Chimia 47 (1993) 247–249

ISSN 0009–4293

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

Christophe Bonny[a]) and **Erhard Stutz***

Abstract. The chloroplast *tuf*A gene codes for the translation elongation factor EF-Tu. In algae like *e.g.* Euglena *tuf*A 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 tuf*A 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 *tuf*A gene. A cDNA1 fragment was used to screen a genomic library and one complete *tuf*A 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 *tuf*A genes which belong to two subfamilies with two members each. *Tuf*A mRNA is absent in dark grown seedlings but light grown seedlings contain a stable *tuf*A 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 organellar genome were transposed to and successfully integrated in the nuclear genome ,*i.e.*, 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 *tuf*A 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* [1] and *Chlamydomonas reinhardtii* [2] contain a single *tuf*A gene in the chloroplast genome but this gene is missing from the chloroplast DNA of all examined land plants [2]. A single nuclear *tuf*A 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 *tuf*A genes were mapped and totally sequenced in *Nicotiana sylvestris* [4]. In the following we report about structure and expression of *tuf*A genes located in the soybean nuclear genome.

Results and Discussion

We used a 2 kb (kilo basepairs) *tuf*A 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 identy on the nucleotide and amino-acid level, respectively, clearly indicating that at least two actively transcribed *tuf*A 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 *tuf*A genes and that the four genes belong to two related families with two members each. Note that soybean is an amphidiploide plant [6] and each parent may have contributed two *tuf*A genes.

A *tuf*A cDNA1 probe (360 bp) was used to screen a genomic DNA library [5]. We show in *Fig. 2* a fragment (2313 positions) containing one complete *tuf*A 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'GCC**AUG**G is in line with the consensus sequence of competent initiator co-

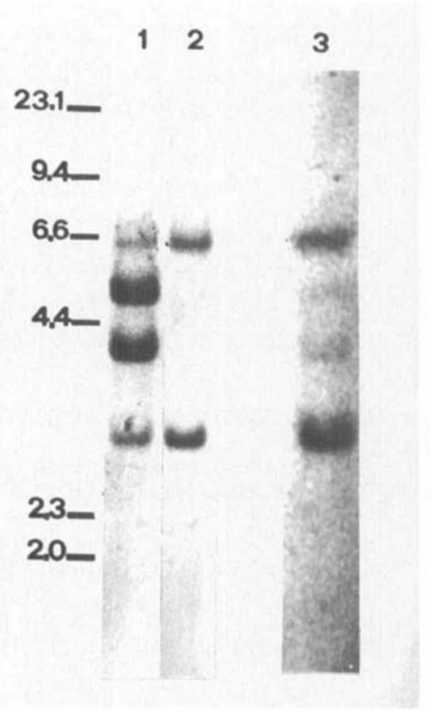

Fig. 1. *Southern blot of soybean nuclear DNA cut with HINDIII.* A filter imprint was first hybridized with the cDNA2 probe (panel 1) and after washing rehybridized with cDNA1 (panel 2); another filter imprint (panel 3) was hybridized with cDNA1; size marker in kb. Hybridization conditions: 6 x SSC, 68°, 12 h.

*Correspondance: Prof. E. Stutz
Laboratoire de Biochimie végétale
Université de Neuchâtel
Chantemerle 18
CH–2000 Neuchâtel
[a]) Present address: Northwestern University
BMBCB
Evanston, Ill. 60208, USA

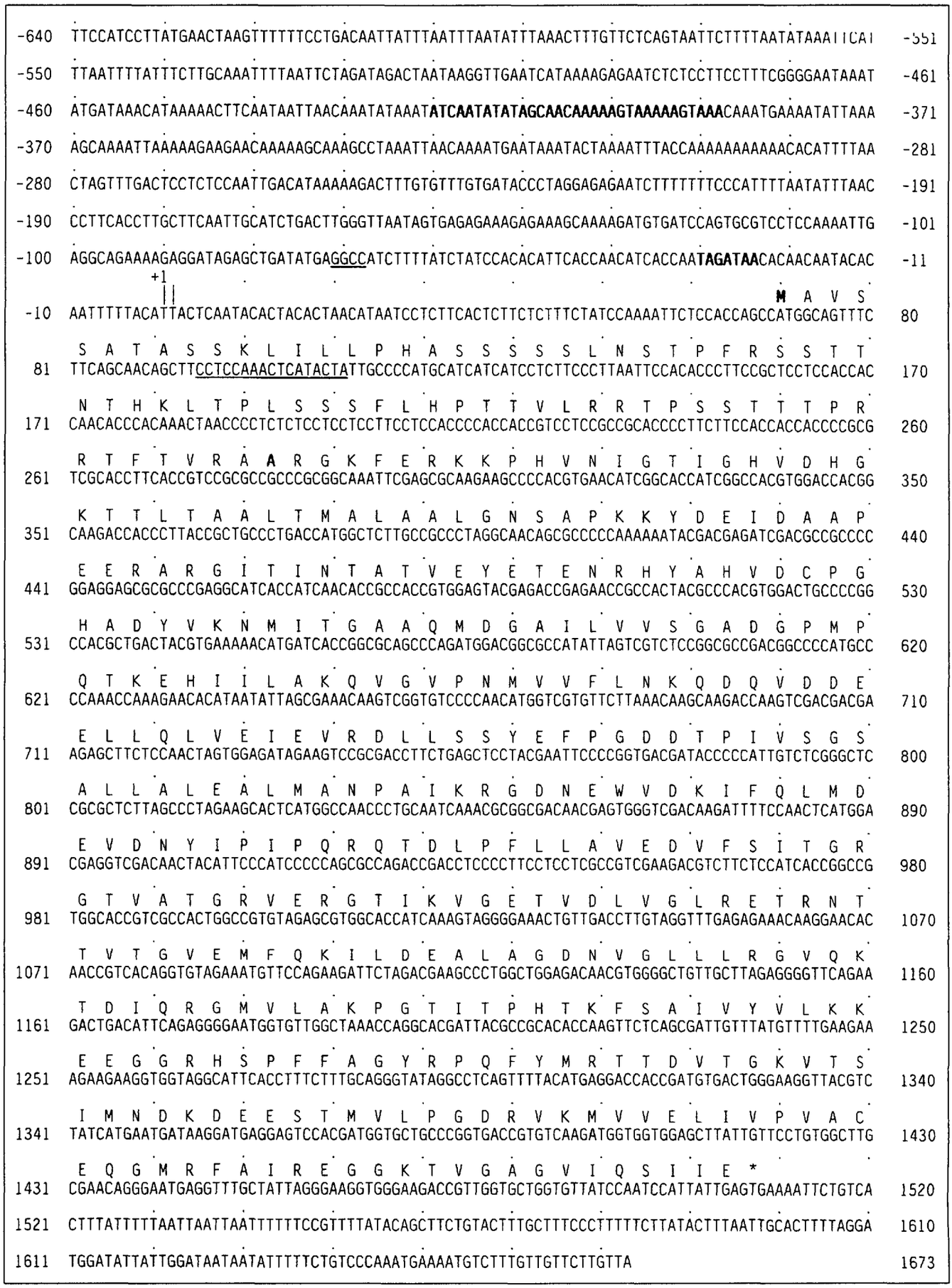

Fig. 2. *Nucleotide sequence of a* tuf*A gene* [15]. The open reading frame is decoded, the first methionine (**M**) after the presumed transcription start site (pos. +1, vertical bars) and the first amino-acid of EF-Tu (**A**) are in bold letters. The last A (pos. 1673) corresponds to the beginning of the poly(A) sequence of the cDNA1. 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 (S1-endonuclease-protection experiments) are underlined.

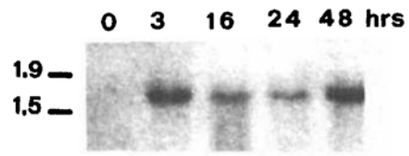

Fig. 3. *Northern blot of soybean RNA.* RNA from soybean seedlings (dark grown, or exposed to light for the indicated time in h) were electrophoretically separated and filter imprints were probed with a 360 bp DNA fragment from cDNA1; size marker (margin) is in kb. Hybridization in 6 x SSC, 50% formamid, 42°, 12 h.

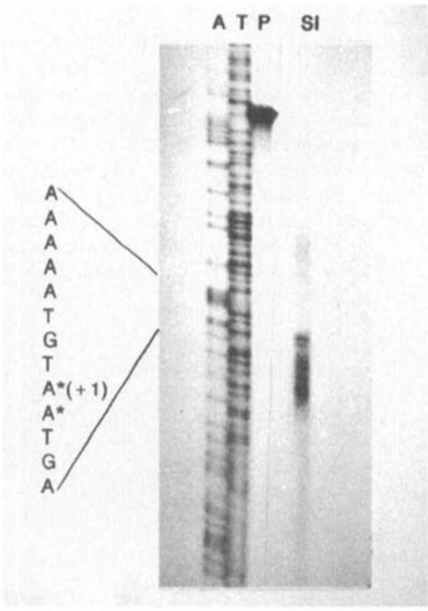

Fig. 4. *S1-Endonuclease-protection experiment.* A primed single-strand DNA probe [11] covering a region starting with the primer (see *Fig. 2*) and extending to the HaeIII site was hybridized to RNA. Protected fragments (panel SI) were separated on a sequencing gel along with the A and T sequencing probes of the corresponding DNA fragment and the untreated DNA probe (panel P). The 13 nucleotides given in the margin are complementary to the sequence (pos. –8 to +5) of *Fig. 2*.

dons [10]. The cDNA1 sequence (932 positions of the 3' part, not shown) perfectly matches the genomic sequence given in *Fig. 2*, *i.e.*, the *tuf*A gene presented in *Fig. 2* is transcribed.

According to Northern experiments (*Fig. 3*) the *tuf*A DNA probe interacts with a stable transcript of *ca.* 1.6 –1.7 kb. The band is absent in RNA samples from dark grown seedlings. Upon illumination the steady state level of the *tuf*A mRNA sharply increases remaining at about the same level over the monitored time (48 h).

In order to identify the transcription start site we made S1-endonuclease-protection studies using *in vitro* synthesized single-strand DNA fragments [11]. The result given in *Fig. 4* was obtained with a DNA probe going from the HaeIII site in the upstream region to the priming site in the coding region (consult *Fig. 2*). The longest protected fragment starts with AATG- what corresponds to the TTAC-sequence marked in *Fig. 2*. Identical results were obtained with longer DNA probes extending to position –520.

Considering the size of the *tuf*A mRNA (Northern) and the length of the transcript according to the S1 experiment we can assume that this *tuf*A gene has no intron in the coding part and most likely none within the promoter region (see below). This clearly distinguishes the *tuf*A gene from the *tef* genes (coding for the eEF-1α translation elongation factor) which are interrupted by an intron both in the coding part and the upstream region [5].

We cloned the entire region upstream of the presumed transcription start site into an expression vector and tested the construct for promoter activity in transgenic tobacco (*F. Maurer, Ch. Bonny, A. Spielmann*, and *E. Stutz*, to be published). The results obtained so far indicate that the upstream region (negative numbering in *Fig. 2*) acts as strong promoter. We mark in *Fig. 2* two regions (bold letters) which may be instrumental in promoter function, *i.e.*, a TATA-box like element near the transcription start site and a longer sequence (pos. –420 to –388) which is very similar to a '*cis*' element in the promoter region of pea *rbc*S 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.* [13]. Fresh leaves frozen in liquid N_2 (5–10 g) are ground to a fine powder and resuspended in 8 ml/g leaves of lysis buffer (8M urea, 0.35M NaCl, 0.05M Tris-Cl, pH 7.5, 0.02M EDTA, 2% sarcosyl and 5% phenol). The suspension is homogenized by gentle shaking. One volume of phenol/$CHCl_3$ 3:1 is added and sodium dodecyl sulfate (SDS) to a final concentration of 0.5%. Subsequent steps of DNA extraction are as described [13]. The yield of RNA-free DNA is *ca.* 0.1 mg DNA/g leaves.

Isolation of RNA. The procedure of *Dean et al.* [14] was followed. Fresh leaves frozen in liquid N_2 (batches of 2 g) are ground to a fine powder and resuspended in 2.25 ml/g leaves of lysis buffer (0.1M NaCl, 1 mM Tris-Cl, pH 7.5, 1mM EDTA, 1% SDS). The suspension is homogenized by gentle shaking before adding one volume of phenol/$CHCl_3$ 1:1). Subsequent steps af RNA extraction are as described [14]. The yield of DNA-free RNA is *ca.* 2 mg/g leaves.

Blotting and Hybridization. Transfer of gel-electrophoretically separated RNA (Northern blot) and DNA (Southern blot) to filters was according to standard procedures [11]. Hybridization of filters with ^{32}P-labelled DNA probes was done in buffer 6 x SSC (SSC = 0.15M NaCl, 0.015M trisodium citrate, pH 7.0) and 1 x D*enhardt*'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.

^{32}P-Labelling of Single-Strand DNA and S1-Endonuclease-Protection Experiments. The protocol of *Sambrook et al.* [11] was followed using a construct (pBluescript II KS-) containing the EcoRI-XbaI fragment (1.3 kb) which carries 590 nucleotides upstream of the start codon (AUG). A DNA primer (17-mer) complementary to nucleotide positions 40 to 25 (see *Fig. 2*) was used to copy the insert in the presence of α^{32}PdATP (400 Ci/mmol, 10 mCi/ml, *Amersham*).

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

[1] P.E. Montandon, E. Stutz, *Nucleic Acids Res.* **1983**, *11*, 5877.
[2] S.L. Baldauf, J.D. Palmer, *Nature* (*London*) **1990**, *344*, 262.
[3] S.L. Baldauf, J.R. Manhart, J.D. Palmer, *Proc. Natl. Acad. Sci. U.S.A.* **1990**, *87*, 5317.
[4] Y. Murayama, T. Matsubayashi, M. Sugita, M. Sugiura, in 'Third International Congress of Plant Molecular Biology', Ed. R.B. Hallick, 1991, Abstract 1949.
[5] F. Aguilar, P.E. Montandon, E. Stutz, *Plant Mol. Biol.* **1991**, *17*, 351.
[6] M.A. Grandbastien, S. Berry-Lowe, B.W. Shirley, R.B. Meagher, *Plant Mol. Biol.* **1986**, *7*, 451.
[7] G. Von Heijne, J. Steppuhn, R.G. Herrmann, *Eur. J. Biochem.* **1989**, *180*, 535.
[8] S. Nagata, Y. Tsunetsugu-Yokota, A. Naito, Y. Kaziro, *Proc. Natl. Acad. Sci. U.S.A.* **1983**, *80*, 6192.
[9] T. Yokota, H. Sugisaki, M. Takanami, Y. Kaziro, *Gene* **1980**, *12*, 25.
[10] M. Kozak, *Nucleic Acids Res.* **1981**, *9*, 5233.
[11] J. Sambrook, E.F. Fritsch, T. Maniatis, 'Molecular Cloning: A Laboratory Manual', Cold Spring Harbor Laboratory, 1989.
[12] R. Fluhr, C. Kuhlemeier, F. Nagy, N.H. Chua, *Science* **1986**, *232*, 1106.
[13] M. Shure, S. Wessler, N. Federoff, *Cell* **1983**, *35*, 225.
[14] C. Dean, T. Tamaki, P.M. Dunsmuir, M. Favreau, C. Katayama, H. Dooner, J. Bedbrook, *Nucleic Acids Res.* **1986**, *14*, 2229.
[15] The nucleotide sequence appears in the EMBL, GenBank and DDBJ Nucleotide Sequence Database under the Accession number X66062 (G. max *tuf*A gene for chloroplast elongationfactor EF-Tu).