Soybean (Glycine max L.)
Nuclear DNA Contains Four
tufA Genes Coding for the
Chloroplast-specific Translation
Elongation Factor EF-Tu

Christophe Bonnaï and Erhard Stutz

Abstract. The chloroplast tufA gene codes for the translation elongation factor EF-Tu. In algae like e.g. Euglena tufA is part of the chloroplast DNA but in land plants the equivalent gene is transposed to the nuclear DNA. Using a previously sequenced Euglena gracilis tufA DNA probe a cDNA library from soybean was screened. Two cDNA clones (cDNA1 and cDNA2) were sequenced (ca. 500 nucleotides) and shown to contain parts of the chloroplast tufA gene. A cDNA1 fragment was used to screen a genomic library and one complete tufA gene was sequenced including 640 nucleotides of the region upstream of the transcription start site which was identified by S1 endonuclease protection experiments. The soybean genome contains four tufA genes which belong to two subfamilies with two members each. TufA mRNA is absent in dark grown seedlings but light grown seedlings contain a stable tufA mRNA of ca. 1.6–1.7 kb. The open reading frame codes for a chloroplast transit peptide (71 amino acids) and a very conservative chloroplast EF-Tu protein.

Introduction

Chloroplasts are considered to be endosymbionts of prokaryotic origin having invaded in an evolutionary past nucleated heterotrophic cells. During evolution parts of the organelar genome were transposed to and successfully integrated in the nuclear genome. Structure and function of the photosynthetic organelle depend on both chloroplast and nuclear DNA encoded gene products. Recent comparative sequence studies indicate that number and kind of genes located on chloroplast genomes from different chlorophytes vary. A case in point is the tufA gene which codes for the translation elongation factor EF-Tu. This GTP requiring protein interacts with aminoacyl-tRNA and delivers it to the ribosomal A-site during the translation elongation step. Photosynthetic algae like e.g. Euglena gracilis and Chlamydomonas reinhardtii [2] contain a single tufA gene in the chloroplast genome but this gene is missing from the chloroplast DNA of all examined land plants [2]. A single nuclear tufA gene was mapped and sequenced in Arabidopsis thaliana [2][3] and multiple copies of this gene seem to exist in the nuclear genome of other Brassicaceae [3]. Two tufA genes were mapped and totally sequenced in Nicotiana sylvestris [4]. In the following we report about structure and expression of tufA genes located in the soybean nuclear genome.

Results and Discussion

We used a 2 kb (kilo basepairs) tufA DNA probe from Euglena gracilis to screen a cDNA library from soybean [5]. Two positive clones were selected and the inserts sequenced (data not shown). Both inserts contained segments coding for the C-terminal part of EF-Tu. A sequence comparison of equivalent parts of the two inserts of ca. 500 nucleotides revealed 84 and 95% sequence identity on the nucleotide and amino-acid level, respectively, clearly indicating that at least two actively transcribed tufA genes exist. Equivalent fragments were cut from both cDNA samples (cDNA1 and cDNA2) and used as probes in Southern experiments (Fig. 1). We see that either probe interacts with four HindIII fragments albeit in a differential manner: cDNA2 (panel 1) gives a strong signal with the two middle sized fragments while cDNA1 (panels 2, 3) preferentially interacts with the shortest and the longest of the four fragments. This indicates that the soybean genome contains four tufA genes and that the four genes belong to two related families with two members each. Note that soybean is an amphidiploid plant [6] and each parent may have contributed two tufA genes.

A tufA cDNA probe (360 bp) was used to screen a genomic DNA library [5]. We show in Fig. 2 a fragment (2313 positions) containing one complete tufA gene coding for the EF-Tu protein and a N-terminal peptide of 71 amino-acids with structural features of a chloroplast transit peptide [7]. The decoded EF-Tu protein starting with A(72) has e.g. 95 and 88% sequence similarity with the A. thaliana [2] and the E. gracilis [1] counterparts and 77 and 84% similarity with the yeast mitochondrial [8] and the bacterial [9] EF-Tu proteins. Note that the sequence 5'GCAUGG is in line with the consensus sequence of competent initiator co-

Correspondence: Prof. E. Stutz
Laboratoire de Biochimie végétale
Université de Neuchâtel
Chamenerie 18
CH-2000 Neuchâtel

*) Present address: Northwestern University
BMBCB
Evanston, Ill. 60208, USA

Fig. 1. Southern blot of soybean nuclear DNA cut with HINDIII. A filter imprint was first hybridized with the cDNA2 probe (panel 1) and after washing rehybridized with cDNA1 (panel 2); another filter imprint (panel 3) was hybridized with cDNA1; size marker in kb. Hybridization conditions: 6 x SSC, 68°, 12 h.
Fig. 2. Nucleotide sequence of a tufA gene [15]. The open reading frame is decoded, the first methionine (M) after the presumed transcription start site (pos. +1, vertical bars) and the first amino-acid of EF-Tu (A) are in bold letters. The last A (pos. 1673) corresponds to the beginning of the poly(A) sequence of the cDNA. A potential TATA box (pos. -30 to -24) and a crucial cis element (pos. -420 to -388) are in bold letters. The DNA sequence complementary to the primer used for the synthesis of single-stranded DNA probes and the upstream HaeIII site (8I-endonuclease-protection experiments) are underlined.
Fig. 3. Northern blot of soybean RNA. RNA from soybean seedlings (dark grown, or exposed to light for the indicated time in h) were electrophoretically separated and filter imprints were probed with a 360 bp DNA fragment from cDNA1; size marker (margin) is in kb. Hybridization in 6x SSC, 50% formamide, 42°C, 2 h.

Fig. 4. S1-Endonuclease-protection experiment. A primed single-strand DNA probe [11] covering a region starting with the primer (see Fig. 2) and extending to the HaeIII site was hybridized to RNA. Protected fragments (panel SI) were separated and the untreated DNA probe (panel P). The 13 nucleotides given in the margin are complementary to the sequence (pos. -420 to -388) which is very similar to a 'cis' element in the promoter region of pea rbcS genes [12] considered to be involved in the light-induction process.

Experimental

Isolation of DNA. DNA from soybean leaves was isolated following the procedure of Shure et al. [11]. Fresh leaves frozen in liquid N2 (5-10 g) are ground to a fine powder and resuspended in 2.25 ml/g leaves of lysis buffer (0.1 M NaCl, 0.015 M Tris-Cl, pH 7.5, 0.02 M EDTA, 2% Sarkosyl and 5% phenol). The suspension is homogenized by gentle shaking before adding one volume of phenol/CHCl3 1:1 and one volume of CHCl3. Finally, the supernatant is extracted with 0.1 M NaCl, 0.015 M Trisodium citrate, pH 7.0 and 1 x Denhardt's (50 x Denhardt's = 1% Ficoll, 1% polyvinylpyrrolidine, 1% bovine serum albumin). For hybridization conditions see legends to Figs. 1 and 3.

Cloning of DNA Fragments and DNA Sequencing. The protocols described in detail in [11] and the instructions of the suppliers of vectors, plasmids, and enzymes (Boehringer, Mannheim and Stratagene, La Jolla) were followed.

We are grateful to Dr. A. Spielmann for helpful advice and Sophie Marc-Martin for technical assistance. This work receives support from the Swiss National Science Foundation (to E.S.) and it is part of the Ph.D. Thesis of Ch. B. submitted to and accepted by the Faculté des Sciences, Université de Neuchâtel.

Received: April 6, 1993

[16] The nucleotide sequence appears in the EMBL, GenBank and DDBJ Nucleotide Sequence Database under the Accession number X66062 (G. max tufA gene for chloroplast elongation factor EF-Tu).