Plant Gene Register

Richard Graham and Robert Thornburg (1997) DNA Sequence of UDP Glucose:Indole-3-acetate Beta-D-Glucosyltransferase from *Arabidopsis thaliana* (Accession No. U81293) (PGR97-044). Plant Physiol. **113**: 1004

DNA Sequence of UDP Glucose:Indole-3-acetate Beta-D-Glucosyltransferase from *Arabidopsis thaliana* (Accession No. U81293)

Richard Graham and Robert Thornburg (*)

Department of Biochemistry and Biophysics, 2232 Molecular Biology Building, Iowa State University, Ames, Iowa 50011

(*) Corresponding Author
FAX: 1-515-294-0453
Email: thorn@iastate.edu

Journal Paper Number J-17203 of the Iowa Agriculture and Home Economics Experiment Station, Ames, Iowa, Project No. 3340, and supported by Hatch Act and State of Iowa funds.

This cDNA clone was requested from the Arabidopsis Biological Resource Center at Ohio State University based upon its previously published identity [GenBank N38403] as a Germin-like protein (Newman *et al*., 1994). However, sequencing of the clone identified the insert as IAA Glucose Synthase, not as the indicated Germin-like protein. The original clone for the Germin-like protein (216C14T7) presumably has been lost. A search of the ABRC revealed another clone (176A10T7) that also encodes IAA Glucose Synthase. It is not clear whether the clone reported herein and 176A10T7 are the same clone or not.

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. Sequence was initiated from known vector sequences. On the basis of these runs, primers specific to the IAA Glucose Synthase sequence were constructed. DNA sequences were performed in duplicate or triplicate for each run. The entire sequence of the cDNA was confirmed on the opposite strand.

The cDNA encoding IAA Glucose Synthase is 1612 nucleotides long. The cDNA is evidently full length. A methionine start codon was found at [26..28] and an opal stop codon is located at [1448..1450]. No obvious polyadenylation signal was observed; however, the best fit (ATGAAA)is found at [1568..1575], which is 21 nucleotides upstream from the polyadenylation site [1589].

The *Arabidopsis thaliana* IAA Glucose Synthase is a 474 amino acid protein that shares 30% identity with the *Zea mays* protein (Altschul *et al*., 1990). If conservative substitutions are permitted then the conserved identity between the *Arabidopsis* and *Zea* proteins is 42.7%.

The maize IAA Glucose Synthase contains a single site of N-glycosylation (Szerszen *et al*., 1994). In contrast, the Arabidopsis IAA Glucose Synthase contains 4 sites of N-glycosylation. One of these putative four Arabidopsis sites of glycosylation is conserved with the maize protein (Asn371). In addition there are two potential sites in the N-terminus as well as one potential site in the C-terminus. The second of these potential N-glycosylation sites is in a region of high identity between the Arabidopsis and Zea enzymes and may not be glycosylated.

The maize IAA Glucose Synthase was identified as having three potential kinase C sites (S/T)-X-(R/K) that

have been proposed to regulate protein action (Szerszen *et al* ., 1994). In the maize protein, one site was located near the N-terminus and two sites were located within 20 amino acids of the C-terminus. The Arabidopsis protein also contains multiple (six) potential protein kinase C sites scattered throughout the primary sequence. However, none of these potential protein kinase C sites is located near any of the potential protein kinase C sites in the maize protein. Therefore, either the three-dimensional structure of the IAA Glucose Synthase places these potential protein kinase C sites into a structural context that can be phosphorylated or perhaps protein kinase C does not regulate the IAA Glucose Synthase as previously hypothesized. The protein does not contain readily discernible N-terminal targeting sequences that would direct the IAA Glucose Synthase toward the secretory pathway, the nucleus, the mitochondria or the chloroplast.

Acidic and basic amino acids were scattered throughout the protein, but one region from amino acid 316 to 342, which corresponded to the most hydrophilic region of the protein when analyzed by hydrophobicity plots, was extremely rich (44.4%) in acidic amino acids. This region was not conserved between the Arabidopsis and Zea proteins.

Table 1. Characteristics of the *Arabidopsis thaliana* cDNA for IAA Glucose Synthase

Organism:
Arabidopsis thaliana cv Columbia

Source:
Clone from the Arabidopsis Biological Resource Center. 216C14T7.

Function of Gene:
UDP Glucose:Indole-3-acetic acid beta-D-glucosyltransferase IAA Glucose Synthase (EC 2.4.1.121) catalyzes the conjugation of indole-3-acetic acid with glucose. Transfers glucose moeity from UDP-glucose to form ester linked IAA Glucose.

Method of Identification:
Sequence comparison to Zea mays IAA Glucose Synthase.

Characteristics of cDNA:
The cDNA is contained on a 1612 nucleotide fragment. A 25 bp 5' untranslated sequence [1..25] A stop codon (TGA) at [1448..1450]. A 164 bp 3' untranslated sequence [1448..1612]. No obvious polyadenylation signal was observed. The best fit (ATGAAA)is found at [1568..1575], which is 21 nucleotides upstream from the polyadenylation site [1589].

Features of the deduced protein:
A 474 amino acid protein that shares 30.0 % identity with the *Zea mays* IAA Glucose Synthase. If conservative substitutions are permitted then the conserved identity between the *Arabidopsis* and *Zea* proteins is 42.7 %. Potential N-glycosylation sites (N)-(X)-(S/T): Asn4, Asn26, Asn371, Asn444 Potential Protein Kinase C regulatory sites (S/T)-(X)-(R/K): Ser180, Ser263, Ser292, Ser306, Ser318, Ser336

Theoretical pI:
5.09

Theoretical MW:
53835 Da

Amino Acid Composition:
Sulfur Containing Amino Acids 4.0%
Acidic Amino Acids 14.4%
Basic Amino Acids 12.5%
Hydrophobic Amino Acids 31.6%

Glutamine 2.5%

Acknowledgments:

This work was supported by grants from the US Department of Agriculture (91-37301-6208) and from the Iowa State Biotechnology Council.

Literature Cited

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. (1990). Basic local alignment search tool. J. Mol. Biol. **215**:403-10.

Newman,T., de Bruijn,F.J., Green,P., Keegstra,K., Kende,H., McIntosh,L., Ohlrogge,J., Raikhel,N., Somerville,S., Thomashow,M., Retzel,E. and Somerville,C. (1994) Genes galore: a summary of methods for accessing results from large-scale partial sequencing of anonymous Arabidopsis cDNA clones. Plant Physiol. **106**:1241-1255

Szerszen, J.B., Szczyglowski, K. and Bandurski, R.S. (1994) iaglu, a gene from Zea mays involved in conjugation of growth hormone indole-3-acetic acid., Science **265**:1699-1701.

 Return to Plant Gene Register Index

Return to Plant Physiology ONLINE