Curr Genet (1998) 34: 212–215

ORIGINAL PAPER

Nathalie A. Lusson · Philippe M. Delavault
Patrick A. Thalouarn

The *rbcL* gene from the non-photosynthetic parasite *Lathraea clandestina* is not transcribed by a plastid-encoded RNA polymerase

Received: 12 January / 8 July 1998

Abstract In the plastome of the obligate root-parasitic plant, *Lathraea clandestina*, the *rbcL* gene has been maintained and is expressed, despite the reduced size and gene content of the plastid genome. Some of the plastid genes involved in translation (e. g. transfer RNAs, ribosomal RNAs and ribosomal proteins) have been sequenced and still appear to code for functional ribosomal components. Indeed, the *16S rRNA* and *rpl20* genes are expressed whilst other necessary tRNA and ribosomal protein-encoding genes have probably been deleted or truncated. Although obtained by PCR, the four *rpo* genes for *Escherichia coli*-like plastid encoded RNA polymerase appear to be pseudogenes. Nevertheless, the *rbcL* gene, with a "–10, –35" prokaryotic-like promoter, is still transcribed. In contrast to photosynthetic plants, *rbcL* transcripts in *Lathraea* are larger in their 5′ region and cover the prokaryotic-like promoter. The transcription initiation site is located near the ATG start codon of the *atpB* pseudogene. Similarity to non-consensus *E. coli*-like plastid promoters suggests that *rbcL* transcription is driven by a nuclear-encoded RNA polymerase.

Key words Parasitic plant · *Lathraea clandestina* · *rpo* genes · *rbcL*

Introduction

The plastid genome of angiosperms is highly conserved in gene content and most, if not all, of its non-housekeeping genes code for proteins involved in photosynthesis and chloro-respiration. It also codes for four subunits of a bacterial-type RNA polymerase (Shinozaki et al. 1986). However, a nuclear-encoded phage-type RNA polymerase (Hedtke et al. 1997) transcribes some plastid genes from distinct promoters (Hübschmann and Börner 1998; Miyagi et al. 1998). The holoparasitic plants, whose plastids are mostly amyloplastids devoid of chlorophyll, obtain their reduced carbon directly from their photosynthetic host. This phenomenon raises the question as to whether holoparasites still need a plastid genome. The presence of a plastid genome, reduced in size and gene-content, was first demonstrated (dePamphilis and Palmer 1990) in *Epifagus virginiana* (Orobanchaceae). In a plastome of 71 kb, all photosynthesis-related genes were found to be absent or else present only as pseudogenes. Moreover, intact genes for the plastid-encoded RNA polymerase were missing and only a truncated pseudogene for *rpoA* was identified (Wolfe et al. 1992). A similar result was obtained in another related Orobanche, *Conopholis americana* (Wimpee et al. 1992), which contains only a short pseudogene for *rpoB* (89 bp) of the four *rpo* genes. However, in another root holoparasite within the Scrophulariaceae, *Lathraea clandestina*, in contrast to all other Orobanchaceae so far studied, the *rpo* genes could be amplified by PCR (Delavault et al. 1996). Moreover the plastome has retained an intact and transcribed *rbcL* gene although a function for Rubisco in its amyloplasts remains unclear (Delavault et al. 1995). Other studies on *Orobanche* spp. (Thalouarn et al. 1994; Wolfe and dePamphilis 1997) suggest that *rbcL* is likely to have been the last among the photosynthetic genes to have lost its functionality. The observed expression of the *rbcL* gene in *Lathraea* raises the question as to whether it is still transcribed by a plastome-encoded RNA polymerase, as in all other land plants.

N. A. Lusson and P. M. Delavault contributed equally to this work

N. A. Lusson · P. A. Delavault · P. A. Thalouarn (✉)
Groupe de Physiologie et Pathologie Végétales,
Université de Nantes, 2 rue de la Houssinière,
BP 92208, F-44322 Nantes cedex 3, France
e-mail: patrick.thalouarn@svt.univ-nantes.fr
Fax: +33-2-51 12 56 12

Communicated by A. Brennicke

Materials and methods

L. clandestina and tobacco plastid DNA were isolated following the method previously described by Delavault et al. (1996). Total RNA

was isolated using the "RNeasy Plant Mini Kit" from Qiagen according to the manufacturer's instruction. DNA amplification was conducted in a thermal cycler (Perkin-Elmer-Cetus). One-hundred to five-hundred nanograms of DNA were amplified as described by Thalouarn et al. (1994). For each gene, two tobacco sequences were used as primers (Shinozaki et al. 1986). PCR-amplified products were cloned in the pCR 2.1 plasmid from the "Original TA cloning kit" (Invitrogen) and sequenced by the chain-termination method using the "T7 sequenase quick-denature plasmid sequencing kit" for double-stranded DNA (Amersham, Life Science). Sequence analysis was carried out using the Mac DNASIS software system (Hitachi Software Engineering). 5′ RACE-PCR was performed using the "5′ RACE System kit" (Life Technologies) adapted from the protocol developed by Frohman et al. (1988). The different antisense-specific primers used were rbcL5′ (5′-TCCATGTACCAGTCGAAGATT-3′), rbcLsp (5′-GGAGTTACTCGGAATGCTGCC-3′) and rbcLrace (5′-TGTTTCGTATTCAGGAGTA-3′). The PCR products were transferred to a nylon membrane and hybridised to a *Lathraea*-specific probe labelled by random-primed incorporation of digoxigenine-labelled deoxyuridine triphosphate (DIG RNA labelling kit, Boehringer Mannheim). Detection was performed with the DIG luminescent detection kit (Boehringer Mannheim). In order to determine the cDNA 5′-end, PCR products were cloned and sequenced as above.

Results

Plastid-encoded genes *rps7, rpl20, trnR* and *trnN* are highly conserved

PCR amplification and sequence analysis have previously shown that the *rpl20* gene is conserved in a functional form on the *Lathraea* LSC (Lusson et al. 1995).

In *Lathraea*, as in tobacco and *Epifagus*, northern-blot hybridization of total RNA with a homologous *rpl20* RNA probe revealed a major transcript of 3.0 kb. This corresponds to a mature tri-cistronic *clpP*-5′*rps12-rpl20* mRNA. The *rps7* gene, located on the intergenic region (IR) (Delavault et al. 1996), is identical in size and 97.6% identical to that of tobacco, with 11 substitutions (GenBank accession number AF030982). We conclude that *rps7* is, in all probability, still functional.

Two tRNA genes, *trnR*-Arg(ACG) and *trnN*-Asn (GUU), found in the IR of tobacco ptDNA, are also present in *Lathraea.* Sequencing of these *trn* genes revealed neither deletion nor insertion and they are nearly identical to those of tobacco and *Epifagus* (GenBank accession numbers; *trnR:* AF03098 and *trnN:* AF030981). Consequently, *Lathraea trnR*-Arg (ACG) and *trnN*-Asn (GUU) can be folded into cloverleaf structures.

rpo genes for a plastid-encoded RNA polymerase are pseudogenes

While heterologous hybrization and PCR amplification had demonstrated that the *rpoA* gene is conserved on the *Lathraea* LSC, sequence analysis showed that this gene contains six additional bases (1020 bp) and is 88% identical to that of tobacco (GenBank accession number AF030983). However, the deduced amino-acid sequence

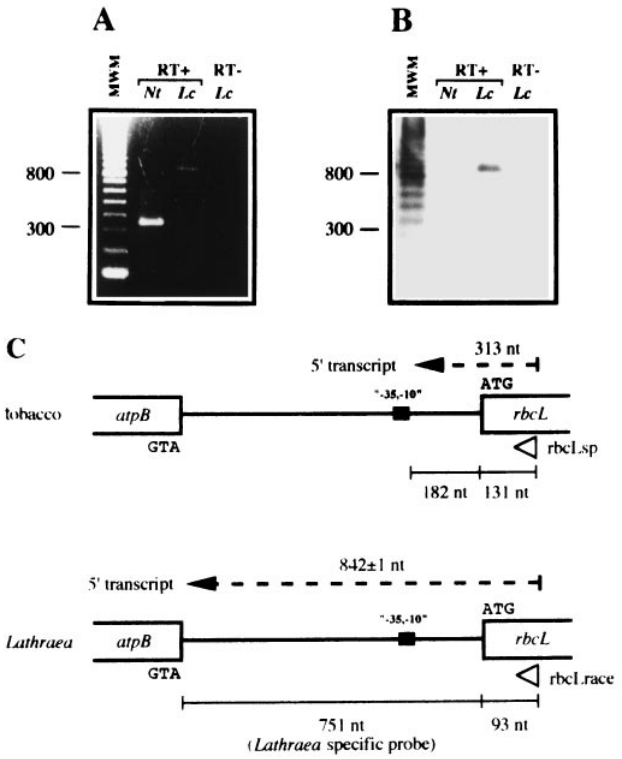

Fig. 1 A–D 5′ RACE analysis of *rbcL* transcripts in tobacco (*Nt*) and *Lathraea* (*Lc*). **A** 5′ RACE products (*RT+*) electrophoresed on an agarose gel. When reverse transcriptase is omitted (*RT–*), no amplified transcript is obtained. Molecular-weight markers (*MWM*) are indicated in bp on the left. **B** Southern-blot analysis of the 5′ RACE products using a *Lathraea*-specific probe (see **C**). **C** schematic representation of the *rbcL-atpB* region showing primer locations, the *Lathraea*-specific probe, and the 5′ transcript termini obtained in *Lathraea* and tobacco. **D** alignment of DNA sequences flanking some plastid NEP promoter transcription-initiation sites. *Vertical bars* indicate identical nucleotides between *Lathraea* and consensus sequences. Positions of the 5′ ends are marked by *filled circles.* The ATG codons of the *atpB Lc* pseudogene and the *Nt* gene are *underlined.* (*1*) Shinozaki et al. (1986); (*2*) Kapoor et al. (1997); (*3*) Hajdukiewicz et al. 1997

reveals the presence of seven internal stop codons, demonstrating that the *rpoA* gene in the *Lathraea* plastome is no longer functional. Furthermore, the intergenic region between *petD* and *rpoA* is deleted in *Lathraea*, and the 3′ regions of both genes overlap by 14 bp.

The *rpoB-C1-C2* operon was partially sequenced from cloned restriction fragments and PCR products (GenBank accession numbers: AF030984, AF039188, AF039189, AF039190, AF039191, AF039192, AF039193, AF039194 and AF039195). Fragments of 338, 435 and 928 nt in the 5′ regions of the three *rpo* genes (*B, C1* and *C2* respectively) exhibit approximately 80% identity with those of

tobacco. However, the deduced amino-acid sequences are only similar to those of tobacco near the N-terminus and totally diverge from codons 16, 107 and 180 respectively. Numerous stop codons are found in these sequences, starting at positions 7, 121 and 184 in *rpoB, C1* and *C2*, respectively.

Unusual length of the 5′ untranslated *rbcL* transcript

Amplification of 5′-end of the *rbcL* transcript with the rbcLsp primer led to PCR products of approximately 850 nt for *Lathraea*, instead of the 350 nt for tobacco expected from the "–10, –35" location (Shinozaki and Sugiura 1982) (Fig. 1 A). Southern probing with the *Lathraea atpB-rbcL* intergenic region (Fig. 1 B) confirms the transcript size in *Lathraea* (the stringency conditions did not reveal tobacco transcripts), suggesting that the 5′-end of *Lathraea rbcL* transcripts is dramatically longer than in other plants (Fig. 1 C). 5′ RACE PCR products for *Lathraea rbcL* were cloned and sequenced and, as expected from the length of the PCR product, the nucleotide sequence was found to cover a large part of the *atpB-rbcL* intergenic region. The transcription initiation site can be localised 2 – 4 nt downstream from the ATG codon of the *atpB* coding region.

Discussion

Some of the results reported in this work suggest the functionality of the *Lathraea* plastid gene-expression apparatus. Firstly, though we did not check for all of them, transfer RNAs (*trnN* and *trnR*) seem to be better maintained on *Lathraea* plastid DNA than on that of *Epifagus*. Besides, the codon usage observed from the Rubisco LSU amino-acid sequences exhibits very little difference between *Lathraea* and tobacco. The high degree of nucleotide-sequence conservation of *rpl20* and *rps7*, as well as *16S rRNA* gene expression, suggest that *Lathraea* plastid ribosomes are not very different in their composition from those of photosynthetic species. We were, however, unable to obtain PCR products of the *rpl2, rpl23, rpl32* and *rps15* genes, thus questioning the capability of *Lathraea* plastid ribosomes.

A RNA polymerase must be active in *Lathraea* plastids since transcription of the *16S RNA, rpl20* and *rbcL* genes is observed. Sequencing of the whole *rpoA* and of large fragments in the 5′ region of *rpoB, C1* and *C2* identifies several stop-codons throughout the four *rpo* genes. Thus, the hypothesis of a plastid-encoded RNA polymerase should be discarded. The observed transcriptional activity must be allocated to a RNA polymerase, imported from the cytoplasm, recognizing a putative additional *rbcL* promoter. The unusual length of the 5′ untranslated extremity implies that the *Escherichia coli*-like "–10, –35" promoter located at the "–168" position in *Lathraea rbcL* is not used, as it is in all the green plants studied so far. Our results suggest that there is a unique transcription initiation site for *Lathraea rbcL*, located in the *atpB-rbcL* intergenic region between 2 and 4 nt from the *atpB* ATG start codon. DNA sequences flanking the novel transcription initiation site do not contain a PEP promoter with "–10, –35" consensus sequences but share a high degree of similarity with the NEP promoter consensus sequence proposed by Hajdukiewicz et al. (1997), with 7 out of 10 identical nt (Fig. 1 D). These features suggest that, in *Lathraea, rbcL* transcription is driven by a NEP, instead of a PEP, similar if not identical to that which transcribes several other plastid genes, and most likely encoded in *Arabidopsis* by the *RpoPt* gene (Hedtke et al. 1997).

These features raise the question of the emergence of a new site for NEP promoter elements in a region which overlaps the *atpB* 5′ upstream and coding regions. This region (particularly the 15 – 25 nucleotides immediately upstream of the *atpB* start codon) is known to be variable in dicots (Manen et al. 1994) in an otherwise better-conserved *atpB-rbcL* intergenic region. This may explain the occurrence of a NEP promoter sequence for *rbcL* in *Lathraea*.

Acknowledgements The authors thank Denneal Jamison and Geoffrey Williams for critical reading of the manuscript.

References

Delavault P, Sakanyan V, Thalouarn P (1995) Divergent evolution of two plastid genes, *rbcL* and *atpB*, in a non-photosynthetic parasitic plant. Plant Mol Biol 29: 1071 – 1079

Delavault P, Russo N, Lusson N, Thalouarn P (1996) Organization of the reduced plastid genome of *Lathraea clandestina*, an achlorophyllous parasitic plant. Physiol Plant 96: 674 – 682

dePamphilis C, Palmer J (1990) Loss of photosynthetic and chlororespiratory genes from the plastid genome of a parasitic flowering plant. Nature 348: 337 – 339

Frohman M, Michael K, Martin G (1988) Rapid production of full-length cDNAs from rare transcripts: amplification using a single gene-specific oligonucleotide primer. Proc Natl Acad Sci USA 85: 8998 – 9002

Hajdukiewicz P, Allison L, Maliga P (1997) The two RNA polymerases encoded by the nuclear and the plastid compartments transcribe distinct groups of genes in tobacco plastids. EMBO J 16: 4041 – 4048

Hedtke B, Börner T, Weihe A (1997) Mitochondrial and chloroplast phage-type RNA polymerases in *Arabidopsis*. Sciences 277: 809 – 811

Hübschmann T, Börner T (1998) Characterisation of transcript initiation sites in ribosome-deficient barley plastids. Plant Mol Biol 36: 493 – 496

Kapoor S, Suzuki J, Sugiura M (1997) Identification and functional significance of a new class of non-consensus-type plastid promoters. Plant J 11: 327 – 337

Lusson N, Delavault P, Thalouarn P (1995) The plastid *rpl20* gene from *Lathraea clandestina* (holoparasite) is conserved in a functional form. C R Acad Sci Paris 318: 659 – 664

Manen J-F, Savolainen V, Simon P (1994) The *atpB* and *rbcL* promoters in plastid DNAs of a wide dicot range. J Mol Evol 38: 577 – 582

Miyagi T, Kapoor S, Sugita M, Sugiura M (1998) Transcript analysis of the tobacco plastid operon *rps2/atpI/H/F/A* reveals the existence of a non-consensus type-II (NCII) promoter upstream of the *atpI* coding sequence. Mol Gen Genet 257: 299 – 307

Shinozaki K, Sugiura M (1982) The nucleotide sequence of the tobacco chloroplast gene for the large subunit of ribulose-1,5-bisphosphate carboxylase/oxygenase. Gene 20: 91 – 102

Shinozaki K, Ohme M, Tanaka M, Wakasugi T, Hayshida N, Matsubayashi T, Zaita N, Chunwongse J, Obokata J, Shinozaki KY, Ohto C, Torazawa K, Meng BY, Sugita M, Deno H, Kamogashira T, Yamada K, Kusuda J, Takaiwa F, Kata A, Tohdoh N, Shimada H, Sugiura M (1986) The complete nucleotide sequence of the tobacco chloroplast genome: its gene organization and expression. EMBO J 5: 2043 – 2049

Thalouarn P, Theodet C, Russo N, Delavault P (1994) The reduced plastid genome of a non-photosynthetic angiosperm *Orobanche hederae* has retained the *rbcL* gene. Plant Physiol Biochem 32: 233 – 242

Wimpee C, Morgan R, Wrobel R (1992) An aberrant plastid ribosomal RNA gene cluster in the root parasite *Conopholis americana*. Plant Mol Biol 17: 161 – 166

Wolfe A, dePamphilis C (1997) Alternate paths of evolution for the photosynthetic gene *rbcL* in four nonphotosynthetic species of *Orobanche*. Plant Mol Biol 33: 965 – 977

Wolfe K, Morden C, Palmer J (1992) Function and evolution of a minimal plastid genome from a nonphotosynthetic parasitic plant. Proc Natl Acad Sci USA 89: 10 648 – 10 652