F  A  G  L  K  V  N  L  A  A  K  V  S  300

 991 GGATTCACTGGTTGTACAACCACCACAGCCACGCTGCGGAGTTGACTGCTGGATATTACAACCTTTTCAAGAGAGATGGTTACCGTCCGA
     G  I  H  W  L  Y  N  H  H  S  H  A  A  E  L  T  A  G  Y  Y  N  L  F  K  R  D  G  Y  R  P  330

1081 TCGCCCGGATGCTCTCAAAACACTACGGCATTCTCAACTTCACTTGCCTTGAGATGAAAGATACCGACAATACAGCTGAAGCCCTAAGTG
     I  A  R  M  L  S  K  H  Y  G  I  L  N  F  T  C  L  E  M  K  D  T  D  N  T  A  E  A  L  S  360

1171 CTCCTCAAGAACTTGTTCAAGAGGTACTGAGCAAGGCATGGAAAGAAGGTATAGAAGTTGCGGGTGAGAACGCATTAGAGACCTATGGAG
     A  P  Q  E  L  V  Q  E  V  L  S  K  A  W  K  E  G  I  E  V  A  G  E  N  A  L  E  T  Y  G  390

1261 CCAAAGGTTACAATCAGATTCTTCTTAACGCGAGGCCTAACGGGGTTAACCCAAACGGTAAGCCGAAGCTTAGAATGTACGGATTTACTT
     A  K  G  Y  N  Q  I  L  L  N  A  R  P  N  G  V  N  P  N  G  K  P  K  L  R  M  Y  G  F  T  420

1351 ACCTTCGGTTATCCGATACTGTCTTTCAAGAAAACAACTTTGAGCTGTTTAAGAAGTTGGTGAGGAAAATGCACGCTGATCAAGATTATT
     Y  L  R  L  S  D  T  V  F  Q  E  N  N  F  E  L  F  K  K  L  V  R  K  M  H  A  D  Q  D  Y  450

1441 GTGGAGACGCAGCGAAGTACGGGCATGAGATTGTGCCGTTGAAAACGTCGAATTCGCAGCTGACGCTGGAGGATATCGCCGACGCGGCTC
     C  G  D  A  A  K  Y  G  H  E  I  V  P  L  K  T  S  N  S  Q  L  T  L  E  D  I  A  D  A  A  480

1531 AGCCAAGTGGAGCATTTAAGTGGGACTCTGAAACCGATTTGAAGGTCGACGGTTAGTTATATTTTATATGATCGATCGTCTAATCACAAA
     Q  P  S  G  A  F  K  W  D  S  E  T  D  L  K  V  D  G  *                                   498

1621 GAAAGTAAGGGAAAGAATCAAGTATTCGTTGATGTTGTTGTTTTGAACAGGCAAACTAAGGCGGTCTTGGCCTTCAATTAGTTTTGCAGA

1711 TTTGATGATATCTCCAGTCTTGATCTGCCCGAGAATATCTTTTCCTGCTATTGTGTATCCGAAGACCGAGAACTGTCCTTCGTCAAAGGA

1801 TAATCCTCCTAAGCCAGAGTTTCTTTTATCGTATAGGTAGAAGAAGAATTGGTAAGGAGAAGAGTATTCTTCTGAGTTTTCGCTGTGCGC

1891 CATTGCTACAGCTCCATAAACCGATAAAGGAAGCACTGGTAACTCCCATCCTGGACACTCAATGGCGTTCTGTAGAGAGGTTCGAATTGT

1981 CCAGAAGGCATTACTTCCAATGGAACACTAACACTCTCCACCTTACCACTCCCCATCCTCTGTGATAACAGCCTGGTTTACAGTATTGAG

2071 TTTGGCTCCATCATATGCACCGCTAGTTACCAGTTTTGCAAAATTCCCTGCGGTTAGAGGTGCTGA. TATCCATCTATAACGATCTGTAC

2161 TGTAGCACTCTTTCTTTGATCACCTCCAGCTTCAGCTGAAAAAGTCGAACCATCAGCTTTCTCAATGGTGATTTCCACAGTTCCTCTTCC

2251 TGCTAACCTGGGATAGTTTAAGTACTGCTGAGGAAGTAGAAATGACAACCCAGATGCCTGTAATAACTCCAGATCCGCAACAGTATCTAG

2341 TGATGCTGCCAGCCAAAGGGACACTTTATCAGGATCTTGTTTCTTAATAGACGTTATAAGTGCTTGTAGTCCACCCTTTCCATCAATCAG

2431 AGTTGTATACAGTTCTGAGCCCTTGTCTTTTAAGTCTAC
```

Figure 1. Nucleotide sequence and deduced amino acid sequence of a cDNA from *A. thaliana* encoding a β-amylase. The open reading frame extends from nucleotides 90 to 1583 and encodes a polypeptide of 498 amino acids. The stop codon TAG is indicated by an asterisk. The locations of four peptide sequences determined by Edman degradation from purified β-amylase are underlined.

with the substrate analog 2,3-epoxypropyl α-D-glucopyranoside. It was determined that a conserved Glu-186 in the soybean sequence and Glu-187 in the sweet potato sequence were labeled. In the corresponding region of the *A. thaliana* sequence, Glu-189 is conserved (Fig. 1), suggesting that it is also located in the active site.

Even though the *A. thaliana* β-amylase is located in the vacuole (J.D. Monroe, J. Preiss, unpublished data), the cDNA does not appear to encode an N-terminal signal sequence as defined by von Hiejne (10). This is supported by good sequence identity between the extreme N terminus of the deduced *A. thaliana* β-amylase and the N-terminal regions of the mature soybean and sweet potato β-amylases (4, 9). Kreis *et al.* (2) also noted the lack of a signal sequence in the barley β-amylase cDNA, although the precise subcellular location of that protein is not known.

ACKNOWLEDGMENTS

We thank Dr. K. Ball and J. Leykam for assistance in peptide sequencing and Dr. N.-H. Chua for the generous gift of the cDNA library.

LITERATURE CITED

1. **Casper T, Lin TP, Monroe J, Bernhard H, Spilatro S, Preiss J, Somerville C** (1989) Altered regulation of β-amylase activity in mutants of *Arabidopsis* with lesions in starch metabolism. Proc Natl Acad Sci USA **86:** 5830–5833
2. **Kreis M, Williamson M, Buxton B, Pywell J, Hejgaard J, Svendsen I** (1987) Primary structure and differential expression of β-amylase in normal and mutant barleys. Eur J Biochem **169:** 517–525
3. **Lin TP, Spilatro SR, Preiss J** (1988) Subcellular localization and characterization of amylases in *Arabidopsis* leaf. Plant Physiol **86:** 251–259
4. **Mikami B, Morita Y, Fukazawa C** (1988) Primary structure and function of β-amylase. Seikagaku (Tokyo) **60:** 211–216 (in Japanese)
5. **Monroe JD, Preiss J** (1990) Purification of a β-amylase that accumulates in *Arabidopsis thaliana* mutants defective in starch metabolism. Plant Physiol **94:** 1033–1039
6. **Nitta Y, Isoda Y, Toda H, Sakiyama F** (1989) Identification of glutamic acid 186 affinity-labeled by 2,3-epoxypropyl α-D-glucopyranoside in soybean β-amylase. J Biochem **105:** 573–576
7. **Okita TW, Greenberg E, Kuhn DN, Preiss J** (1979) Subcellular localization of the starch degradative and biosynthetic enzymes of spinach leaves. Plant Physiol **64:** 187–192
8. **Preiss J** (1982) Regulation of the biosynthesis and degradation of starch. Annu Rev Plant Physiol **33:** 431–454
9. **Toda H** (1989) Sequence analysis of sweet potato β-amylase. Denpun Kagaku **36:** 87–101 (in Japanese)
10. **von Heijne G** (1985) Signal sequences: the limits of variation. J Mol Biol **184:** 99–105
11. **Ziegler P, Beck E** (1986) Exoamylase activity in vacuoles isolated from pea and wheat leaf protoplasts. Plant Physiol **82:** 1119–1121