Curr Genet (1994) 26: 557–563

Current Genetics

Structure and expression characteristics of the chloroplast DNA region containing the split gene for $tRNA^{Gly}$ (UCC) from mustard (*Sinapis alba* L.)

Karsten Liere, Gerhard Link

Arbeitsgruppe Pflanzliche Zellphysiologie und Molekularbiologie, Fakultät für Biologie, Ruhr-Universität Bochum; D-44780 Bochum, Germany

Received: 28 February 1994 / 18 April 1994

Abstract. The mustard chloroplast gene *trnG*-UCC is split by a 717-bp group-II intron. Northern hybridization and RNase protection experiments suggest cotranscription with the upstream *psbK-psbI* operon, but not with the downstream *trnR*-UCU gene. The ends of most RNase-protected fragments between *psbI* and *trnG* correlate with the position of two potential stem-loop structures in this region, which could act as RNA processing elements. However, one RNA 5′ end, approximately 75 bp upstream of the *trnG* 5′ exon, does not so correlate and is preceded by prokaryotic-type '–10' and '–35' sequence elements. This suggests the possibility that a fraction of the *trnG* transcripts is initiated here. All precursor transcripts spanning the *trnG* region seem to have a common 3′ end, which was located 117 bp downstream from the 3′ exon, immediately after a stem-loop region. During seedling development, the major 0.8–0.9-kb *trnG* precursor transcripts show a transient maximum level at around 48 h after sowing, at a time when the mature tRNA begins to accumulate to constant levels. No significant differences in transcript patterns were observed either in the light or in darkness.

Key words: Plastid gene expression – RNA processing – Group-II intron – *trnG*

Introduction

Chloroplasts contain their own genetic system, with genes for at least 50–100 proteins as well as for the organelle rRNAs and tRNAs, while other chloroplast components are nuclear-encoded (Bogorad and Vasil 1991 a, b; Sugiura 1992). RNA maturation mechanisms, as part of the plastid gene expression program, comprise a diverse array of processes such as internal cutting of polycistronic transcripts, *cis*- and *trans*-splicing, and RNA editing, as well as formation of the mature 5′ and 3′ ends (Weil 1987; Gruissem 1989 a, b; Hoch et al. 1991; 1992). Much of this work has been done with plastid tRNA genes, which are transcribed both as mono- or poly-cistronic precursors (Ohme et al. 1985; Shinozaki et al. 1986). Sequence elements similar to the prokaryotic stigma^{70} promoter have been found in front of a number of these genes (Sugiura and Wakasugi 1989). Others lack, or do not require, such 5′ promoter sequences, but reveal internal elements reminiscent of the A and B boxes of nuclear genes transcribed by RNA polymerase III (Gruissem et al. 1986; Geiduschek and Tocchini-Valentini 1988; Jahn 1992). It is well-established that processing of tRNA transcripts with 5′ and 3′ sequences, as well as splicing of intron sequences, requires nuclease activities (Deutscher 1984; Greenberg et al. 1984; Wang et al. 1988). However, in the case of plastid tRNA precursors, the exact mechanism and sequential order of these processing and splicing events remain to be clarified (see, e.g., Delp et al. 1991).

On tobacco chloroplast DNA the split *trnG* gene coding for $tRNA^{Gly}$(UCC) is located downstream from *psbK* and *psbI* (encoding the two small photosystem II-associated K and I polypeptides) and immediately upstream of *trnR*-UCU (Shinozaki et al. 1986). Transcript analysis of this region revealed cotranscription of *trnG* with *psbK* and *psbI*, but not with *trnR*, suggesting that *trnG* expression in tobacco is co-regulated with *psbK-psbI* by usage of the far-upstream promoter of that operon (Meng et al. 1991). However, an additional promoter was found directly in front of *trnG*, with the initiation site only 6 ntd upstream of the *trnG* 5′ exon (Deno and Sugiura 1984; Meng et al. 1991). These results, and the observation that steady-state levels of *trnG* transcripts are much higher than those of *psbK-psbI*, led Meng et al. (1991) to suggest that the two promoters may be differentially regulated in tobacco.

In case of mustard (*Sinapis alba*), cotranscription of the *psbK* and *psbI* genes was reported by Neuhaus and Link (1990). Here we have analyzed the adjacent region that contains the *trnG* and *trnR* genes. We present the nucleotide sequence and we have searched for sequence ele-

This paper is dedicated to Professor Peter Sitte on the occasion of his 65th birthday

Correspondence to: G. Link

ments that might serve as sites for transcriptional initiation and termination as well as for RNA processing and splicing. We addressed the question of whether *trnG* is co-transcribed with *psbK* and *psbI* by using RNase protection and Northern experiments for transcript analysis. Furthermore, we compared levels of *trnG* transcripts during mustard seedling development in order to define their mode of expression.

Materials and methods

Plant material. Mustard seedlings (cv 'Albatros') were grown on moist filter papers at 25 °C in the dark or under white light from mercury discharge lamps (photon fluence rate 250 $\mu mol \cdot m^{-2} \cdot s^{-1}$).

Plasmid DNA. All plasmids used for the experiments are subclones of pSA364 (Link and Langridge 1984). Plasmid pBSE996 contains the 996-bp *Eco*RI fragment cloned into pBluescript (Stratagene). Transcription of the *Hin*dIII-linearized plasmid with T7 RNA polymerase produces a *trnG*-specific probe. Plasmid pSPTEB395 contains the *trnG* 5′ exon, 247 bp of the *trnG* intron, and 123 bp of the 5′ flanking region. It was constructed by inserting the 395-bp *Eco*RI-*Bam*HI fragment of pBSE996 into pSPT19 (Pharmacia LKB). Plasmid pSPTSE140 has 10 bp of the *trnG* intron, the 3′ exon, and 80 bp of the 3′ flanking region. It was constructed by cloning the 140-bp *Eco*RI-*Sal*I fragment of pBSE996 ito pSPT18. Plasmid pBSBE125 contains 125 bp of the *trnG* 5′ flanking region and consists of a 125-bp *Eco*RI-*Bgl*II fragment of pSA364-EX0.9 (Neuhaus and Link 1990) inserted blunt end into the *Sma*I site of pBluescript. Plasmid pBSEH440 has 440 bp of the *trnG* 3′ flanking region, which contains the *trnR* coding region. It was constructed by inserting the 440-bp *Eco*RI-*Hin*dIII fragment of pSA364 into pBluescript.

Isolation of RNA, Northern analysis. The procedure for extracting chloroplast RNA and total cellular RNA from mustard cotyledons was described previously (Hughes et al. 1987). For Northern analysis, RNA (10 µg) was heat-denatured, separated electrophoretically on 1.7% (w/v) formaldehyde-agarose gels and transferred to nitrocellulose filters (Schleicher and Schuell BA85) using standard procedures (Sambrook et al. 1989). RNA samples were hybridized with labelled RNA probes prepared by transcription of linearized plasmids with T7, T3 or SP6 RNA polymerases (Melton et al. 1984). All hybridization and washing steps were carried out as previously described (Zinn et al. 1983; Hughes et al. 1987).

RNA protection assays. RNase protection mapping of in-vivo transcripts was done with 50 µg of mustard plastid RNA and 1×10^6 cpm or ^{32}P-labelled in-vitro transcripts as described by Sambrook et al. (1989).

DNA sequencing. Plasmids pBSE996 and pBSEH440 were subcloned into pSPT18/19 and pBluescript vectors and sequenced by the dideoxy chain-termination method (Sanger et al. 1977). Computer analysis of the nucleotide sequences was performed by using the Beckman MicroGenie (Queen and Korn 1984) and the IntelliGenetics PC/Gene programs.

Results

Sequence analysis

In previous work we have studied the gene organization and in-vivo expression of the mustard chloroplast DNA region containing the gene cluster *psbK-psbI, trnQ* and

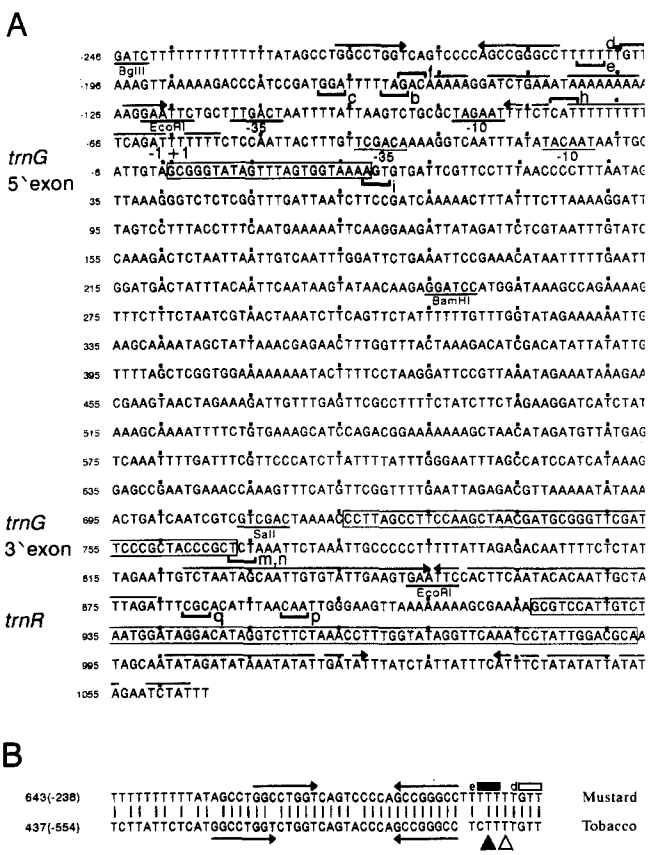

Fig. 1. **A** nucleotide sequence (non-coding strand) of the mustard chloroplast *trnG* and *trnR* genes and their flanking regions. Coding regions are *boxed*. Putative '–35' and '–10' promoter elements are *underlined. Brackets above* the sequence mark the approximate location of 5′ ends, *brackets below* the sequence mark 3′ ends of transcripts inferred from RNAse protection experiments (Fig. 4). Inverted repeat sequences are indicated by *horizontal arrows* (EMBL acession nos. X 17616, X 54725, X 77046). **B** comparison of a region containing a potential stem-loop structure (*horizontal arrows*) upstream of *trnG* from mustard and tobacco (Meng et al. 1991). *Closed triangle and closed bar,* 3′ ends; *open triangle and bar,* 5′ ends; *numbers* correspond to the *psbI* coding region; *numbers in brackets* correspond to the *trnG* coding region

trnS (Neuhaus 1989; Neuhaus et al. 1990; Neuhaus and Link 1990). We now report on the region adjacent to that cluster on the 3′ site of the *psbI* gene. The nucleotide sequence (Fig. 1 A) reveals a putative *trnG* coding region, 772 bp downstream from the *psbI* coding region (Neuhaus et al. 1990), followed by the putative gene for $tRNA^{Arg}$(UCU) at a distance of 154 bp from the *trnG* 3′ exon on the same strand. The *trnG* 5′ exon is identical in mustard and tobacco, whereas in the 3′ exon the third nucleotide in mustard is a 'T' instead of the 'C' in tobacco. In both genes the coding region is interrupted by an intron at an equivalent position in the D-stem region. 5′ upstream of *trnG* are two regions that resemble prokaryotic-type (sigma^{70}) promoters. Each of these two regions, located 72 and 12 bp upstream of the *trnG* 5′ exon, respectively, contain two sequence motifs that resemble the '–35' and '–10' sequence motifs of prokaryotic-type promoters ('TTGACT ... TAGAAT' and 'TCGACA ... TACAAT'). The more-proximal putative promoter (i.e., located more

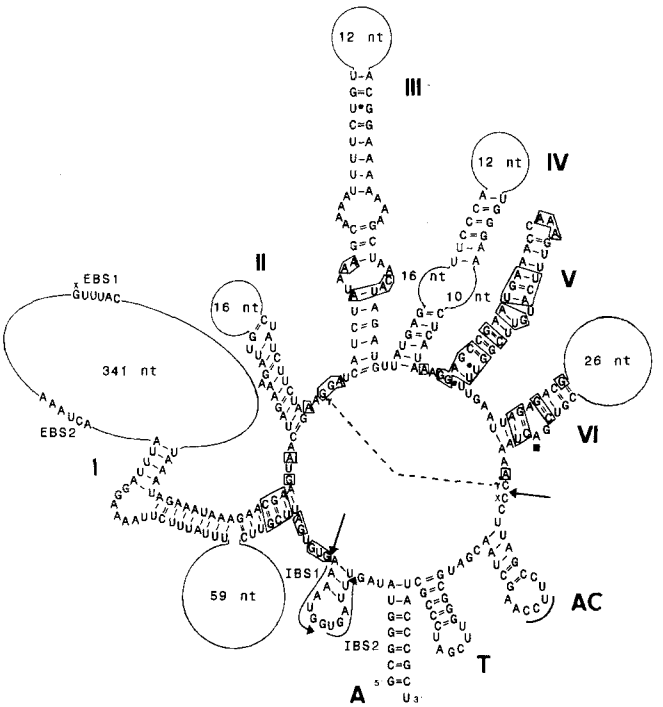

Fig. 2. Secondary structure model for the *trnG*-UCC intron. The *square* denotes the conserved adenine residue which is the proposed nucleophile during lariat formation and intron excision (Van der Veen et al. 1986; Burke 1989). *Straight arrows* mark the putative splice junctions and conserved sequence motifs are *boxed*. The secondary structure is arranged according to the criteria established for group-II introns by Michel and Dujon (1983) and structural elements with known or suspected importance are indicated (Michel et al. 1989; Saldanha et al. 1993). The 'guide pair' is marked by '*x*' symbols. *EBS*, exon-binding site; *IBS*, intron-binding site. γ, γ', proposed tertiary base pair interaction. The anticodon stem-loop (*AC*), T-stem-loop (*T*) and the acceptor stem (*A*) of mature $tRNA^{Gly}$ are also indicated

closely to the *trnG* coding region) was also noted in tobacco (Meng et al. 1991), while the distal promoter (i.e., located more distant to the *trnG* coding region) is within a region that is not conserved between mustard and tobacco (data not shown). No prokaryotic-type promoter elements could be found in front of the *trnR* gene.

Two inverted repeats are located upstream of the *trnG* 5′ exon, with the 5′ end of their more-distal copy at approximately 60 and 210 bp, respectively. The latter repeat is conserved in tobacco, where it is localized 515 bp upstream of the *trnG* 5′ exon (Fig. 1 B). Two other inverted repeats are located downstream from *trnG*, at a distance of 50 and 230 bp from the 3′ exon.

Figure 2 shows that the 717-nt-long mustard intron has characteristic features consistent with a classification as a group-IIB intron as proposed by Michel et al. (1989). These include the sequence elements 'GUGUG' and 'AAC' that conform to the consensus sequences for the conserved elements [('GUGYG') and ('RAY')] at the 5′ and 3′ intron boundaries, respectively (Michel et al. 1989; Saldanha et al. 1993). The intron sequence can be folded into the typical six stem-loop structural (Michel and Dujon 1983), including stem-loop I with the putative exon-binding sites and the highly conserved stem-loop V and stem-loop VI. The latter contains a single 'A' residue exposed in a small bulge eight nucleotides from the 3′ splice junction, which might be the branch point for lariat formation (Peebles et al. 1986; Schmelzer and Schweyen 1986; Van der Veen et al. 1986; Jarrell et al. 1988; Burke 1989). In addition, three characteristic types of elements for proposed tertiary base-pair interactions can be observed. First, exon-intron binding sites: the exon-binding site 1 (EBS1) is complementary to the last five nucleotides of the 5′ exon (intron-binding site 1 – IBS1) and EBS2 is capable of base-pairing with IBS2, located one nucleotide upstream of IBS1. Second, gamma-gamma' interaction: there is a 'G' residue immediately preceding the conserved sequence 'GA' in the single-stranded region between stem-loop II and III (gamma position); this 'G' is thought to base-pair with the last 'C' of the intron at the 3′ intron-exon junction (gamma' position). Third, 'guide pair' interaction: the term 'guide pair' designates two additional base-pairing residues, i.e., the 'G' preceding EBS1 and the first 'C' of the exon at the 3′ intron-exon junction (Michel et al. 1989; Jacquier and Jacquesson-Breuleux 1991; Saldanha et al. 1993). Both characteristic residues are present in the *trnG* sequence (marked by crosses in Fig. 2).

Transcript levels of the plastid trnG gene during mustard seedling development

To estimate transcript levels, total RNA was extracted from seeds and from cotyledons of seedlings up to 84 h after sowing, separated on gels, blotted and then hybridized. The *trnG*-specific probe (Fig. 3 C) detects different transcripts with sizes in the range of 0.8–0.9 kb and (less than) 0.1 kb, both in dark-grown and light-grown seedlings (Fig. 3 A, panels 1 and 2). While the small '0.1-kb' transcript is likely to represent the mature *trnG* tRNA, the transcripts detected in the 0.8–0.9 kb range might reflect one or possibly two different, but closely-related, *trnG* precursors. In some experiments, but not in others, we noted a partial separation of the 0.8–0.9 kb signal into two components, using independently-prepared probes. Both the 0.8- and 0.9-kb portions of the signal are detected by the same probes and therefore contain closely-related transcripts (data not shown).

The levels of the 0.8–0.9 kb *trnG* precursor(s) begin to increase at approximately 12–24 h after sowing. At 48 h, maximum signal intensity is reached. Thereafter, the levels of these transcripts decrease again, indicating rapid transient accumulation of the *trnG* precursor(s), as has been shown before for the *trnK* precursor transcripts (Hughes et al. 1987). The levels of the mature tRNA-related transcript at 0.1 kb begin to increase slightly later than the precursor transcripts at approximately 24–30 h after sowing. In contrast to the latter, they seem to accumulate to a constant level at 84 h after sowing. Throughout this entire period, levels of both the 0.8–0.9-kb and 0.1-kb transcripts in light-grown seedlings (Fig. 3 A, panel 2) are not significantly different from those in dark-grown seedlings (Fig. 3 A, panel 1).

Longer exposure of the autoradiographs (Fig. 3 A) revealed additional signals at 2.5 kb and 1.1 kb, with RNA

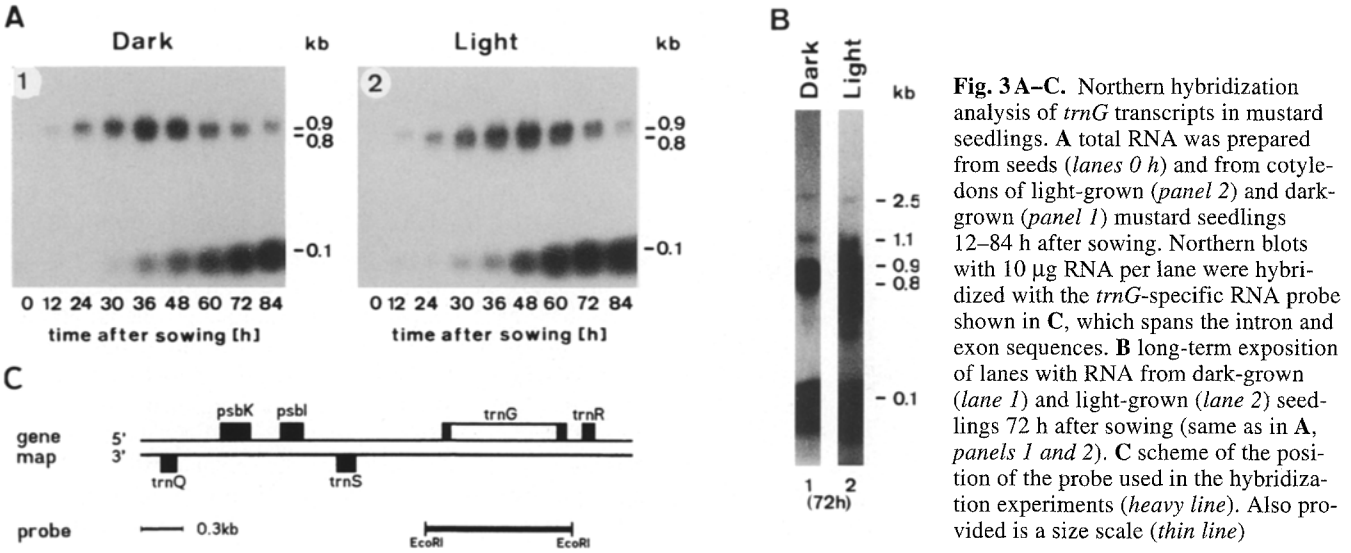

Fig. 3 A–C. Northern hybridization analysis of *trnG* transcripts in mustard seedlings. **A** total RNA was prepared from seeds (*lanes 0 h*) and from cotyledons of light-grown (*panel 2*) and dark-grown (*panel 1*) mustard seedlings 12–84 h after sowing. Northern blots with 10 µg RNA per lane were hybridized with the *trnG*-specific RNA probe shown in **C**, which spans the intron and exon sequences. **B** long-term exposition of lanes with RNA from dark-grown (*lane 1*) and light-grown (*lane 2*) seedlings 72 h after sowing (same as in **A**, *panels 1 and 2*). **C** scheme of the position of the probe used in the hybridization experiments (*heavy line*). Also provided is a size scale (*thin line*)

from either dark-grown or light-grown seedlings. A representative pattern is shown in Fig. 3 B for RNA at 72 h after sowing. Although the hybridization signals at the position of these large transcripts were generally weak, and variable from one preparation to another, their relative strength compared to the 0.8–0.9-kb band was highest at the later stages 60–84 h after sowing, indicating that they accumulate to constant levels (data not shown). The same mode of expression was previously found for the preceding *psbK-psbI* gene cluster (Neuhaus and Link 1990) and a transcript in this size range was also noted using *psbK/psbI* probes (Neuhaus and Link 1990, unpublished data). It is therefore conceivable that the 2.5-kb RNA species reflects a tricistronic precursor transcript (Meng et al. 1991). The smaller 1.1-kb species might represent a processed *trnG* intermediate of the putative 2.5-kb *psbK-psbI-trnG* cotranscript.

5′ RNase mapping of trnG transcripts

To locate in-vivo transcript ends in the 5′ region of the *trnG* gene, a sensitive RNase protection assay was carried out (Sambrook et al. 1989). When the ^{32}P-labelled RNA probe cEB395 (Fig. 4 B) was hybridized to chloroplast RNA, followed by treatment with RNAses A and T1, four resistant fragments in the range of 395 nt ('g'), 345 nt ('h'), 145 nt ('i') and 25 nt ('k') could be detected (Fig. 4 A, lane 6) which are not observed after hybridization of cEB395 to *E. coli* tRNA (Fig. 4 A, lane 5). The intensive band at 25 nt ('k') most likely corresponds to the 5′ *trnG* exon, thereby reflecting part of the mature tRNA. Transcripts consisting of 5′ non-coding and exon sequences, but lacking intron sequences, could be reflected by the 145-nt fragment ('i'), indicating a RNA 5′ end at a position upstream of the *Eco*RI site of cEB395 (Fig. 4 B). However, another possible 5′ end of the *trnG* transcripts seems to map approximately 75 nt upstream of the *trnG* 5′ exon, as was inferred from the presence of a 345-nt RNase-protected fragment ('h'). This fragment is interpreted to contain the entire 5′ exon and 247-nt-long portion of the intron that is covered by the probe. In accordance with the Northern analysis (Fig. 3), fragment 'h' might represent the 5′ portion of the transiently-expressed 0.8–0.9 kb transcript spanning the *trnG* gene. Directly upstream of 5′ end 'h' are the prokaryotic-type promoter motifs (Fig. 1 A), suggesting that this might be a transcription start site. A 395-nt RNA fragment ('g') would be expected if the entire chloroplast sequence within cEB395 is protected by in-vivo transcripts. Hence, band 'i' points to another more-upstream 5′ end that is not covered by this probe.

To further map these predicted 5′ termini of *trnG* precursors, RNA probe cBE125 (Fig. 4 B) was used in similar RNase protection experiments. Hybridization signals were detected that centered around 125 nt ('a'), 92 nt ('b'), 85 nt ('c'), 70 nt ('d'), 50 nt ('e') and 35 nt ('f') (Fig. 4 A, lane 2). The protected products around 125 nt (band 'a') are of the size expected if there are transcripts that cover the entire chloroplast-RNA-specific portion of the probe cBE125 (Fig. 4 B), indicating 5′ end(s) of *trnG* transcripts in the more-distant region not covered by this probe. In previous studies (Neuhaus and Link 1990) of the adjacent region towards the preceding *psbI* gene, no indications were obtained for the presence of a RNA 5′ end. Considering the 2.5-kb transcript detected by Northern analysis (Fig. 3 B), fragment 'a' might indicate cotranscription of *trnG* with *psbK-psbI* from a far-upstream transcription initiation site.

The 3′ end of the *psbK-psbI* precursor was previously mapped 670 nt downstream from the *psbI* reading frame (Neuhaus and Link 1990). It seems likely that the protected fragment 'e' in Fig. 4 A, lane 2, reflects this 3′ end. If the conserved sequence element within this region (Fig. 1 B this paper; Meng et al. 1991) is a processing site, then the (weak) fragment 'd' might represent the 190-nt *trnG* portion of a processed *psbK-psbI-trnG* precursor detected by the probe. Unlike the situation in tobacco, a second potential stem-loop structure is present in the *trnG* 5′ flanking region approximately 50 nt downstream from the more distal one (Fig. 1 A). Immediately in front of the proximal

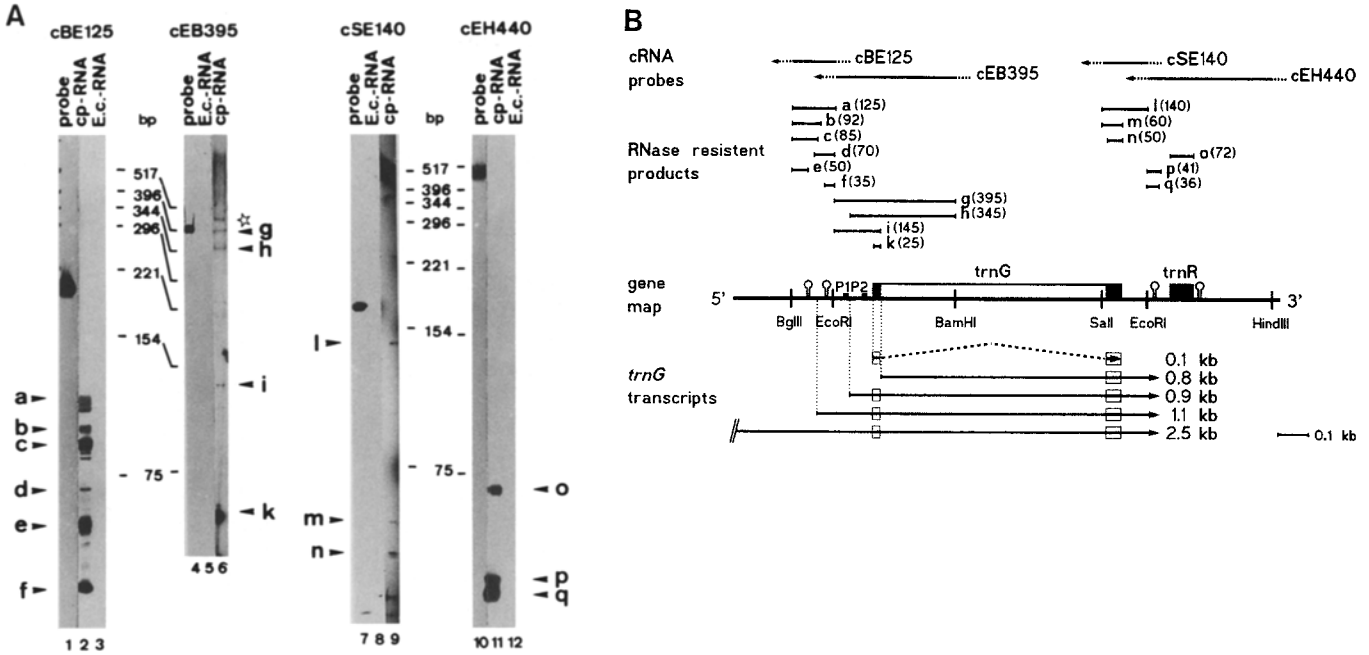

Fig. 4 A, B. RNase protection mapping of in-vivo transcripts of the *trnG* 5′ and 3′ regions. **A** following RNA-RNA hybridization and treatment with RNases, resistant products were separated on a 6% denaturing polyacrylamide gel. The probes (shown in **B**) were cBE125 (*lanes 1–3*), cEB395 (*lanes 4–6*), cSE140 (*lanes 7–9*), cEH440 (*lanes 10–12*). *Lanes 1, 4, 7, 10,* ^{32}P-labelled RNA probes without RNase treatment. *Lanes 3, 5, 8, 12,* RNA probes hybridized to 50 µg of *Escherichia coli* tRNA and subsequently treated with RNases. *Lanes 2, 6, 9, 11,* RNase-resistant products following hybridization with 50 µg of chloroplast RNA. The *asterisk* denotes a band of unknown origin (*lane 6*). **B** scheme showing the organization of the *trnG* gene and its flanking regions, including *trnR*. Coding regions are depicted as *filled boxes*. The *lines above* indicate the putative position of the protected RNA fragments shown in **A**. *Arrows on top,* ^{32}P-labelled in-vitro cRNA probes with *broken lines* representing polylinker and vector sequences. Below the gene map is a summary of the proposed in-vivo transcripts, as deduced from the RNA protection and Northern hybridization experiments (Fig. 3)

element there could be another processing site as indicated by the RNase-protected bands 'b' and 'c'. The 3′ ends of these two fragments roughly match the 5′ end of fragment 'f' (Fig. 4B). Hence, this pattern of protected fragments is consistent with another RNA 5′ end approximately 150 nt upstream of the *trnG* coding region. The 1.1-kb transcript detected by the Northern analysis (Fig. 3B) can be viewed as originating from processing at this suggested site, thus reflecting the processed *trnG* transcript portion of the primary tricistronic precursor.

3′ RNase mapping of trnG transcripts

To locate transcripts that extend into the *trnG* 3′ region, RNase protection assays were carried out, using an RNA probe that contained a small portion of the intron, the 3′ exon, and 80 nt of the 3′ flanking region (Fig. 4B, probe cSE140). As shown in Fig. 4A (lane 9), three RNase-resistant products are generated. The 50-nt band 'n' roughly coincides with the size of the *trnG* 3′ exon and therefore seems to represent the mature, spliced tRNA. Fragment 'm' is 60 nt in length and might represent the *trnG* 3′ exon together with 10 nt of the adjacent intron sequences covered by the probe, reflecting a (3′) unspliced but 3′ processed *trnG* transcript. The 140-nt fragment is colinear with the entire chloroplast-specific region of the probe, suggesting that in vivo the unprocessed *trnG* transcripts extend further downstream.

To localize the 3′ ends of the unprocessed *trnG* precursor, we used probe cEH440, which covers the downstream region adjacent to probe cSE140 and includes the coding region for the *trnR* gene (Fig. 4B). As shown in Fig. 4A, lane 11, three resistant products were obtained with sizes of 72 nt ('o'), 41 nt ('p'), and 36 nt ('q'), respectively. Bands 'p' and 'q' might reflect the common 3′ ends of the *trnG* precursor(s) in vivo and define them to be 117 nt and 121 nt downstream from the *trnG* 3′ exon (Fig. 1A). Immediately upstream, this region contains sequences capable of forming a stem-loop structure, which would serve as a termination signal (Rosenberg and Court 1979). Band 'o' might represent the *trnR* transcript, indicating that *trnR* is not co-transcribed with the *psbK-psbI-trnG* gene cluster. The size of this fragment correlates with that of the 72-bp *trnR* coding region (Fig. 1A).

Discussion

Northern analysis of the mustard *trnG* gene revealed major precursor transcripts that are 0.8–0.9 kb in length. It is suggested that this size range contains two related, but distinct, transcripts: those which span the entire distance between the putative transcription start site of the *trnG* gene (Fig. 1A, bracket 'h') and the mapped 3′ end (Fig. 1A, brackets 'p' and 'q'), and others that might be splicing intermediates consisting of the intron, the 3′ exon and the 3′ unprocessed portion (Fig. 4B). The putative

transcription start site is suggested to map 75 nt upstream of the *trnG* 5′ exon. The prokaryotic-type promoter elements, 'TTGACT' (–35) and 'TAGAAT' (–10), found directly upstream indicate that this transcript 5′ end might be a transcriptional initiation site. Further sequence analysis revealed a second prokaryotic-type promoter structure located more closely to the *trnG* coding region. This is similar to the situation in tobacco, where Meng et al. (1991) determined a transcriptional inition site 6 bp upstream of *trnG*. However, our RNase mapping experiments in mustard did not indicate the presence of a 5′ end in this region. Conversely, in the tobacco sequence there are no additional promoter elements in front of the *trnG* gene. In the case of the *trnK* gene encoding $tRNA^{Lys}$ (UUU) from mustard and tobacco, two suggested prokaryotic-type promoter structures are located upstream of the *trnK* coding region (Sugita et al. 1985; Neuhaus and Link 1987). It remains to be established whether there is a function for and a utilization of this dual promoter architecture in the developmental expression and regulation of those genes, perhaps comparable to the suggested differential usage of *cis*-elements within the *psbA* promoter (Eisermann et al. 1990; Tiller and Link 1993).

In contrast to the work on tobacco, in addition to the putative mature *trnG* 3′ end we detected another mustard *trnG* transcript end approximately 120 nt downstream from the 3′ exon. This distal site is preceded by an inverted repead structure (delta G = –42 kcal) that might serve as a transcription termination or processing signal (Figs. 1 A and 4). No precursor transcripts with 5′ and 3′ extensions or upstream promoter consensus sequences could be detected in the case of the adjacent *trnR* gene. This could point to rapidly-processed, and hence undetected, large co-transcripts of *trnG* and *trnR*. On the other hand, as demonstrated for spinach *trnR*-UCU and *trnS*-UGA (Gruissem et al. 1986) as well as for *Chlamydomonas reinhardtii trnE* (Jahn 1992), no upstream promoter elements are required by several chloroplast tRNA genes. Sequence analysis of these genes and *trnR* and *trnS* from mustard (this paper; Neuhaus et al. 1990) revealed internal sequence elements resembling the A- and B-block of nuclear genes transcribed by RNA polymerase III (Galli et al. 1981).

The pattern of fragments generated in the RNase protection assays shows no fragments pointing to *trnG* RNA species that consist of 5′ exon plus intron sequences or of the 3′ exon plus the 3′-terminal sequences (Fig. 4, RNA probes cEB395 and cSE140). On the other hand, fragments 'i' and 'm' might reflect RNAs consisting of the 5′ exon with 5′ flanking sequences and the 3′ exon together with intron sequences, respectively (Fig. 4). This suggests that the removal of the 3′-terminal sequences precedes splicing of the intron sequences, while splicing of the 5′ intron sequences seems to occur prior to processing of the 5′ sequences. This would be consistent with the view that 5′-terminal cleavage by the chloroplast RNase P is one of the final chloroplast tRNA processing events, whereas 3′ endonucleolytic cleavage probably occurs as the first step prior to the splicing of the intron sequences (Delp et al. 1991).

Similar to the situation in tobacco (Meng et al. 1991), the mustard *trnG* gene might be cotranscribed with the preceding *psbK-psbI* genes, as indicated by the 2.5-kb transcripts detected in the Northern experiments (Fig. 3 B). The 2.5-kb precursor might be processed to the 1.1-kb species by cleavage in a region approximately 640 kb downstream from the *psbI* coding region. This region contains two potential stem-loop structures, one of which is conserved between mustard and tobacco (Fig. 1). Apart from a suggested role of inverted repeat elements in the termination of transcription (Weil 1987; Westhoff and Herrmann 1988), they were shown to be involved in RNA processing and stabilization (Stern and Gruissem 1987; Stern et al. 1989, 1991; Adams and Stern 1990). In contradiction with their surrounding regions, these stem-loops seem to act as recognition sites for protein-factors (Hsu-Ching and Stern 1991 a, b; Nickelsen and Link 1991, 1993; Schuster and Gruissem 1991). Nevertheless, the in-vivo analyses of Blowers et al. (1993) indicated not only that 3′ sequences play a role in transcript stabilization, but that 5′ sequences also affect the differential degradation or stabilization of chloroplast transcripts (Salvador et al. 1993). Thus, it might be that different 5′ regions of *trnG* transcripts, generated at different processing sites, are involved in the post-transcriptional regulation of *trnG* gene expression.

What might be the role for dual transcription of *trnG* both from its own promoter and as part of the tricistronic *psbK-psbI-trnG* operon? This cotranscription could be viewed as the result of a 'leaky' (inefficient) termination signal downstream from *psbI* and may have no special function. Alternatively, the utilization of two distinct promoters could be an instrument for differential regulation of *trnG* gene expression. The latter mechanism was shown for the *psbD-psbC* operon from barley, which is differentially regulated via promoter switching in a light/dark-dependent manner (Sexton et al. 1990 a, b).

Our Northern hybridization experiments, however, did not show evidence for differential accumulation of *trnG* transcripts during mustard seedling development in either the light or in darkness. The 0.8–0.9-kb *trnG* transcripts become detectable early at 12–24 h after sowing and then accumulate transiently. In contrast, the preceding *psbK* and *psbI* genes are expressed in a mode that is characterized by late accumulation to constant levels ('constitutive mode') (Neuhaus and Link 1990). Both the 2.5-kb tricistronic transcript and the 1.1-kb suggested processed intermediate seem to follow this latter mode of expression. Hence, the proximal *trnG* promoter and the distal promoter upstream of *psbK-psbI* may be differentially regulated, with a portion of *trnG* precursors being cotranscribed and thus accumulating during the late phase of seedling development. However, since these sequences represent only a small fraction within the entire transcript population in the 0.8–0.9 kb size range, this effect is not visible in the Northern experiment.

Evidence was obtained that a portion of the *psbA* transcripts is cotranscribed with the preceding split *trnK* gene (Nickelsen and Link 1991). While the *psbA* transcript levels show an initial light-independent increase until 30–40 h, and thereafter accumulate to high levels only in light-grown seedlings, the *trnK* gene is expressed transiently with peak levels around 40 h (Hughes et al. 1987).

This is reminiscent of the so-called 'super-operons' in purple nonsulfur bacteria (Wellington et al. 1992) which are cotranscribed, but differentially-expressed, groups of genes, thus serving as a common target for graded responses to environmental conditions. In this context, it is notable that two other split tRNA genes, *trnA*-UGC and *trnI*-GAU, which are cotranscribed within the *rrn* operon, also show transient light-independent transcript accumulation (unpublished data). Hence, it might be a common feature of transiently-expressed chloroplast genes that they are split by a group-II intron. It will be interesting to see if the putative maturase that is encoded by the *trnK* intron (Neuhaus and Link 1987) plays a common role in the regulation of expression of these split tRNA genes.

Acknowledgements. We thank U. Müller and C. Wittig for expert technical assistance. We are indebted to H. Neuhaus-Kück and J. Nickelsen, who both made significant contributions during the initial part of the project. This work was supported by the Deutsche Forschungsgemeinschaft and the Fonds der Chemischen Industrie, FRG.

References

Adams CC, Stern DB (1990) Nucleic Acids Res 18:6003–6010
Blowers AS, Klein U, Ellmore GS, Bogorad L (1993) Mol Gen Genet 238:339–349
Bogorad L, Vasil IK (1991 a) The molecular biology of plastids. Cell culture and somatic cell genetics, vol 7 A. Academic Press, San Diego
Bogorad L, Vasil IK (1991 b) The photosynthetic apparatus: molecular biology and operation. Cell culture and somatic cell genetics, vol 7 B. Academic Press, San Diego
Burke JM (1989) Methods Enzymol 180:533–545
Delp G, Igloi GL, Kössel H (1991) Nucleic Acids Res 19:713–716
Deno H, Sugiura M (1984) Proc Natl Acad Sci USA 81:405–408
Deutscher MP (1984) Crit Rev Biochem 17:45–71
Eisermann A, Tiller K, Link G (1990) EMBO J 9:3981–3987
Galli G, Hofstetter H, Birnstiel ML (1981) Nature 294:626–631
Geiduschek EP, Tocchini-Valentini GP (1988) Annu Rev Biochem 57:873–914
Greenberg BM, Gruissem W, Hallick RB (1984) Plant Mol Biol 3:97–109
Gruissem W, Elsner-Menzel C, Latshaw S, Narita JO, Schaffer MA, Zurawski G (1986) Nucleic Acids Res 14:7541–7557
Gruissem W (1989 a) Cell 56:161–170
Gruissem W (1989 b) Chloroplast RNA: transcription and processing. In: Marcus A (ed) The biochemistry of plants. Vol. 15, Molecular biology. Academic Press, San Diego
Hoch B, Maier RM, Appel K, Igloi GL, Kössel H (1991) Nature 353:178–180
Hsu-Ching C, Stern DB (1991 a) Mol Cell Biol 11:4380–4388
Hsu-Ching C, Stern DB (1991 b) J Biol Chem 266:24205–24211
Hughes JE, Neuhaus H, Link G (1987) Plant Mol Biol 9:355–363
Jacquier A, Jacquesson-Breuleux N (1991) J Mol Biol 219:415–428
Jahn D (1992) Arch Biochem Biophys 298:505–513
Jarrell KA, Peebles CL, Dietrich RC, Romiti SL, Perlmann PS (1988) J Biol Chem 263:3432–3439
Link G, Langridge U (1984) Nucleic Acids Res 12:945–958
Melton DA, Krieg PA, Rebagliati MR, Maniatis T, Zinn K, Green MR (1984) Nucleic Acids Res 12:7035–7057
Meng BY, Wakasugi T, Sugiura M (1991) Curr Genet 20:259-264
Michel F, Dujon B (1983) EMBO J 2:33–38
Michel F, Umesono K, Ozeki H (1989) Gene 82:5–30
Neuhaus H (1989) Nucleic Acids Res 17:444
Neuhaus H, Link G (1987) Curr Genet 11:251–257
Neuhaus H, Link G (1990) Curr Genet 18:377–383
Neuhaus H, Pfannschmidt T, Link G (1990) Nucleic Acids Res 18:368–368
Nickelsen J, Link G (1991) Mol Gen Genet 228:89–96
Nickelsen J, Link G (1993) Plant J 3:537–544
Ohme M, Kawogashira T, Shinozaki K, Sugiura M (1985) Nucleic Acids Res 13:1045–1056
Peebles CL, Perlman PS, Mecklenburg KL, Petrillo ML, Tabor JH, Jarrell KA, Cheng H-L (1986) Cell 44:213–223
Queen C, Korn LJ (1984) Nucleic Acids Res 12:581–599
Rosenberg M, Court D (1979) Annu Rev Biochem 13:319–353
Saldanha R, Mohr G, Belfort M, Lambowitz AM (1993) FASEB J 7:15–24
Salvador ML, Klein U, Bogorad L (1993) Proc Natl Acad Sci USA 90:1556–1560
Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning: a laboratory manual, 2nd edn. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York
Sanger F, Nicklen S, Coulson ARC (1977) Proc Natl Acad Sci USA 74:5463–5467
Schmelzer C, Schweyen RJ (1986) Cell 46:557–565
Schuster G, Gruissem W (1991) EMBO J 10:1493–1502
Sexton TB, Jones JT, Mullet JE (1990 a) Curr Genet 17:445–454
Sexton TB, Christopher DA, Mullet JE (1990 b) EMBO J 9:4485–4494
Shinozaki K, Ohme M, Tanaka M, Wakasugi T, Hayashida N, Matsubayashi T, Zaita N, Chungwongse J, Obokata J, Yamaguchi-Shinozaki K, Deno H, Kamogashira T, Yamada K, Kusuda J, Takaiwa F, Kato A, Tohdoh N, Shimada H, Sugiura M (1986) EMBO J 5:2043–2049
Stern DB, Gruissem W (1987) Cell 51:1145–1157
Stern DB, Jones H, Gruissem W (1989) J Biol Chem 264: 18742–18750
Stern DB, Radwanski ER, Kindle KL (1991) Plant Cell 3:285–297
Sugita M, Shinozaki K, Sugiura M (1985) Proc Natl Acad Sci USA 82:3557–3561
Sugiura M (1992) Plant Mol Biol 19:149–168
Sugiura M, Wakasugi T (1989) Crit Rev Plant Sci 8:89–101
Tiller K, Link G (1993) EMBO J 12:1745–1753
Van der Veen R, Arnberg AC, Van der Horst G, Bonen L, Tabak HF, Grivell LA (1986) Cell 44:225–234
Wang MJ, Davis MW, Gegenheimer P (1988) EMBO J 7:1567–1574
Weil JH (1987) Plant Sci 49:149–157
Wellington CL, Bauer CE, Beatty JT (1992) Can J Microbiol 38:20–27
Westhoff P, Herrmann RG (1988) Eur J Biochem 171:551–564
Yamaguchi-Shinozaki K, Shinozaki K, Sugiura M (1987) FEBS Lett 215:132–136
Zinn K, DiMaio D, Maniatis T (1983) Cell 34:865–879

Communicated by H. Kössel