

Plant Gene Register

Nucleotide and Protein Sequences of a Cytoplasmic Ribosomal Protein S15a Gene from *Arabidopsis thaliana*¹

Peta C. Bonham-Smith* and Maurice M. Moloney

Biomira Inc., c/o Plant Biotechnology Institute, 110 Gymnasium Place, Saskatoon, Saskatchewan, Canada (P.C.B.-S.); and Department of Biological Sciences, University of Calgary, University Drive NW, Calgary, Alberta, Canada (M.M.M.)

The eukaryotic ribosome is composed of 70 to 80 ribosomal proteins (r-proteins) together with three or four rRNAs. In yeast, the rate of ribosome formation reflects that of cellular growth, with a balance in the production of ribosomal constituents being maintained in part by the activity of the respective promoters (Mager and Planta, 1991). Only a small number of genes or cDNAs encoding r-proteins have been characterized in monocotyledonous or dicotyledonous plants (Bonham-Smith et al., 1992, and refs. therein; Taylor et al., 1992; Hwang and Goodman, 1993; Joanin et al., 1993; Lu et al., 1993). In all plants studied to date, r-protein transcripts have been found to be more abundant in actively dividing tissues (Bonham-Smith et al., 1992, and refs. therein); however, the factors involved in this developmental regulation of r-protein gene expression have yet to be determined. Isolation and characterization of additional r-protein genes will help to identify the regulatory elements required for coordinate r-protein and rRNA synthesis and subsequent ribosome biogenesis. We report here the isolation and nucleic acid sequence of a genomic clone for *Arabidopsis thaliana* (var Columbia) r-protein S15a (Table I).

A *Brassica napus* (cv Westar) r-protein S15a cDNA (pA813; Bonham-Smith et al., 1992) was used to screen a genomic library of *A. thaliana* in EMBL3A (Stratagene, La Jolla, CA). After tertiary screening and endonuclease mapping, a 2702-nucleotide *Sall* (5′)-*EcoRI* (3′) genomic fragment was cloned. The cloned fragment contains a 1662-nucleotide region 5′ to the putative start of translation, 651 nucleotides of coding sequence interrupted by two introns (259 nucleotides), and a 3′ region of 389 nucleotides. Placement of the two introns resulted from an alignment with the *B. napus* r-protein S15a cDNA sequence (Bonham-Smith et al., 1992). Possible TATA and CAAT boxes are located at −364 (TATAGAA) and −458 (GCTCCCAATCT), respectively. An alignment of the 5′ UTR of the *B. napus* cDNA and the equivalent region of the *A. thaliana* gene (start of transcription unknown) shows a number of short sequence motifs common to both UTRs. The 3′ terminal 100 nucleotides of the two 5′ UTRs are 43% identical in sequence. The absence of a recognizable consensus poly-

Table I. Characteristics of the ribosomal protein S15a gene from *Arabidopsis thaliana*

Organism:	<i>Arabidopsis thaliana</i> Columbia.
Genome Location:	Nuclear genome.
Function:	Encodes a cytosolic ribosomal protein, S15a.
Method of Identification:	Comparison of deduced amino acid sequence with that of r-protein S15a from <i>Brassica napus</i> (100%), rat (76%), human (S24, 70%), and the fungi <i>Saccharomyces cerevisiae</i> (S24, 80%) and <i>Kluyveromyces marxianus</i> (S24, 79%).
Techniques:	Genomic library screening, isolation of phage DNA, and restriction fragment subcloning into pBluescript vector, dideoxy sequencing of both strands using standard and custom-made sequencing primers.
Features of Genomic Fragment Containing The R-Protein S15a Sequence:	1662 nucleotides 5′ to the putative start of translation; 651 nucleotides of coding sequence interrupted by two introns (259 nucleotides: placement of the two introns from alignment with <i>B. napus</i> r-protein S15a sequence); 389 nucleotides 3′ to the translation termination codon; possible TATA box (TATAGAA) at nucleotide 1298; possible CAAT box (GCTCCCAATCT) at nucleotide 1204.
Structural Features of Deduced Protein Sequence:	The coding region encodes a polypeptide of 130 amino acids with a calculated mol wt of 14,778 D and a +9 net charge.
Expression Characteristics:	Not determined.
Antibody:	Not available.

adenylation signal sequence T/AAATAAA/G in the 3′ UTR has been reported previously for r-proteins in *A. thaliana*, *B. napus*, and maize (Bonham-Smith et al., 1992, and refs. therein). However, in this gene an AATAAC motif at position 947 may be sufficient to function as a polyadenylation signal.

A Southern blot of restriction enzyme-digested *A. thaliana* genomic DNA probed with the *rps15a* sequence showed only one copy of this gene in the *A. thaliana* genome (data not shown). The deduced amino acid sequence of the AtS15a protein shows 100% identity with that for *B. napus* r-protein

¹ This work was supported by the Strategic Grants Program (award STR0024990 to M.M.M.) from the Natural Sciences and Engineering Research Council of Canada.

* Corresponding author; fax 1-306-975-4839.

S15a (Bonham-Smith et al., 1992), 80 and 79% with the fungi *Saccharomyces cerevisiae* (Leer et al., 1985) and *Kluyveromyces marxianus* (Bergkamp-Steffens et al., 1992), respectively, 76% with rat (over the first 38 nucleotides, I.G. Wool and Y.L. Chan, personal communication), and 70% with human (Schwabe, 1991). The function of r-protein S15a and its interactions with other r-proteins or rRNA are unknown; however, the degree of sequence conservation among the human, rat, yeast, and plant r-proteins suggests an evolutionary constraint on the form or function of this protein within the ribosome.

Received March 1, 1994; accepted April 10, 1994.

Copyright Clearance Center: 0032-0889/94/106/0401/02.

The GenBank accession number for the sequence reported in this article is L27461.

LITERATURE CITED

- Bergkamp-Steffens GK, Hoekstra R, Planta RJ** (1992) Structural and putative regulatory sequences of *Kluyveromyces* ribosomal protein genes. *Yeast* **8**: 903–922
- Bonham-Smith PC, Oancia TL, Moloney MM** (1992) Cytoplasmic ribosomal protein S15a from *Brassica napus*: molecular cloning and developmental expression in mitotically active tissue. *Plant Mol Biol* **18**: 909–919
- Hwang I, Goodman HM** (1993) Cloning of an *Arabidopsis* ribosomal protein S28 cDNA. *Plant Physiol* **102**: 1357–1358
- Joanin P, Gigot C, Philipps G** (1993) cDNA nucleotide sequence and expression of a maize cytoplasmic ribosomal protein S13 gene. *Plant Mol Biol* **21**: 701–704
- Leer RJ, van Raamsdork-Durin MMC, Kraakman P, Mager WH, Planta RJ** (1985) The genes for yeast ribosomal proteins S24 and L46 are adjacent and divergently transcribed. *Nucleic Acids Res* **13**: 701–709
- Lu G, Wu K, Ferl RJ** (1993) A cDNA for *Arabidopsis* cytosol ribosomal protein S11. *Plant Physiol* **102**: 695–696
- Mager WH, Planta RJ** (1991) Coordinate expression of ribosomal protein genes in yeast as a function of cellular growth rate. *Mol Cell Biochem* **104**: 181–187
- Schwabe G** (1991) *H. sapiens* mRNA for ribosomal protein (homologous to yeast S24). GenBank data base, accession No. X62691
- Taylor MA, Mad Arif SA, Pearce SR, Davies HV, Kumar A, George LA** (1992) Differential expression and sequence analysis of ribosomal protein genes induced in stolon tips of potato (*Solanum tuberosum* L.) during the early stages of tuberization. *Plant Physiol* **100**: 1171–1176