Acta Biochimica Polonica

Vol. 44 No. 2/1997

259–274

QUARTERLY

Nucleotide sequence of nuclear tRNA^{Gly} genes and tRNA^{Gly} pseudogenes from yellow lupin (*Lupinus luteus*): Expression of the tRNA^{Gly} genes *in vitro* and *in vivo**⊡

Przemysław Nuc, Katarzyna Nuc, Zofia Szweykowska-Kulińska[1], Jerzy Pawełkiewicz✉

Institute of Biochemistry and Biotechnology, University of Agriculture, Wołyńska 35, 60-637 Poznań, Poland
[1]Institute of Molecular Biology and Biotechnology, Adam Mickiewicz University, Międzychodzka 5, 60-637 Poznań, Poland

Received: 26 March, 1997

Key words: genes, expression, organization, glycine

A nuclear DNA fragment (7.8 kb) from yellow lupin (*L. luteus*) was sequenced and shown to contain tRNA^{Gly}(GGC) genes and tRNA^{Gly}(GGC) pseudogenes organized in three tandemly repeated units: of 2565 bp and 2564 bp, and one, truncated from its 3′ end, of 1212 bp. Each unit contains an identical pair of a tRNA^{Gly} gene and a pseudogene, both having the same polarity. The nucleotide sequence of the gene appears colinear to *L. luteus* cytoplasmic tRNA^{Gly}(GGC) primary structure. All three genes are efficiently transcribed in HeLa-cell nuclear extract giving two primary transcripts. The main, longer primary transcripts have each an extremely long 3′ trailer of about 100 nucleotides, the structure of which is specific only for tRNA^{Gly} genes and pseudogenes (80% homology) of the studied tandem (but not for other tRNA^{Gly} genes of the yellow lupin genome) as it has been shown by Southern hybridization. This distinctive feature allowed to isolate putative tRNA^{Gly} precursor(s) encoded by at least one of the three tRNA^{Gly}(GGC) genes from *L. luteus* seedlings.

Nuclear tRNA genes are present in eucaryotes in multiple copies. Some of them are dispersed throughout the genome, the others form clusters at single chromosomal sites, which consist either of the same or different tRNA gene species [1]. This conclusion is based mainly on studies on animal genes, though scarce data confirm such an organization also for plants. For example, *Arabidopsis thaliana* tRNA genes specific for tyrosine, serine [2] and cysteine [3] appear in clusters, several intron-containing methionine tRNA genes of this plant are probably also clustered, whereas others are dispersed,

*This work was supported by the State Committee for Scientific Research (KBN grant: PB-6P 204 050 04).

⊡The nucleotide sequence data reported in this paper will appear in the EMBL, GenBank and DDBJ Nucleotide Sequence Databases under the accession number Z49255.

✉Correspondence to: Jerzy Pawełkiewicz, Tel: 061 48-72-04

Abbreviations: kb, kilo base pair; nt, nucleotide; dNTP, deoxynucleoside-5′-triphosphate; PCR, polymerase chain reaction; M-MLV, Moloney mouse leukemia virus; DEPC, diethylpyrocarbonate; SSC, 0.15 M sodium chloride, 0.015 M Na3-citrate, pH 7.0; SDS, sodium dodecyl sulphate.

like $tRNA^{Tyr}$ genes in *Nicotiana rustica* genome [4]. Moreover, the number of the defined tRNA gene copies in a genome can be adjusted to the required supply for its expression product, as for example, alanine tRNA gene of silkworm [5]. An earlier report from our laboratory has shown that the level of $tRNA^{Gly}$ is relatively high in crude tRNA isolated from yellow lupin (*L. luteus*) seedlings and is proportional to the glycine content in the plant proteins [6]. Having these data in mind we became interested in the isolation of $tRNA^{Gly}$ gene(s), from *L. luteus* genomic library, to study their organization and expression.

In this paper we describe a $tRNA^{Gly}$ gene cluster consisting of three $tRNA^{Gly}$(GGC) genes and three $tRNA^{Gly}$(GGC) pseudogenes arranged in tandem units. This is the first complete nucleotide sequence of eucaryotic tRNA gene tandem with its 3′ and 5′ flanking regions. Moreover, this is the longest and evolutionarily the youngest, so far known tandem of genes coding for a single tRNA species. These yellow lupin $tRNA^{Gly}$ genes are efficiently transcribed in a HeLa cell transcription system giving mature $tRNA^{Gly}$. Moreover, it has been shown that at least one of them is transcribed in yellow lupin seedlings.

MATERIALS AND METHODS

Plant materials, genomic library, HeLa-cell nuclear extract and primers. Seeds of six *L. luteus* cultivars were obtained from Dr T. Nijaki (Stacja Hodowli Roślin, Wiatrowo, Poland). The *L. luteus* cv. Ventus genomic library, cloned in EMBL 3 phage was a generous gift from Professor A.B. Legocki (Institute of Bioorganic Chemistry, Poznań, Poland). Synthetic oligodeoxyribonucleotides were provided by Professor W. Markiewicz or purchased from IZOGEN (Netherlands) and Promega (U.S.A.). The following oligodeoxyribonucleotides were used in this study:

G1 (5′-GCACCAGTGGTCTAGTGG-3′) and G2 (5′-TGCACCAGCCGGGAATCG-3′) corresponding to the 5′ distal region and complementary to the 3′ distal region of the *L. luteus* $tRNA^{Gly}$ gene, respectively; Z-21 (5′-GAGCTATGATGATAATTAGCA-3′) homologous to nucleotides 9–30 beyond the 3′ end of the LtG gene coding strand (Fig. 1, 3), and nucleotides 11–32, downstream of MtG and HtG genes (Fig. 1, 3), Z-21c (5′-TGCTAATTATCATCATAGCTC-3′), complementary to Z-21, and its biotinylated version, b-Z-21c (5′-biotin-TGCTAATTATCATCATAGCTC-3′); G-23 (5′-GCCACGGTACAGACCCGGGTTCG-3′), homologous to nucleotides 36–58 of *L. luteus* cytoplasmic $tRNA^{Gly}$(GGC) [7], and biotinylated b-G-23c (5′-biotin-CGAACCCGGGTCTGTACCGTGGC-3′), complementary to the same sequence, pUC/M13 reverse sequencing primer 22-mer (5′-TCACACAGGAAACAGCTATGAC-3′) and pUC/M13 forward (–47) sequencing primer (5′-CGCCAGGGTTTTCCCAGTCACGAC-3′).

Genomic library screening. About 8×10^5 recombinant phages of the *L. luteus* genomic library were screened with a 5′-^{32}P-labelled synthetic oligodeoxyribonucleotide probe G-23. Bacteriophage plaques were transferred to Hybond N nylon filters, denatured and baked according to a standard method [8]. Hybridizations were performed essentially as described in [8].

Isolation of total RNA and nuclear DNA from L. luteus seedlings. DNA and RNA were isolated from etiolated 3-day old seedlings. Total *Lupinus* RNA was extracted from fresh tissue and ground in liquid nitrogen, essentially as described in [9] (a phenolic method). Nuclei were isolated by the Jofuku & Goldberg method [10]. Lysis of nuclear fraction was carried out in 7 M urea, 0.1% (w/v) sodium sarcosine at 50°C, for 15 min and followed by chloroform and subsequent phenol/chloroform extraction. The extracted DNA was precipitated with 0.6 vol. of isopropanol. When necessary, DNA was further purified by RNase A and/or proteinase K treatment [11].

L. luteus glycine pre-tRNA purification using a solid phase DNA probe. The biotinylated oligodeoxyribonucleotide probes b-Z-21c and b-G-23c, were used to isolate $tRNA^{Gly}$(GGC) precursors from *Lupinus* total RNA.The probes were purified by electrophoresis on 15% polyacrylamide gel in the presence of 8 M urea and were bound to

streptavidin coated magnetic beads (Magnetosphere, Promega) as described in [12]. The RNA fraction isolated using b-Z-21c probe was further purified by hybridization to b-G-23c probe. Total *L. luteus* RNA, 5 mg, dissolved in 1 ml TE buffer (10 mM Tris/HCl, 1 mM EDTA, pH 8.0) was heated for 5 min at 90°C and immediately mixed with solid phase DNA resin, suspended in 0.5 ml 6×SSC, 1 mM EDTA, 0.1% SDS and incubated for 20 min at 40°C. The hybridized solid phase DNA resin was separated from the RNA containing solution by using a magnetic separation stand (Promega). The hybridized resin was washed with 2×SSC until the absorbance of the washing solution at 260 nm was almost zero.

Then pre-tRNAs were eluted from the resin by four washes with 100 µl of TE buffer, at 90°C. The TE washing solution, containing pre-tRNAs^{Gly} was collected, mixed with 1 µl of glycogen solution (20 µg/ml) and 40 µl of 3 M sodium acetate, pH 5.5, then phenol:chloroform extracted and precipitated with 2.5 vol. of ethanol, at –20°C. All solutions were prepared in DEPC treated water.

Preparation of DNA probes for Southern hybridizations by PCR. About 10 pg of the *XbaI/Hin*dIII DNA fragment containing the MtG tRNA^{Gly} gene (Fig. 1) was used to amplify the tRNA gene sequence, probe G-72, by PCR, with 200 µM each (dCTP, dGTP and dTTP), and 20 µCi [α-^{32}P]dATP (3000 Ci/mmol), 500 ng of each the relevant primers: G1 and G2, and 2 units of Taq DNA polymerase (Amersham), using the Gene Amp 2400 PCR system (Perkin Elmer). The PCR conditions were as follows: denaturation at 94°C for 30 s, annealing at 41°C for 30 s and extension at 72°C for 30 s, repeated 20 times. Ten picograms of pMtG recombinant plasmid DNA, containing the same *XbaI/Hin*dIII fragment cloned in pGEM-3Zf(+) phagemid, was used to prepare by PCR a double stranded ^{32}P-labelled probe, designated Z-317, comprising the nucleotide sequence from nt 9 to nt 326 downstream the tRNA^{Gly} gene of the M unit. Primers: Z-21 and pUC/M13 reverse sequencing primer 22-mer were used to amplify this sequence under conditions described above, except that the extension time was 40 s and annealing temperature 38°C. The synthesized double stranded ^{32}P-labelled probes designated G-72 and Z-317 were electrophoretically separated in 4% polyacrylamide gel, eluted from gel slices for 30 min at 70°C and denatured for 5 min at 95°C in 50 mM KCl, 10 mM Tris/HCl, pH 9.0, 0.1% Triton X-100, and then added to the hybridization solution.

Southern hybridizations. Southern blot analyses were performed by a standard method [11]. *L. luteus* DNA, 5 µg, was digested by appropriate restriction enzymes and separated by electrophoresis in a 1% (w/v) agarose gel. DNA was transferred onto a Hybond N nylon membrane. Each filter was washed in 3×SSC, 0.1% SDS for several hours at 65°C and then for about 2 h in prehybridization solution: 6×SSC, 1×Denhardt's solution, 100 µg/ml denatured herring sperm DNA, 0.5% SDS, 0.05% sodium pyrophosphate at 40°C. Then the filters were incubated for 20 h at 55°C in hybridization mixture: 4×SSC, 1×Denhardt's solution, 0.05% sodium pyrophosphate, containing ^{32}P-labelled probe G-72 or Z-317. The filters which were hybridized to the probe G-72 were washed for 30 min in 0.2×SSC, 0.05% sodium pyrophosphate at 70°C, and those hybridized to the probe Z-317, at 60°C. Autoradiography was done with X-ray film for 1–5 days.

Transcription of L. luteus tRNA^{Gly} genes in HeLa-cell nuclear extract. To characterize the LtG, MtG and HtG as functional tRNA^{Gly} genes, three restriction fragments: *SalI/Hin*dIII of 756 bp which contains the LtG gene with it's 320 bp long upstream sequence, and two *XbaI/Hin*dIII fragments of 1004 bp and 1006 bp, containing MtG and HtG genes with their 5′ flanking sequences of 621 and 620 bp, respectively, were cloned in pGEM-3Zf(+) vector, and named pLtG, pMtG and pHtG. Three DNA fragments, containing the same genes, but with shorter upstream sequences, were also cloned in the same vector: *SpeI/Hin*dIII fragment, containing the LtG gene with its 59 bp upstream nucleotide sequence (named pLtGSpD) and two *XhoI/Hin*dIII restriction fragments, containing MtG and HtG genes with their 141 bp upstream sequences (pMXtG and pHXtG). To test whether the two poly(T) tracts; T_5 and

T_6 (Fig. 3) do serve as effective transcription termination signals, the *SalI/Sac*I DNA fragment containing the LtG gene, deprived of T_5 and T_6 by *Sac*I digestion (5 bp upstream T_5) and cloned in pGEM-3Zf(+) vector (named pLtGSaD), was also assayed in HeLa-cell transcription system. The *SpeI/Hin*dIII restriction fragment of 451 bp of the L unit was cloned in pGem-5Zf(+) vector. The cloned *Lupinus* DNA contains the LtG gene and its 5′ flanking sequence of 70 bp.

In vitro transcription activity of the three genes was studied in a heterologous HeLa-cell extract which has been shown to be an efficient expression system for many other plant tRNA genes [3, 13, 14]. The nuclear extract was prepared from HeLa S-3 cells according to Dignam *et al.* [15]. Transcription reactions were performed as described by Stange & Beier [16] in the presence of 4 mM $MgCl_2$ [14]. Two different nucleoside triphosphate mixes were used: a standard 10 × NTP mix, containing: 1 mM ATP, 1 mM CTP, 1 mM GTP and 2.5 μCi [α-^{32}P]UTP (3000 Ci/mmol), and the 10 × γNTP mix, containing: 1 mM CTP, 1 mM GTP, 1 mM UTP and 2.5 μCi [γ-^{32}P]ATP (3000 Ci/mmol). The γNTP mix was used to visualize only the $tRNA^{Gly}$ gene transcription products, with unprocessed 5′ ends. Transcription products were separated by electrophoresis in 8% polyacrylamide gel containing 8 M urea.

Pre-tRNA^{Gly} maturation in a cell-free wheat germ S-23 extract. Cell-free wheat germ S-23 extract was prepared from wheat embryos as described by Stange & Beier [16]. *In vitro* processing of p1 $tRNA^{Gly}$ precursor was performed in the incubation mixture containing in a total volume of 90 μl:18 μl S23 extract (30 mg protein/ml), 20 mM Tris/HCl, pH 7.4, 100 mM potassium acetate, 6 mM $Mg(OOCH_3)_2$, 80 μM spermine, 0.4% Triton X-100, 1 mM ATP, 0.1 mM CTP and 5×10^3 c.p.m. of pre-tRNA labelled during the transcription assay in HeLa-cell nuclear extract with [α-^{32}P]UTP.

Primer extension analysis. *In vitro* synthesized RNA was eluted from polyacrylamide gel [17], mixed with 20 μl of 2.5% linear polyacrylamide per 400 μl [18] and precipitated by adding 3 vol. of ethanol (−70°C) and 10 min centrifugation at room temperature, washed with 70% ethanol, dried and dissolved in 10 μl of TE buffer, containing 0.1 pmol of ^{32}P-end-labelled primer Z-21c. The mixture was heated at 90°C for 2 min. Then 10 μl of 10 mM Tris/HCl, pH 8.0, 250 mM NaCl, 1 mM EDTA was added and the mixture was incubated at 30°C for 10 min. After the primer annealing the mixture was phenol:chloroform extracted. Nucleic acids were precipitated with 3 vol. of ethanol (−70°C) and centrifuged for 10 min at room temperature. The precipitate was washed twice with 70% ethanol and dissolved in 4 μl of 5×M-MLV reaction buffer (250 mM Tris/HCl, pH 8.3, 250 mM KCl, 50 mM $MgCl_2$, 50 mM dithiothreitol, 2.5 mM spermidine). The primer was extended for 1 h at 37°C by the addition of 12.5 μl H_2O, 2 μl of 2 μM dNTPs, 1 μl RNasin (Promega, 40 u/μl) and 0.5 μl M-MLV reverse transcriptase (50 units, Promega). The reaction mixture was phenol:chloroform extracted, precipitated with 2.5 vol. of ethanol, dissolved in 95% formamide, 10 mM NaOH, 0.5% bromophenol blue, 0.5% xylene cyanole, heated at 90°C for 5 min and run on a sequencing gel. Sequencing ladders were generated using 5′-^{32}P-labelled pUC/M13 forward (−47) sequencing primer, with Taq DNA polymerase (Promega) on double stranded plasmid pGEM-3Zf(+) template.

DNA sequencing. In order to obtain overlapping sequences of the whole 7.8 kb *SalI/SalI L. luteus* DNA fragment contained in a selected EMBL 3 recombinant phage, 30 restriction fragments were subcloned in pUC and pGEM vectors and 3 series of nested deletion clones were prepared. Nucleotide sequence determination was performed by the dideoxy chain termination method [19] with Taq Track sequencing system (Promega).

RESULTS

Tandem arrangement of $tRNA^{Gly}$ genes and pseudogenes

About 80×10^3 clones from lupin genomic library were screened as described in Materials and Methods section. After twice re-

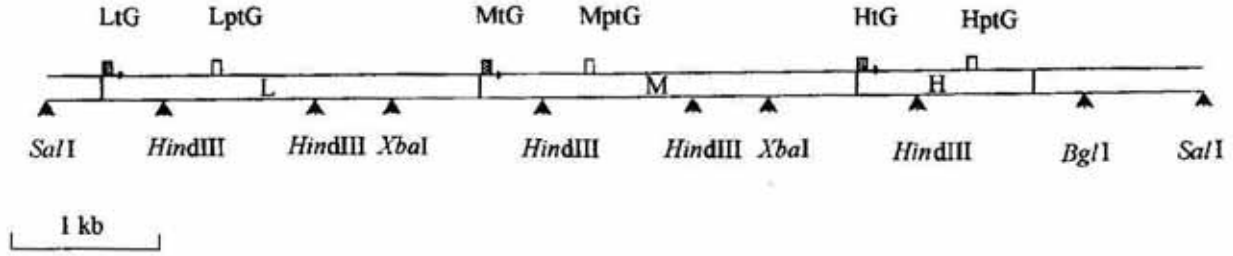

Figure 1. Schematic presentation of the 7.8 kb *SalI* DNA fragment derived from the nuclear genome of *L. luteus* cv. Ventus.

L, M and H identify the two full length and one truncated, repeating units, respectively. Stippled boxes identify the three tandemly repeated $tRNA^{Gly}$ genes, designated LtG, MtG and HtG, open boxes represent the three $tRNA^{Gly}$ pseudogenes (LptG, MptG, HptG). Horizontal arrows indicate transcribed sequences and the direction of transcription. Restriction sites are indicated by vertical arrowheads.

peated screening, from 20 recombinants giving the highest hybridization signals two were chosen. Restriction analyses and Southern hybridizations have shown that they contain the same 7.8 kb *SalI* fragments hybridizing to the G-23 oligonucleotide probe, homologous to nt 36–58 of $tRNA^{Gly}$(GGC) from yellow lupin cytoplasm. Therefore only one clone was used for further studies.

We have found that the 7.8 kb *SalI* restriction fragment contains three tandemly repeated units, tentatively named: L of 2565 bp, M of 2564 bp and H of 1212 bp, truncated at its 3′ end (Fig. 1). Analogous nucleotide sequences of the three units are homologous in 99.2%. Each unit L, M and H contains a $tRNA^{Gly}$(GGC) gene and a downstream $tRNA^{Gly}$(GGC) pseudogene having the same polarity, which are spaced by intergenic sequences of 671, 670 and 668 bp, respectively.

The tandemly repeated sequences start 8 bp upstream of the $tRNA^{Gly}$ gene of the unit L and end 394 bp downstream of the pseudogene of the unit H (Fig. 3). The mature $tRNA^{Gly}$ encoding sequences of the three genes are identical. The three $tRNA^{Gly}$(GGC) genes show 100% nucleotide sequence homology to the two other known plant nuclear $tRNA^{Gly}$ genes, from *Sorghum bicolor* [20] and *Oryza sativa* [21] and are colinear with

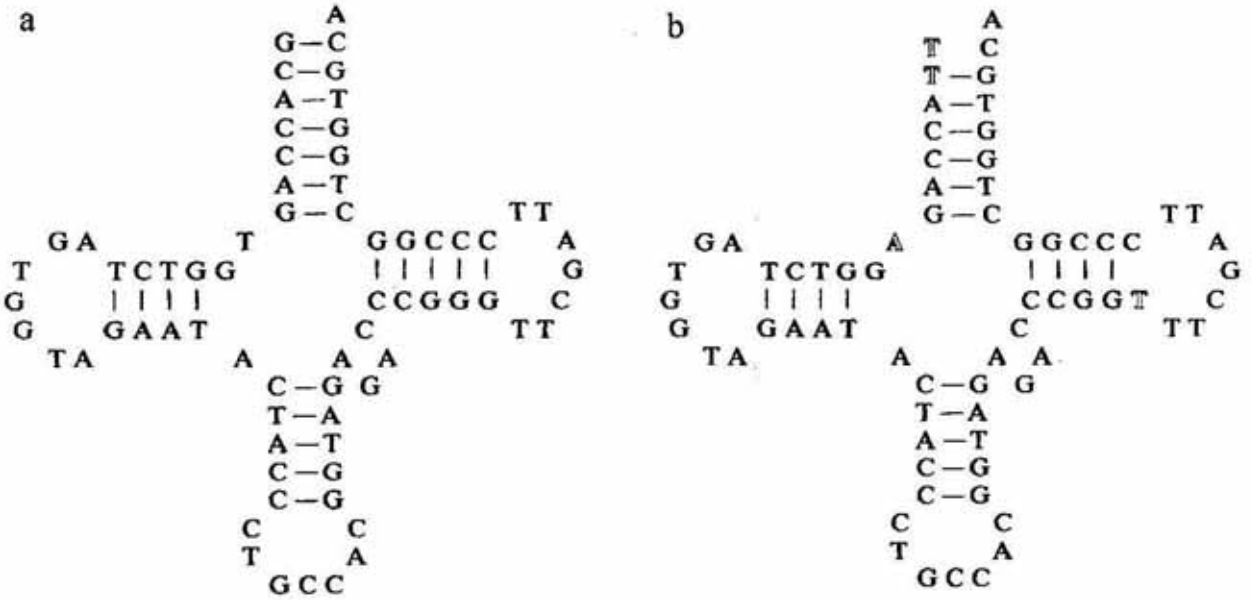

Figure 2. Secondary structures of $tRNAs^{Gly}$ encoded by nuclear $tRNA^{Gly}$(GGC) genes designated LtG, MtG, HtG (a), and $tRNA^{Gly}$(GGC) pseudogenes designated LptG, MptG, HptG (b), which differ from genes by four nucleotide exchanges, indicated by outlined letters.

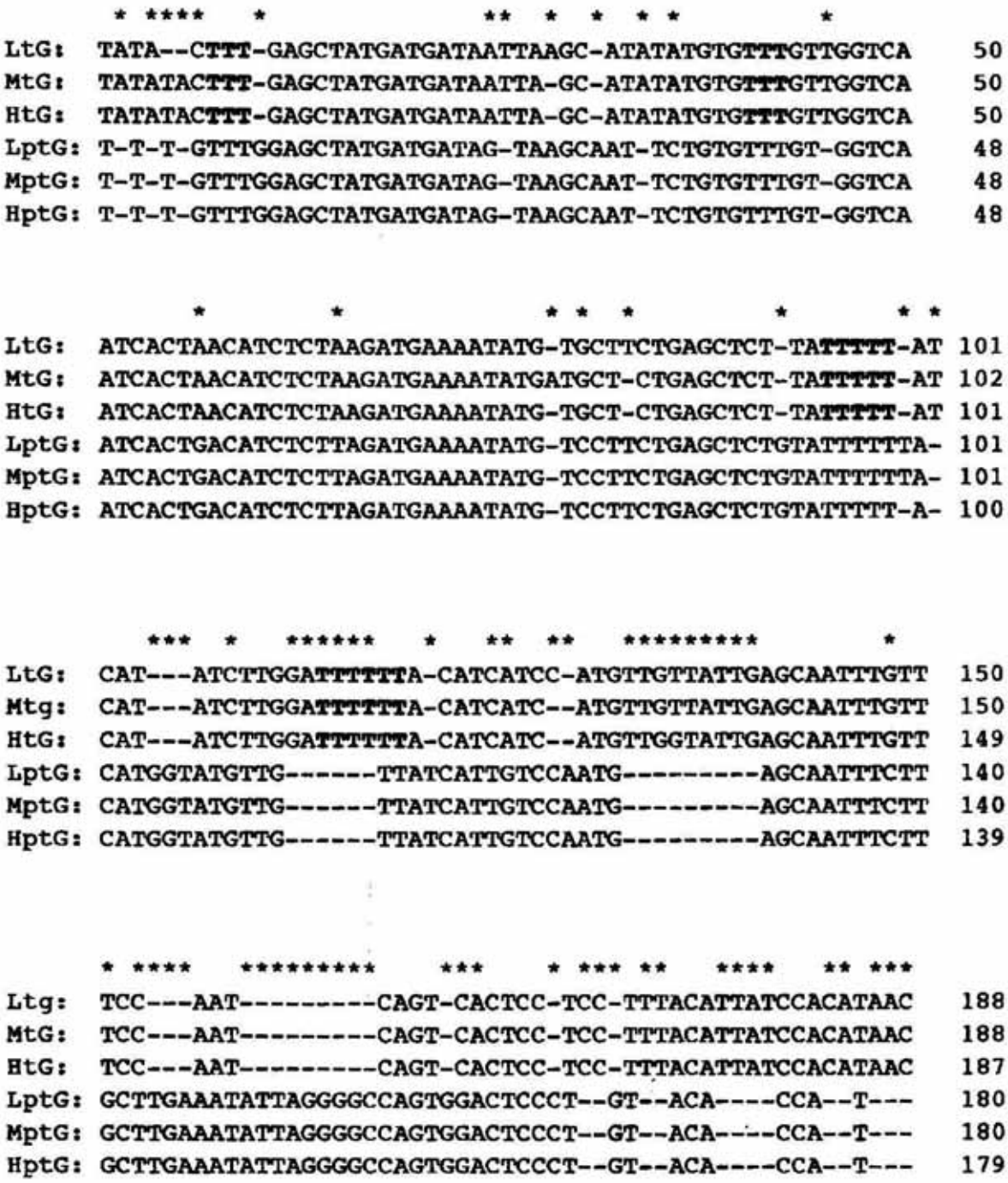

Figure 3. Nucleotide sequences (coding strands) of 3′ flanking regions of the three tRNA^{Gly} genes and these pseudogenes.

Sequences are aligned to achieve maximal homology. Positions where one nucleotide differs in at least one sequence are indicated by asterisks above the LtG 3′ flanking region sequence. Insertions or deletions appear as gaps. The nucleotide sequence (gene) name is indicated at the left. The nucleotide numbering starts just beyond the 3′ end of each mature tRNA or pseudo-tRNA encoding sequence. RNA polymerase III termination signals appear in bold print.

the published *L. luteus* cytoplasmic tRNA^{Gly}(GGC) [7]. Similarly, tRNA^{Gly}(GGC) pseudogenes in this tandem are identical with each other and differ from the genes in four positions. C1→T1 and G2→T2 nucleotide exchanges disrupt at least the first base pair in the acceptor stem. Two transversions, T8→A8 and G51→T51 extinguish the first invariant position of A-box and the second position of B-box of the internal polymerase III promoter, respectively [22]. G51→T51 exchange also disrupts the last base pair of the TΨC arm (Fig. 2). The tRNA^{Gly}(GGC) genes and pseudogenes are surrounded by A:T-rich sequences (63% A:T) and their 5′ flanking sequences contain TATA-like elements. The CAA triplet occurring just upstream of many different plant tRNA genes [23] was also found in 5′-adjacent regions of three tRNA^{Gly} genes, but not upstream of tRNA^{Gly} pseudogenes. In contrary to the different 5′ flanking sequences of the tRNA^{Gly} genes and pseudogenes, their 3′-adjacent regions of about 100 bp (Fig. 3) are highly homologous (over 80% homology). These homologous regions roughly comprise the 3′ trailer sequence of the longest primary tRNA^{Gly} precursors obtained during transcription of the three genes in HeLa-cell extract (Fig. 6). Downstream sequences of genes and pseudogenes contain only a few runs of short T-clusters. The first ones of five and six consecutive T residues are separated from 3′ end of the mature tRNA encoding sequences by at least 94 bp spacers (Fig. 3).

Southern hybridizations of *L. luteus* DNA restriction fragments with G-72 and Z-317 DNA probes revealed that these downstream sequences, common to the sequenced glycine tRNA genes and similar sequences located downstream of pseudogenes, are unique for the studied tandem. They seem not to appear in close proximity to other tRNA^{Gly}(GGC) genes, or possible tRNA^{Gly}(GGC)-like sequences in yellow lupin nuclear genome (Fig. 4, 5). It has also been shown that this tandem occurs as a single copy per lupin haploid genome (Fig. 4 and 5).

To estimate the copy number of tRNA^{Gly}(GGC) genes in the *L. luteus* genome, Southern blot analysis was performed. DNA was hydrolysed with five pairs of restriction enzymes which do not cut inside the tRNA^{Gly}(GGC) gene coding sequence, and hybridized to G-72 probe, which is gene specific. From 22 to 29 hybridizing restriction fragments were detected (Fig. 5A). In each lane, three of them giving as a rule the strongest signals, derive from the 3 units of the studied tandem (Fig. 5B). These fragments contain tRNA^{Gly}(GGC) gene and its pseudogene (compare with Fig. 1). This implies that there

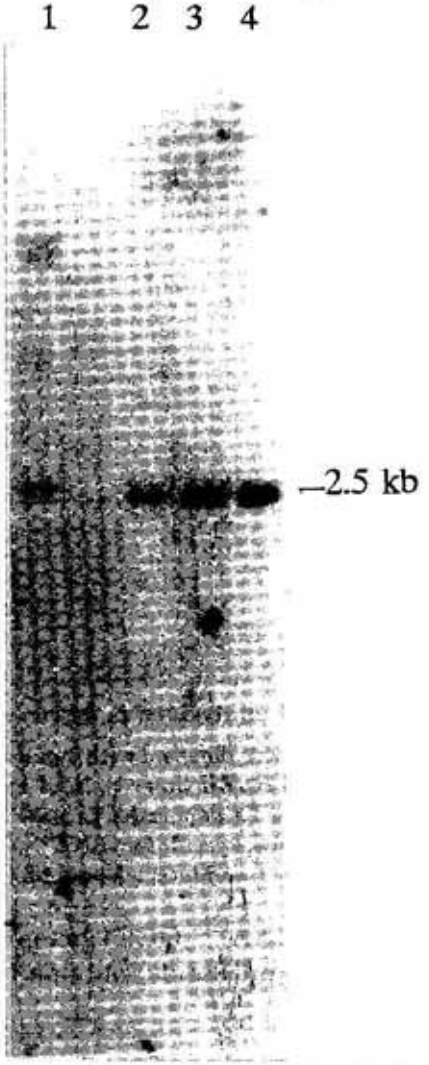

Figure 4. Southern blot analysis of *L. luteus* cv. Ventus, nuclear DNA.

Lupinus DNA, 5 μg, (lane 1) and 25, 50 and 75 pg of pMtG DNA (lanes 2, 3, 4), respectively, were digested with *Xba*I restriction enzyme, separated in 0.9% agarose gel, transferred onto nylon membrane and hybridized to ^{32}P-labelled Z-317 probe; 25 pg of pMtG DNA corresponds to the amount of *Lupinus* 2565 bp *Xba*I restriction fragment, analogous to the plasmid insert, in 5 μg of nuclear DNA from *L. luteus* cv. Ventus (2.212 pg DNA per diploid genome) when it appears in one copy per haploid genome (M. Olszewska, personal communication).

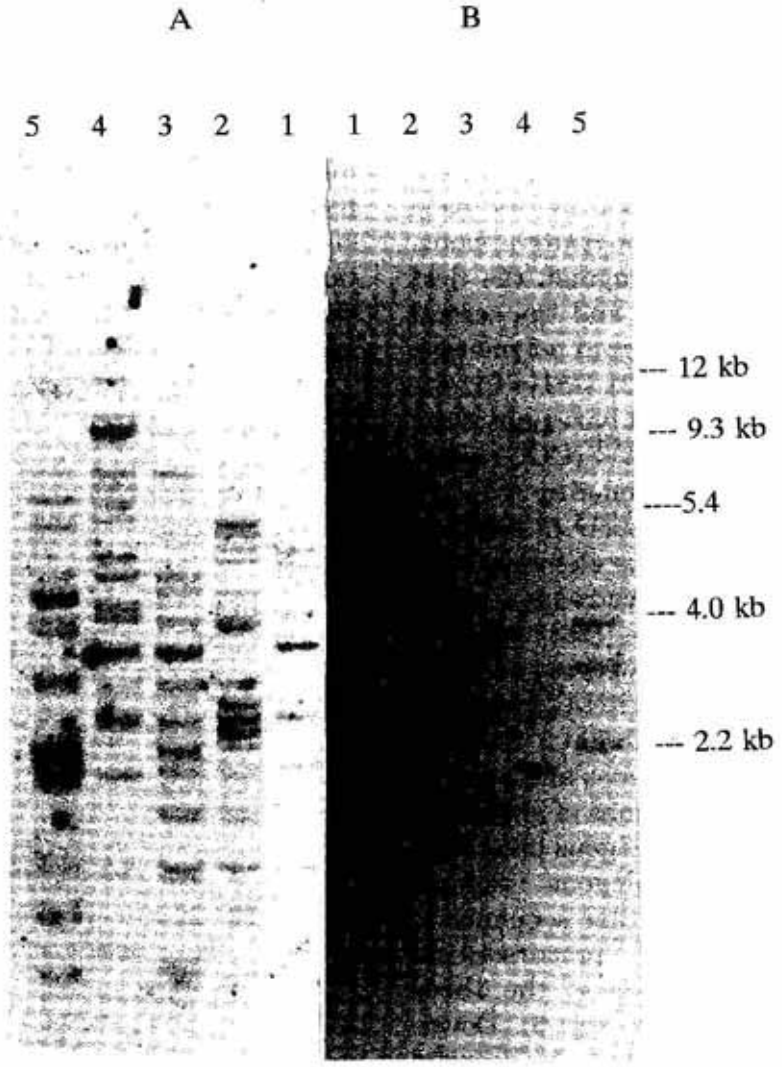

Figure 5. DNA blot analysis of *L. luteus* cv. Ventus nuclear DNA.

Lupinus DNA (5 µg) was digested with 5 pairs of restriction enzymes which do not cut inside tRNA^{Gly}(GGC) coding sequences; *Xba*I and *Bst*EII (lane 1), *Nde*I and *Bst*EII (lane 2), *Nde*I and *Bgl*II (lane 3), *Xba*I and *Bgl*II (lane 4), *Xba*I and *Nde*I (lane 5), separated in 0.9% agarose gel, transferred onto nylon membrane and hybridized to ^{32}P-labelled G-72 probe, specific for tRNA^{Gly}(GGC) (panel A), and to Z-317 probe (B), specific for 3' flanking regions of LtG, MtG and HtG (see Materials and Methods, and Fig. 1).

are at least 26 tRNA^{Gly}(GGC) genes or similar sequences (faint bands) outside the tandem.

Transcription of the *Lupinus* tRNA^{Gly}-(GGC) genes in HeLa-cell nuclear extract and maturation of the two main LtG gene primary transcripts in a wheat germ S-23 extract

To characterize LtG, MtG and HtG as functional tRNA^{Gly} genes their *in vitro* transcription has been analysed in a heterologous HeLa-cell nuclear extract. During *in vitro* transcription of the three individually subcloned genes, about 7 transcripts (or RNA fractions) are produced (Fig. 6), also when α-amanitin (5 µg/ml) is added to the reaction mixture (not shown); this indicates that the observed transcripts are synthesized by RNA polymerase III. All three genes are transcribed in this system with a similar efficiency, comparable to that of tRNA^{Tyr} gene from wheat [24] (Fig. 6), irrespective of their different upstream sequences (the tandemly repeated sequence starts 8 bp upstream of the LtG gene). On the other hand, the removal of the MtG gene upstream nucleotide sequence (from –621 to –141) markedly lowers the efficiency of its transcription, while almost the same upstream sequence can be removed from analogous pHtG template with

Figure 6. Comparison of *in vitro* transcription patterns of the three tRNA^{Gly}(GGC) genes in HeLa-cell extract.

Products were separated on 8% denaturing polyacrylamide gels. Recombinant plasmid DNAs, containing LtG, MtG and HtG genes, were transcribed in HeLa-cell *in vitro* transcription system, as described in Materials and Methods section. Lanes 1–8 contain transcripts produced from: (1) pHtG, (2) pHXtG, (3) pMtG, (4) pMXtG, (5) pLtG, (6) pLSptG, (7) pLSatG, (8) ptTyRsaII [24]. The main bands of the transcription and processing products of LtG, MtG and HtG genes are designated p1 to p7; p4 and p6 were not observed among the MtG gene transcription products.

no significant effect (Fig. 6). Thus it seems that this change of MtG *in vitro* transcriptional activity may be due rather to an overall conformational change of the plasmid template than to any specific upstream sequence [21]. Furthermore, the rate of processing of MtG gene transcription products is faster in HeLa-cell extract, and slightly different than of the products synthesized on the two remaining tRNA^{Gly} genes. To test whether the T_5 and T_6 tracts (Fig. 3) are indeed responsible for transcription termination of the longest transcripts synthesized on the three tRNA^{Gly}(GGC) gene templates, an *in vitro* transcription assay was also performed on pLtGSaD template (see Materials and Methods section), deprived of these two RNA polymerase III transcription termination signals. As expected, a longer primary transcript than the original p1 (180 nt) was obtained (Fig. 6, lane 7).

The two transcription products: p1 (180 nt) and p5 (86 nt), appear first during the time course of the transcription experiment (Fig. 7), thus they seem to be the products of transcription termination rather, than of further processing. The *in vitro* transcription carried out in the presence of [γ-^{32}P]ATP has shown that both fractions, p1 and p5, contain adenosine triphosphate as the first base (Fig. 8). These, and primer extension data together (see below), imply that the two poly(T) tracts: the proximal CTTTG, located 5 bp downstream of LtG gene (7 bp downstream of MtG and HtG), and the distal T_5, starting 94 bp beyond the 3′ end of LtG and HtG (95 bp downstream MtG), respectively, serve as

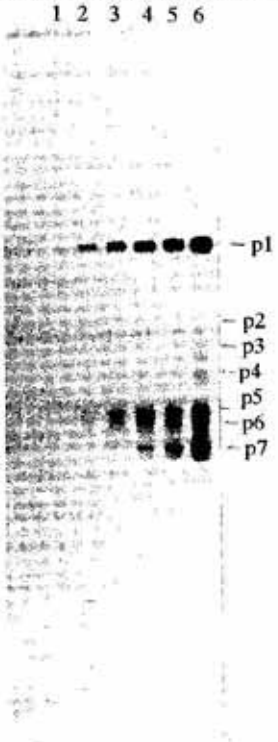

Figure 7. Time course of the transcription of pLtG gene.

The pLtG DNA was incubated in HeLa-cell nuclear extract for: (1) 15 min, (2) 30 min, (3) 60 min, (4) 90 min, (5) 120 min and (6) 180 min. The products of transcription and processing were separated on 8% polyacrylamide, 8 M urea gel. p1-p7 designate the transcription and processing products.

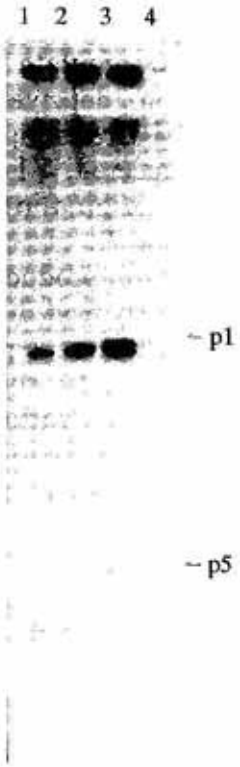

Figure 8. Fractionation of the *in vitro* transcription products labelled with [γ-^{32}P]ATP during transcription in HeLa-cell extract.

In this case the transcription was carried out in the presence of [γ-^{32}P]ATP instead of [α-^{32}P]UTP, to visualize only those transcripts which have unprocessed 5′ ends. Lane 4: 60 min transcription without the template carried out as a control; lanes: 1, 2 and 3 contain pLtG transcripts obtained during 30, 60 and 90 min of incubation at 30°C, respectively. The bands corresponding to p1 and p5 are indicated on the right.

main transcription termination signals in this system.

The two primary transcripts (p1 and p5) are also efficiently processed in wheat germ S-23 extract to the mature size tRNA^{Gly} (Fig. 9). Other RNA fractions, probably intermediates of this heterologous transcription and processing system, were not further characterized.

Primer extension analysis of transcription initiation site of transcript p1

The primer extension analysis was performed using the 5′-^{32}P-labelled oligodeoxynucleotide Z-21c complementary to nucleotides located 9–30 bp downstream of the MtG coding strand. The longest primary transcript p1 (180 nt), obtained after 90 min transcription of pMtG template in HeLa-cell

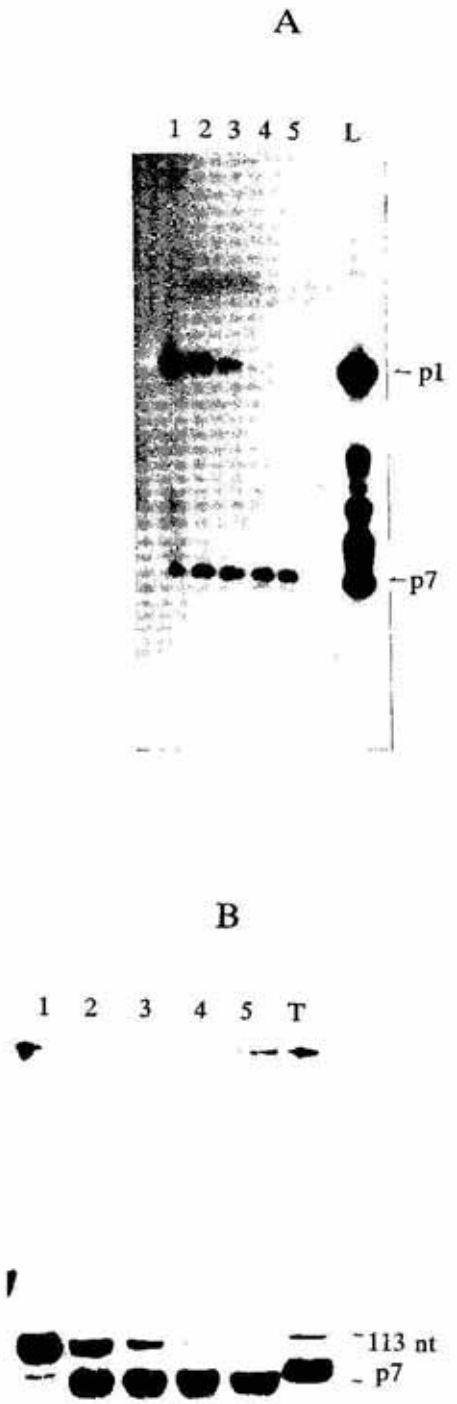

Figure 9. *In vitro* maturation of LtG gene primary transcripts, p1 and p5 in wheat germ S-23 extract.

The primary transcripts were recovered from a preparative gel and incubated in a wheat germ extract. Aliquots of the reaction mixture containing p1 (panel A) and p5 (panel B) were removed after (1) 0, (2) 30, (3) 60, (4) 90 and (5) 120 min of incubation at 30°C, and separated on 8% polyacrylamide, 8 M urea gel. Lane T contains the *Arabidopsis* tRNA^{Tyr} 113 nt precursor incubated in the same conditions for 60 min; lane L, HeLa-cell transcription products of the LtG tRNA^{Gly}(GGC) gene.

nuclear extract, was used as a template. This analysis revealed that the initiation of transcription occurs at the A residue positioned in a canonical R-Y-R sequence 5 bp upstream of the MtG gene (Fig. 10, PE-1). Since the transcription initiation site is located within the 8 bp long, MtG 5′-adjacent region common to all three tRNA^{Gly} genes of the tandem giving nearly the same *in vitro* transcription products, it can be assumed that the transcription of LtG and HtG genes initiates at analogous sites.

Figure 10. Analysis of transcription initiation site of the MtG gene p1 (180 nt) transcript, by primer extension.

The tRNA^{Gly}(CCG) precursor was recovered from a preparative gel, hybridized to ^{32}P-end-labelled primer Z-21c and reverse transcribed with M-MLV reverse transcriptase as described in Materials and Methods section. Primer extension products were fractionated in 8% sequencing gel against pGEM -3Zf(+) sequencing ladders, generated using 5′-^{32}P-labelled pUC/M13 forward (–47) sequencing primer. Lanes: A, C, G, T designate sequencing ladders, lane 1 contains primer extension products, PE-1 and PE-2, indicated on the right.

The second, by 5 nt shorter, faint band of an unexpectedly observed primer extension product (Fig. 10, PE-2) corresponds to the mature 5′ end of the p1 precursor.

Purification of glycine tRNA^{Gly}(GGC) precursors originating from the tandemly organized tRNA^{Gly} gene(s), from *L. luteus* seedlings

In order to test whether tRNA^{Gly} genes of the studied tandem are transcriptionally active *in vivo*, their precursors were purified from *Lupinus* RNA by two consecutive hybridizations to oligodeoxyribonucleotide probes bound to magnetic beads. Total RNA, 5 mg, isolated from 3 day old, etiolated, *L. luteus* seedlings was hybridized to b-Z-21c probe, complementary to nucleotides 9–30 beyond the 3′ end of MtG gene coding strand. This RNA fraction was then hybridized to b-G-23c probe, complementary to nucleotides 36–58 of *L. luteus* cytoplasmic tRNA^{Gly}-(GGC). The purified and 5′-end-^{32}P-labelled RNA was separated in 8% polyacrylamide gel containing 8 M urea (Fig. 11). The RNA fraction hybridizing to b-Z-21c and G-23c probes has the same length (180 nt) as the p1 primary transcript produced from the three tRNA^{Gly}(GGC) genes in a HeLa-cell transcription system.

These data imply that at least one of the three tRNA^{Gly} genes is transcribed *in vivo* in yellow lupin seedlings to give a precursor of the same size as that obtained in HeLa-cell extract, probably terminated at the T_5, RNA polymerase III transcription termination signal. *In vivo* transcription of these three genes (or one of them) giving a pre-tRNA(s) with unique 3′ trailer, raises the possibility of further, direct studies on the expression of a higher plant, single, nuclear tRNA gene, belonging to a multigene family.

DISCUSSION

In this report it has been shown that three tRNA^{Gly}(GGC) genes and tRNA^{Gly}(GGC) pseudogenes of *L. luteus* genome are organ-

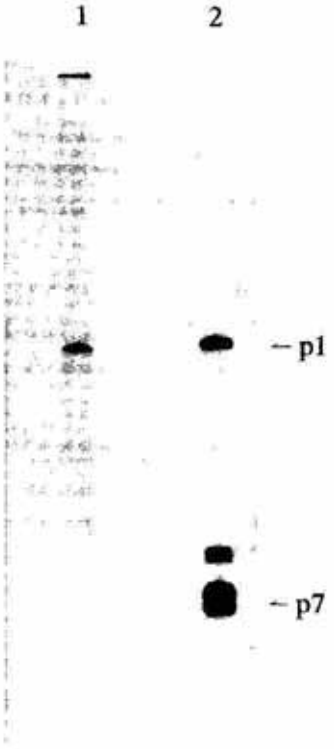

Figure 11. Electrophoretic separation of pLtG transcripts from *L. luteus* seedlings (lane 1), obtained during 90 min incubation in HeLa-cell extract (lane 2) and 5′-^{32}P-labelled RNA fraction, purified by hybridization to Z-21c and G-23c oligonucleotide probes.

ized in three tandemly repeated units of 2565, 2564 and 1212 bp. Each unit contains an identical pair of an tRNA^{Gly}(GGC) gene and a downstream of tRNA^{Gly}(GGC) pseudogene, both having the same polarity. The tandemly repeated sequence starts 8 bp upstream of the tRNA^{Gly} gene of the first (2565 bp) unit and ends 394 bp downstream of the tRNA^{Gly} pseudogene of the last (1212 bp) unit, truncated at its 3′ end. The nucleotide sequences of 5′ and 3′ flanking sequences of the tandem were determined and it has been shown that this tandem occurs in one copy per haploid genome of *L. luteus*.

The tRNA^{Gly}(GGC) gene and the pseudogene, contained in each of the three tandemly repeated units are, except their 3′-adjacent sequences of about 100 bp, surrounded by different nucleotide sequences. Therefore, it can be assumed that this tandem appeared by a two step process: a defective tRNA^{Gly} gene duplication, either by an unequal crossing-over with the loss of the pseudogene 5′ flanking sequence, or by reverse transcription of a partly processed pre-tRNA^{Gly}, followed by a subsequent tandem amplification of the 2.5 kb DNA fragment which contained a preexisting pair of a tRNA^{Gly} gene and its pseudogene. The 99.2% nucleotide sequence homology of the three units implies that this amplification event could take place a short time ago. One can estimate the age of this tandem to be about 800000 years [2, 25]. Furthermore, since the tandemly repeated sequence starts in close proximity to the putative transcription initiation site of the LtG gene (Fig. 1), it can be postulated that this tandem amplification event could have been transcription dependent [26]. Hence it seems interesting which one of the three tRNA^{Gly}(GGC) genes of this tandem is the most ancient. The nucleotide sequence comparison of 3′ flanking regions of the tRNA^{Gly}(GGC) genes (Fig. 3) reveals that the MtG 3′ flanking region shows the most conservative structure among the three tRNA^{Gly} gene 3′-adjacent sequences. On the other hand, the LtG 3′ flanking region exhibits the greatest similarity to analogous 3′-adjacent sequences of the three pseudogenes (Fig. 3). Thus, the LtG gene, which is located just at the 5′ end of the tandem seems to be the best candidate for being considered as the ancestor of the first pseudogene and the most ancient tRNA^{Gly} gene of this tandem. Since it has been shown that the at least 26 remaining tRNA^{Gly}(GGC) genes of *L. luteus* genome have different 3′ flanking sequences than the three tRNA^{Gly}(GGC) genes of the tandem, the first tRNA^{Gly}(GGC) pseudogene could have indeed appeared by a defective duplication of the upstream tRNA^{Gly} gene (LtG).

The pseudogene appearance by the tandem duplication seems rather unlikely because the upstream sequences of the gene and the pseudogene are different. However the pseudogene could have appeared by a recombination event, but strikingly, homologous sequences of the genes and the pseudogenes roughly comprise the regions of tRNA^{Gly}(GGC) genes, which are transcribed as well *in vitro* as in *in vivo*. This may imply that the pseudogene appeared rather by reintegration of a reverse transcribed p1, or a longer pre-tRNA^{Gly} with a mature 5′ end, possibly a transcript of the ancestral upstream tRNA^{Gly}(GGC) gene.

The nucleotide sequence and primer extension data support to some extent this hypothesis. It has been shown that a fraction of p1, obtained in HeLa-cell extract, has a mature 5′ end (Fig. 10). Furthermore, the upstream nucleotide sequence of the pseudogene together with its two first nucleotides differs from the analogous region of the gene. The lack of the two first nucleotides of the pseudogene could be ultimately explained to be an effect of retroviral integrase processing of reverse transcribed p1, as a donor molecule, having a mature 5′ end [27].

The arrangement of the $tRNA^{Gly}$(GGC) gene and its downstream pseudogene is reminiscent of human nuclear $tRNA^{Gly}$(GGC) gene and a $tRNA^{Gly}$(GGC) pseudogene [28]. Three aspects of this similarity should be considered. First, the arrangement of the gene and the pseudogene is almost identical in both cases, namely, the pseudogene having the same orientation is located about 0.7 kb downstream of the gene and the intergenic spacer contains very similar poly(A)/poly(T) tracts. Secondly, in both cases only the transcribed sequence is duplicated. Thirdly, the main changes between the duplicated sequences locate just beyond the 3′ end of the $tRNA^{Gly}$ encoding sequence. In the case of human pair of the $tRNA^{Gly}$(GGC) gene and pseudogene, reverse transcription mediated integration of a downstream pseudogene was also discussed [28], since the three different types of repetitive sequence elements in higher eucaryotes may be retroposons that were derived from several tRNA molecules, and two of these elements have a great similarity to human $tRNA^{Gly}$(GGC) [29]. This human $tRNA^{Gly}$(GGC) is also quite homologous to the identifier sequence elements located in the intervening sequences of brain specific genes in rats [28, 30]. Another example of such an arrangement of the tRNA gene and its pseudogene is a pair of $tRNA^{His}$ gene and its downstream, defectively duplicated, transcribed sequence found in *Drosophila melanogaster* [31]. These three examples of strikingly similar arrangement of tRNA genes and pseudogenes in nuclear genomes of those different organisms are either purely accidental or reflect a similar route of tRNA gene duplication in higher eucaryotes.

Tandem organization of tRNA genes appears in several organisms examined so far. However, there are significant differences both in gene composition and in the number of repeating units. *Xenopus laevis* genome contains a set of tandemly, about 150 times, repeated units each of which includes eight different tRNA genes [32]. A cluster of at least 20 individual units with two $tRNA^{Tyr}$ and one $tRNA^{Ser}$ genes appears in *Arabidopsis thaliana* [2].

The only known, another example of tandemly organized $tRNA^{Gly}$ genes appears in *Escherichia coli* genome, where three repeated $tRNA^{Gly}$(GGC) genes having the same orientation are separated by 36 and 35 bp intergenic spacers of almost identical nucleotide sequence [33].

5′ Flanking sequences of the studied genes contain motifs that are considered as extragenic control elements for transcription. They include TATA-like boxes and short TA base pair blocks surrounding the start sites of transcription. Quite recently Ulmasov & Folk [34] reported that upstream TATA-like elements are required for efficient and accurate expression of plant tRNA genes. LtG has a TATAATAA sequence at the position –27 with respect to the mature tRNA coding sequence. The two following genes MtG and HtG have a motif TAAGATAA at the same position and an additional sequence TATAAGAA at the position –48. Moreover, all the genes contain short TA blocks at –10 (or –11), located near the presumable start sites of their transcription, which were shown to improve the transcription efficiency and position the transcription initiation site of tRNA genes in yeasts [35].

In the present work it was shown that all studied $tRNA^{Gly}$ genes are transcribed *in vitro* in a HeLa-cell system. Among the synthesis products two primary transcripts were identified, indicating the existence of two main transcription termination sites. A T_5 sequence located 94 bp downstream of the mature $tRNA^{Gly}$(GGC) encoding sequence is surely a termination signal, giving the 180 nt long and the main transcript. The second, weaker termination signal would be a motif CTTTG located 5 and 7 bp downstream of the LtG and the two MtG and HtG genes, respec-

tively, responsible for the synthesis of the shorter, 86 bp primary transcripts.

Usually the number of tRNA genes, specific for a single amino acid in genomes of eucaryotes is relatively high. Glycine tRNA genes of *L. luteus* appear in at least 29 loci. The determination whether an individual gene or a group of genes are expressed *in vivo* is quite a problem. Ulmasov & Folk [34] have described a method for the determination of tRNA gene transcription ability *in vivo*, that relies upon translational suppression of nonsense codons in a reporter gene. The examined tRNA gene is mutagenized to alter it's anticodon, and the test is carried out in carrot protoplasts. Our way to demonstrate expression *in vivo* of studied $tRNA^{Gly}$ genes was more direct. We have shown that at least one of the three genes is in fact transcribed in lupin, namely in developing seedlings. We took advantage of particular structural features of the gene family. First of all 3′ flanking regions of the three $tRNA^{Gly}$ genes have appeared unique among all tRNA(GGC) genes in yellow lupin genome. Moreover, their primary transcripts are long because the main transcription termination signals T_5 and T_6 are distributed far away from 3′ ends of these genes, as we have shown in *in vitro* experiments. These properties have enabled us to isolate primary transcripts directly from crude RNA preparations by hybridization to magnetic carrier with an attached oligodeoxynucleotide, complementary to the unique 3′ trailer of these pre-tRNAs. The isolated RNA fraction has the size characteristic of the main primary transcript produced *in vitro* in HeLa-cell extract, and hybridizes to oligodeoxynucleotides complementary to both 3′ flanking sequence and mature $tRNA^{Gly}$ sequence.

Transcriptional activity of at least one of the three $tRNA^{Gly}$ genes raises the possibility to compare its (their) transcription efficiency in different organs or tissues of yellow lupine. Since the long primary transcripts of the three $tRNA^{Gly}$ genes, likely transcribed in lupin, differ from each other by point mutations, it is also possible to study the transcriptional activity of the three individual tRNA genes organized in a tandemly repeated sequence *in vitro* and *in vivo*.

We thank Professor A.B. Legocki for providing the *L. luteus* genomic library, Professor W. Markiewicz for oligodeoxynucleotide synthesis, Dr T. Nijaki for *L. luteus* seeds and Dr M. Sikorski for useful practical advice.

REFERENCES

1. Sprintzl, M., Steegborn, C., Hübel, F. & Steinberg, S. (1996) Compilation of tRNA sequences and sequences of tRNA genes. *Nucleic Acids Res.* **24**, 68–72.
2. Beier, D., Stange, N., Gross, H.J. & Beier, H. (1991) Nuclear $tRNA^{Tyr}$ genes are highly amplified at a single chromosomal site in the genome of *Arabidopsis thaliana*. *Mol. Gen. Genet.* **225**, 72–80.
3. Urban, C., Smith, K.N. & Beier, H. (1996) Nucleotide sequence of nuclear $tRNA^{Cys}$ genes from *Nicotiana* and *Arabidopsis* and expression in HeLa cell extract. *Plant Mol. Biol.* **32**, 549–552.
4. Stange, N. & Beier, H. (1986) A gene for the major cytoplasmic $tRNA^{Tyr}$ from *Nicotiana rustica* contains 13 nucleotides long intron. *Nucleic Acids Res.* **14**, 8691.
5. Underwood, D.C., Knickerbocker, H., Gardner, G., Cundliffe, D.P. & Sprague, U.K. (1988) Silk gland-specific $tRNA^{Ala}$ genes are highly clustered in the silkworm genome. *Mol. Cell Biol.* **8**, 5505–5512.
6. Kędzierski, W. (1980) Correlation between tRNA population and amino acid composition of proteins in plants. *Plant Sci. Lett.* **21**, 15–21.
7. Barciszewska, H., Barciszewski, J., Kuchino, J.G. & Nishimura, S. (1986) The nucleotide sequence of two glycine tRNAs from *L. luteus* seeds. *Nucleic Acids Res.* **14**, 9525.
8. Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A. & Struhl, K. (1987) *Current Protocols in Molecular Biology;* pp. 6.4.1–6.4.10, John Wiley & Sons, Inc.

9. Ainsworth, C., Beynon, J. & Buchanan-Wollaston, V. (1991) *Techniques in Plant Molecular Biology*; pp. 66–79, Wye College, Ashford.

10. Jofuku, K.D. & Goldberg, R.B. (1988) Analysis of plant gene structure; in *Plant Molecular Biology: A Practical Approach* (Shaw, C.W., ed.) pp. 37–66, IRL Press.

11. Sambrook, J., Fritsch, E.F. & Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual;* 2nd edn., Cold Spring Harbor Laboratory Press.

12. Wakita, K., Watanabe, Y., Yokogawa, T., Kumazawa, Y., Nakamura, S., Ueda, T., Watanabe, K. & Nishikawa, K. (1994) Higher order structure of bovine mitochondrial $tRNA^{Phe}$ lacking the "conserved" GG and TΨC sequence as inferred by enzymatic and chemical probing. *Nucleic Acids Res.* **22**, 347–353.

13. Beier, D. & Beier, H. (1992) Expression of variant nuclear *Arabidopsis* $tRNA^{Ser}$ genes and pre-tRNA maturation in HeLa, yeast and wheat germ extracts. *Mol. Gen. Genet.* **233,** 201–208.

14. Stange, N., Beier, D. & Beier, H. (1991) Expression of nuclear $tRNA^{Tyr}$ genes from *Arabidopsis thaliana* in HeLa cell and wheat germ extracts. *Plant Mol. Biol.* **24**, 865–875.

15. Dignam, J.D., Lebowitz, R.M. & Roeder, R.G. (1983) Accurate transcription initiation by polymerase II in a soluble extract from isolated mammalian nuclei. *Nucleic Acids Res.* **11**, 1475–1489.

16. Stange, N. & Beier, H. (1987) A cell-free plant extract for accurate pre-tRNA processing, splicing and modification. *EMBO J.* **6**, 2811–2818.

17. Reyes, V.M. & Abelson, J.N. (1989) *In vitro* synthesis of end-mature, intron containing transfer RNAs. *Methods Enzymol.* **180A**, 63–69.

18. Gaillard, C. & Strauss, F. (1990) Ethanol precipitation of DNA with linear polyacrylamide as a carrier. *Nucleic Acids Res.* **18**, 378.

19. Sanger, F., Nickeln, S. & Coulson, A.R. (1997) DNA sequencing with chain terminating inhibitors. *Proc. Natl. Acad. Sci. U.S.A.* **74**, 5463–5467.

20. Ma, D.P. & Yang, Y.W. (1988) Nucleotide sequence of a $tRNA^{Gly}$ gene from *Sorghum bicolor*. *Nucleic Acids Res.* **16**, 3588.

21. Reddy, P.S. & Padayatty, J.D. (1988) Effect of 5′ flanking sequences and changes in 5′ internal control region on the transcription of rice $tRNA^{Gly}_{GCC}$ gene. *Plant Mol. Biol.* **11**, 575–583.

22. Clarkson, S.G. (1983) Transfer RNA genes; in *Eucaryotic Genes: their Structure, Activity and Regulation* (McLean, N., Gregory, S.P. & Flavell, R.A., eds.) pp. 239–261, Butterworth, London.

23. Choisne, N., Carneiro, V., Pelletier, G. & Small, I. (1994) Effect of 5′ flanking sequence on expression of a leucine tRNA gene of *Phaseolus vulgaris* in tobacco protoplasts; in *Abstracts of the 14th International Congress of Plant Molecular Biology,* p. 475, Amsterdam.

24. Szweykowska-Kulińska, Z. (1992) Two intron-containing pre-$tRNAs^{Tyr}$ from *Triticum aestivum* are efficiently processed and spliced in homologous cell-free extract. *Acta Biochim. Polon.* **39**, 283–288.

25. Kimura, M. (1977) Preponderance of synonymous changes as evidence for the neutral theory of molecular evolution. *Nature* **267**, 275–276.

26. DePamphilis, M.L. (1993) Origins of DNA replication that function in eucaryotic cells. *Curr. Opinion Cell Biol.* **5**, 434–441.

27. Vora, A.C., McCord, M., Fitzgerald, M.L., Inman, R.B. & Grandgenett, P. (1994) Efficient concerted integration of retrovirus-like DNA *in vitro* by avian myeloblastis virus integrase. *Nucleic Acids Res.* **22**, 4454–4461.

28. Pirtle, I.L., Shortridge, R.M. & Pirtle, M. (1986) Nucleotide sequence and transcription of a human glycine $tRNA_{GCC}$ gene and a nearby pseudogene. *Gene* **43**, 155–167.

29. Lawrence, C.B., McDonnel, D.P. & Ramsey, W.J. (1985) Analysis of repetitive sequence elements containing tRNA-like sequences. *Nucleic Acids Res.* **13**, 4239–4252.

30. Sutcliffe, J.G., Milner, R.J., Gottesfeld, M.J. & Lerner, R.A. (1984) Identifier sequences are transcribed specifically in brain. *Nature* **308**, 237–241.

31. Cooley, L., Schaack, J., Johnson Burke, D., Thomas, B. & Söll, D. (1984) Transcription factor binding is limited by the 5′-flanking regions of a *Drosophila* $tRNA^{His}$ gene and a $tRNA^{His}$ pseudogene. *Mol. Cell Biol.* **4**, 2714–2722.

32. Müller, F., Clarkson, S.G. & Galas, D.J. (1987) Sequence of a 3.18 kb tandem repeat of *Xenopus laevis* DNA containing 8 tRNA genes. *Nucleic Acids Res.* **13**, 4239–4252.

33. Komine, Y., Adachi, T., Inokuchi, H. & Ozeki, H. (1990) Genomic organization and physical mapping of the transfer RNA genes in *Escherichia coli J. Mol. Biol.* **212**, 579–598.

34. Ulmasov, B. & Folk, W. (1995) Analysis of the role of 5′ flanking sequence elements upon *in vivo* expression of the plant $tRNA^{Trp}$ genes. *Plant Cell* **7**, 1723–1734.

35. Fruscoloni, P., Zamboni, M., Panetta, G., De Paolis, A. & Tocchini-Valentini, G.P. (1995) Mutational analysis of the transcription start site of the yeast $tRNA^{Leu}$ gene. *Nucleic Acids Res.* **23**, 2914–2918.