Gene expression in plastids of higher plants: evolutionary and functional aspects of different RNA polymerases – coordinated assembly of multiprotein-complexes

Inaugural-Dissertation der Fakultät für Biologie
der Ludwig-Maximilians-Universität München

Gene expression in plastids of higher plants: evolutionary and functional aspects of different RNA polymerases – coordinated assembly of multiprotein-complexes

Inaugural-Dissertation der Fakultät für Biologie
der Ludwig-Maximilians-Universität München

vorgelegt von
Juliana Legen
aus Wolfratshausen
November, 2002

1. Gutachter: Prof. Dr. Reinhold G. Herrmann
2. Gutachter: Prof. Dr. Hans Ulrich Koop

Tag der Mündlichen Prüfung: April, den 8, 2003

1. Introduction

Genetic information in plant cells is found in three different cellular compartments, nucleus, mitochondria and plastids. The latter two are acquisitions of formerly free-living organisms which have been integrated into a host cell by two independent successive endosymbiotic events. Whereas mitochondria trace back phylogenetically to the α-proteobacteria lineage, plastids are descendants of free-living cyanobakterial-like organisms. During evolution, the formerly autarkic gene expression apparatus of plastids - together with the corresponding machineries of mitochondria and the nucleus/cytosol - became the integrated, compartimentalised genetic system of the plant cell in which the 'subgenomes' are regulated syntonically (Herrmann, 1997; Herrmann and Westhoff, 2001).

Due to the eubacterial ancestry of plastids, their genetic machinery in many, but not all, respects resembles that of prokaryotes. During evolution from a free-living organism to the present-day organelle two major trends took place: (i) a massive reduction of the genome size, mainly due to gene loss and gene transfer to the nucleus and (ii) an pronounced increase in the complexity of gene expression.

1.1. Structure and organisation of plastid chromosomes

The genetic potential of plastids is encoded in a reiterated circular chromosome, generally in the range of 120 to 160 kbp, depending on the organism. Sequences of plastid chromosomes of a variety of vascular plants and algae have been completely determined (for a list of plastid chromosomes see **http://www.ncbi.nlm.nih.gov/PMGifs/Genomes/plastids_tax.html**). Some 80 copies of an identical chromosome are found in the single plastid of the unicellular alga *Chlamydomonas* and even 5000 – 10.000 copies in the 50-100 chloroplasts of a mesophyll cell of vascular plants. Thus, with respect to the plastid chromosome plant cells are highly polyploid. The organization of plastid chromosomes and their operons is well conserved in vascular plants. With the exception of some legumes, the circular DNA molecules contain a large inverted repeated segment (IR_A and IR_B) of species-specific sizes encoding the rDNA operon. The repeat regions are separated by a small (SSC) and a large (LSC) single copy region encoding a specific set of conserved operons (Fig. 1).

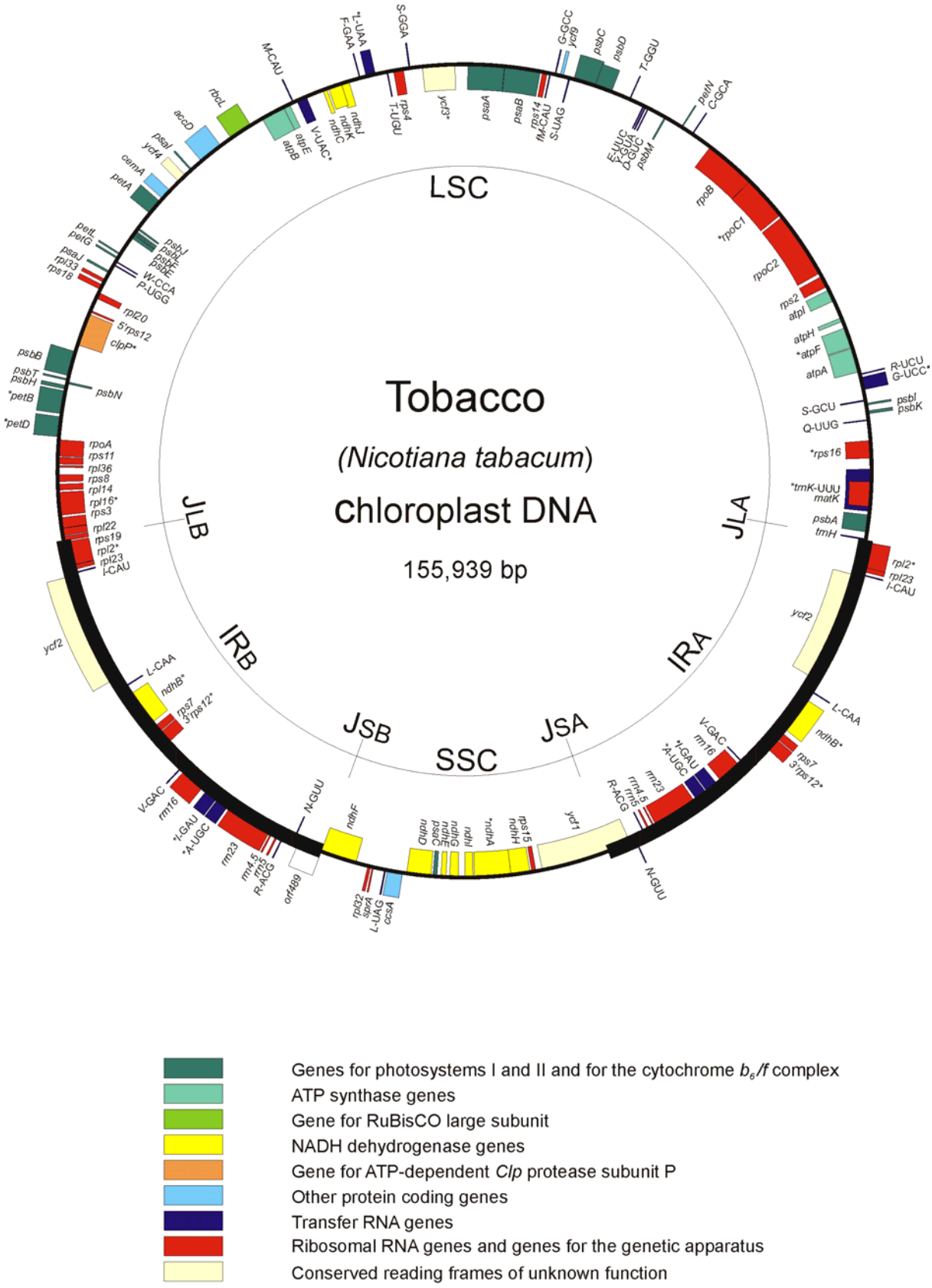


Figure 1. Gene map of the tobacco plastid chromosome (Shinozaki et al., 1986, redrawn by Schmitz-Linneweber, see also Table 1).

Typical plastid chromosomes of vascular plants encode some 120 genes which are usually organised in polycistronic transcription units that are located on both DNA strands (for a review see Sugita and Sugiura, 1996). This corresponds to approximately 5% or even less of the coding potential of the ancestral cyanobacterial genome that may have been in the order of approximately 3.000 genes. During evolution some of the genes got lost, e.g. because they where only useful for the autarcic life style, most genes of the endosymbiont, however, have been transferred to the nucleus and the respective gene products have now to be posttranslationally imported back into the plastids or are found elsewhere in the cell (see Bruce, 2001 and references therein). The genetic information of plastid chromosomes can be divided into three principal classes: (i) genes that contribute to the decoding of genetic information within the organelle, i.e., four genes for subunits of one of the organelle RNA polymerases (see below), genes for ribosomal RNAs and proteins, tRNAs, and factors involved in translational processes, altogether approximately 60 loci, (ii) genes for components of the photosynthetic apparatus, predominantly for constituent polypeptides of the thylakoid membrane complexes (approximately 40 loci), and (iii) open reading frames of unknown function (hypothetical chloroplast reading frames, *ycf*s) (Table 1).

Table 1.

Classification of the genes encoded by the *Nicotiana tabacum* plastid chromosome

Genes coding for RNAs

Genes coding for ribosomal RNAs

rrn23[#], *rrn16*[#], *rrn5*[#], *rrn4.5*[#]

Genes for transfer RNAs

trnA(UGC)[$#], *trnC*(GCA), *trnD*(GUC), *trnE*(UUC), *trnF*(GAA), *trnG*(GCC), *trnG*(UCC)[$], *trnH*(GUG), *trnI*(CAU)[#], *trnI*(GAU)[$#], *trnK*(UUU)[$], *trnL*(CAA)[#], *trnL*(UAA)[$], *trnL*(UAG), *trnfM*(CAU), *trnM*(CAU), *trnN*(GUU)[#], *trnP*(UGG), *trnQ*(UUG), *trnR*(ACG)[#], *trnR*(UCU), *trnS*(GCU), *trnS*(GGA), *trnS*(UGA), *trnT*(GGU), *trnT*(UGU), *trnV*(GAC)[#], *trnV*(UAC)[$], *trnW*(CCA), *trnY*(GUA)

Other RNA genes

sprA

Polypeptide coding genes

Genes coding for ribosomal polypeptides

rps2, rps3, rps4, rps7[#], *rps8, rps11, rps12*[§†], *rps14, rps15, rps16*[$], *rps18, rps19*
rpl2[$#], *rpl14*, *rpl16*[$], *rpl20*, *rpl22, rpl23*[#], *rpl32, rpl33, rpl36*

Genes coding for subunits of the transcriptional apparatus

rpoA, rpoB, rpoC1[$], *rpoC2*

Genes coding for polypeptides of the photosynthesis apparatus

rbcL
psaA, psaB, psaC, psaI, psaJ
psbA, psbB, psbC, psbD, psbE, psbF, psbH, psbI, psbJ, psbK, psbL, psbM, psbN, psbT, psbZ (*ycf9*)
petA, petB[$], *petD*[$], *petG, petL, petN* (*ycf6*)
atpA, atpB, atpE, atpF[$], *atpH, atpI*
ycf3[§], *ycf4*, *cemA (ycf10)*

Genes coding for subunits of the NADH dehydrogenase

ndhA[$], *ndhB*[$#], *ndhC*, *ndhD*, *ndhE*, *ndhF*, *ndhG*, *ndhH*, *ndhI*, *ndhJ*, *ndhK*

Other polypeptide coding genes

accD, *clpP*[§], *ccsA* (*ycf5*)

Conserved reading frames

ycf1, *ycf2*[#], *ycf14* (*matK*; partialy similar to intron maturase)

Pseudogenes

infA, ycf15[#]

[$] Gene with an intron; [§] Gene with two introns; [†] Gene coding *trans*-spliced exons; [#] Gene located in the inverted repeat (IR) regions. For abreviations and nomenclature of genes, see in Stoebe et al. (1998).

Different from the general organisation, only rudimentary 35 - 75 kb chromosomes are found in plastids of organisms that have lost their photosynthetic capability (and genes), such as the apicomplexean parasite *Plasmodium falciparum,* the heterotrophic euglenoid alga *Astasia longa,* and the parasitic Orobanchaean species *Epifagus virginiana* (Siemeister and Hachtel, 1990; Wolfe et al., 1992; Wilson et al., 1996).

1.2. Expression of plastid chromosomes

Contrary to prokaryotes, transcripts of plastid encoded operons are extensively processed into complex sets of overlapping RNA species, finally often into monocistronic RNAs (Barkan, 1988; Westhoff and Herrmann, 1988). Several instances have been noted, where endonucleolytic cleavage of di- or polycistronic transcripts is a prerequisite for efficient translation of mRNAs (Barkan et al., 1994; Hirose and Sugiura, 1997). Essential in higher plant plastids as well are various other transcript maturation processes which are not (i.e. RNA editing) or less commonly (i.e. RNA splicing) associated with prokaryotic systems (Maier et al., 1996; Sugita and Sugiura, 1996). It appears that during organelle evolution distinct RNA processing steps evolved and spread out in the originally cyanobacterial-type genetic system of the ancestral plastid.

In additon to regulatory steps at the level of transcription and transcript processing, the accumulation and assembly of plastid-encoded gene products is controlled at the translational and posttranslational level as well (for a review see Choquet et al., 2000; Zerges, 2000).

1.2.1. Transcription

In analogy to the evolution of posttranscriptional processing, plastid transcription itself developed into a highly sophisticated process, again representing not merely typical prokaryotic features. In particular, transcription in higher plant plastids is performed by (at least) two different RNA polymerases of different phylogenetic origin. Besides the plastid-encoded RNA polymerase (PEP) which resembles eubacterial enzymes, a nuclear-encoded phage-type RNA polymerase (NEP) is involved in transcribing the plastid chromosome (summarised in Hess and Börner, 1999). The core subunit of the PEP enzyme is specified by the four plastid genes *rpoA*, *rpoB*, *rpoC*1 and *rpoC*2 which display significant similarity to the three genes encoding the core of the eubacterial RNA polymerase (summarised in Igloi and Kössel, 1992). Sigma-like factors, which are essential for transcription initiation by PEP, originate in nuclear genes and are posttranslationally imported into the organelle (Allison, 2000). Additional PEP-associated polypeptides of nuclear origin with developement-dependent compositions have been identified in mustard (Pfannschmidt et al., 2000). Whereas PEP traces back to the phylogenetic ancestor of plastids, the phylogenetic origin of NEP is still unclear.

The nuclear genome of higher plants encodes a small family of genes homologous to the RNA polymerase genes of bacteriophages, such as T3 and T7, designated as the NEP gene family. It consists of three members in *Arabidopsis thaliana* (Hedtke et al., 1997, 2000): *Arabidopsis RpoT;*1 with orthologs in the majority of eukaryotic organisms (Cermakian et al., 1996), encodes the core subunit of a phage-type mitochondrial RNA polymerase. A second gene, designated *RpoT*;3 encodes a plastid-targeted phage-type RNA polymerase (Hetdke et al. 1997, 1999). The latter NEP locus seems to have evolved by duplication of *RpoT;*1, acquiring a targeting signal for the new organelle (summarised in Hess and Börner, 1999). Surprisingly, the characterisation of a third NEP gene, designated *RpoT*;2 revealed dual-targeting of the protein to mitochondria as well as to plastids (Hedtke et al., 2000). Consequently, in *Arabidopsis*, both mitochondria and plastids have two NEP-type enzymes at their disposal.

Complicating plastid transcription even more, very recently evidence for a further nuclear-encoded RNA polymerase-type in plastids, that is different from the phage-type NEP enzymes and probably involved in transcription of the rDNA operon has been obtained (Bligny et al., 2000).

The presence of diverse RNA polymerases operating in plastids of higher plants is complemented by the existence of different enzyme-specific promoters and possibly also termination signals preceeding and following plastid transcription units, respectively. The PEP enzyme is known to initiate transcription from -10/-35 eubacterial-type promoters (Igloi and Kössel, 1992). Analyses of transcription in mutants with disrupted PEP genes or in plastids of non-photosynthetic tissue culture cells led to the identification of NEP-specific promoters that are reminiscent of promoters recognised by mitochondrial and T3/T7 phage RNA polymerases (Hajdukiewicz *et al.*, 1997; Kapoor *et al.*, 1997). *In vitro* studies identified a small three nucleotide motif ("CRT") at position –6 to –8 as the NEP promoter core (Liere and Maliga, 1999). Based on analyses of wild-type and PEP-deficient plastids, transcription units have been operationally grouped into three principal classes. Some genes or operons have been suggested to be transcribed by either PEP or NEP, whereas others appeared to be transcribed by both enzymes (Hajdukiewicz et al., 1997). It was proposed that NEP preferentially drives transcription of genes for components of the plastid genetic system, whereas PEP transcribes genes for constituents of the photosynthetic machinery (Hajdukiewicz *et al.*, 1997; Maliga, 1998). Consistently, the accumulation of transcripts of a selected set of photosynthesis-related genes was shown to be reduced dramatically in plants lacking PEP (Allison et al., 1996; Hajdukiewicz et al., 1997). Analysis of run-on transcription activities in PEP-deficient plastids, however, revealed that most segments of the plastid chromosome, independent of the encoded gene class, are transcribed even in the absence of PEP (Krause et al., 2000). Thus, the functional and phylogenetic implications of multiple transcription machineries in the plastids of higher plants is obviously much more complex than initially supposed.

1.2.2. Posttranscriptional RNA processing

Plastid-encoded RNAs are generally subject to a cascade of posttranscriptional processing steps including endonucleolytic cleaveage of polycistronic RNA chains and

complex maturation of the transcript ends (for review see Sugita and Sugiura, 1996; Barkan and Goldschmidt-Clermont, 2000; Monde et al., 2000). In addition, from a number of transcripts introns have to be removed by RNA splicing. Introns are widespread in higher plant plastid chromosomes. Spinach for instance contains 19 intervening sequences spread over 16 plastid-encoded genes (Schmitz-Linneweber et al., 2000). With the exception of one (that in the tRNA-Leu gene represents a group I intron) all these introns belong to group II (Michel and Ferat, 1995). Group II intervening sequences, which have also been found to interrupt reading frames in fungal and plant mitochondria, rarely also in bacteria (Ferat and Michel, 1993), fold into a characteristic, evolutionary conserved secondary structure. Although some of these introns have been reported to exhibit self-splicing activity *in vitro*, splicing of plastid group II introns involves additional, in individual cases intron-specific splicing factors (Jenkins et al., 1997). With the exception of the plastid-encoded *matK* reading frame, which exhibits significant sequence similarity to fungal genes for intron maturases, all splicing factors must originate in nuclear genes (Vogel et al., 1999). Although both plastid-type group I and group II introns have been found in cyanobacteria (Thompson et al., 1995; Besendahl et al., 2000), the majority of them in plastids are thought to have been lost and gained at several branches during cell evolution. Whereas group II introns are not (any more?) present in plastids of most algae lineages, they appear to be widespread in those of euglenoids and vascular plants (Thompson et al., 1995).

Maturation of plastid transcripts frequently includes posttranscriptional modifications of individual nucleotides within an mRNA by a process referred to as RNA editing. Editing sites in plastid chromosomes are generally found in reading frames, in some cases even in translation start and stop codons. The only few editing sites found in intergenic regions are all placed upstream to start codons. It is conceivable that editing at those positions can lead to the restoration of an RNA structure facilitating translation of the downstream reading frames (Drescher et al., 2002). Of the different types of RNA editing found in various organisms, to date only C-to-U conversions, and in few cases also U-to-C conversions are known to affect plastid transcripts (Maier et al., 1996; Bock et al., 1997). Remarkably, editing in plastids displays significant similarities to editing of plant mitochondrial transcripts (Maier et al., 1996). In both cases, site-specific changes from C-to-U and rarely also from U-to-C occur in RNAs transcribed from the organellar chomosomes. These changes take place predominantly at the second position of

codons. The number of editing sites, however, is much lower in plastids. Circumstantial evidence suggests that the process in both plant organelles is evolutionary linked. In both organelles, editing has been found to be widespread within vascular plants occuring in all major groups, and also in Bryophyta (Malek et al., 1996; Freyer et al., 1997). An unusual degree of U-to-C conversions can be found in both, mitochondria and plastids, of the hornwort *Anthoceros formosae* (Yoshinaga et al., 1996), and editing appears to be absent in both organelles of the liverwort *Marchantia polymorpha*. Editing apparently also does not exist in plastids of algae nor in cyanobacteria, representing the phylogentic ancestor of plastids. Taken together, this has suggested that the probably common phylogenetic origin of RNA editing in both plant organelles resides in the early ancestors of land plants. The higher frequency of editing in plant mitochondria, in turn, indicates that editing has entered the genetic system of the organelles successively. If correct, this would be an interesting analogy to the nuclear-encoded RNA polymerase (NEP). However, so far no plant species has been found in which editing occurs in only mitochondria.

In addition to endonucleolytic cleavage of polycistronic transcripts, RNA splicing and editing, RNA stabilisation/destabilisation has been discovered as a further regulatory point in plastid gene expression (for review see Monde et al., 2000). It was found that RNA stability (defined as the time required for half the initial amount of the RNA measured to disappear) varies significantly between individual transcripts and that relative transcription rates and accumulation of a given transcript often do not correlate. In many cases fluctuation in transcription rates do not account for the observed changes in RNA amounts. Sequence elements located in the 5' and 3' UTR as well as in coding regions have been shown to play a critical role in the stability/instability of individual transcripts. Mechanistically, in the initial step RNA chains are endonucleolytically cleaved at specific sites. The cleavage products are then polyadenylated at the 3' end followed by 3'–to–5' degradation and/or exonucleolytical degradation from their 5' termini (reviewed in Hayes et al., 1999; Schuster et al., 1999). The components involved are all nuclear-encoded (Monde et al. 2000 and references therein). It has been shown that the half-life of individual plastid RNAs is regulated differentially in a development-specific manner. For example, compared to spinach grown in dark a significantly increased stability of *psbA* mRNA is observed in light-grown plants (Deng and Gruissem, 1987).

1.2.3. Translational control of plastid gene expression

Translation in plastids exhibits many eubacterial-type characteristics, like 70S ribosomes and formylated initiator tRNAs (fMet-tRNA). However, only few plastid mRNAs contain Shine-Dalgarno (SD)-like sequences necessary for translation initiation in eubacteria, and if they do, it is often shifted to a more upstream position, untypical for eubacterial translation systems (Sugiura et al., 1998). In some instances, novel sequence elements not found in prokaryotes seem to replace or extend the function of SD sequences (e.g. Hirose and Sugiura, 1996). Plastid translation has been shown to be regulated in response to light and developmental programs (for review see Zerges, 2000). *Trans*-acting translational factors are encoded by the nucleus and interact with *cis* elements located in the 5' and 3' UTRs of plastid mRNAs. Only little is known of how translation of mRNAs is regulated mechanistically. As in the case of transcription rates and accumulating RNA amounts, frequently no direct correlation exists between mRNA amounts and corresponding polypeptides. It has been demonstrated that under certain developmental conditions substantial amounts of thylakoid proteins are translated, even from only minute mRNA amounts (Herrmann et al., 1992).

The plastid genome of tobacco encodes 30 tRNA species that can read all sense codons by expanded wobble codon-anticodon recognition. Therefore, there is principally no need for import of tRNAs from the cytoplasm, in contrast to plant mitochondria (Sugita and Sugiura, 1996). The finding that parasitic plants, like *Epifagus virginiana*, encode only 13 tRNAs in their reduced plastid genome, but nevertheless are able to translate mRNAs containing all codons (Morden et al., 1991; Taylor et al., 1991), however, suggests that at least in those plants nuclear-encoded tRNAs are involved in plastid translation as well.

1.3. Thylakoid membrane-located multiprotein complexes

The thylakoid membrane of plastids contains four major multiprotein complexes, involved in photosynthetic electron transfer and ATP synthesis: photosystem I and II, the cytochrome b_6f complex and the ATP synthase (Fig. 2). In addition, a fifth multiprotein complex, the NADH dehydrogenase, suggested to be involved in protection against photooxidative stress, is located in thylakoid membranes of vascular plants (Casano et al., 2000). These complexes are built from polypeptides

either encoded in the nucleus or in the plastid genome which have to be assembled into multiprotein complexes in defined stoichiometry. Assembly takes probably place at the site of function, in the thylakoid membrane, and demands highly coordinated mechanistic steps.

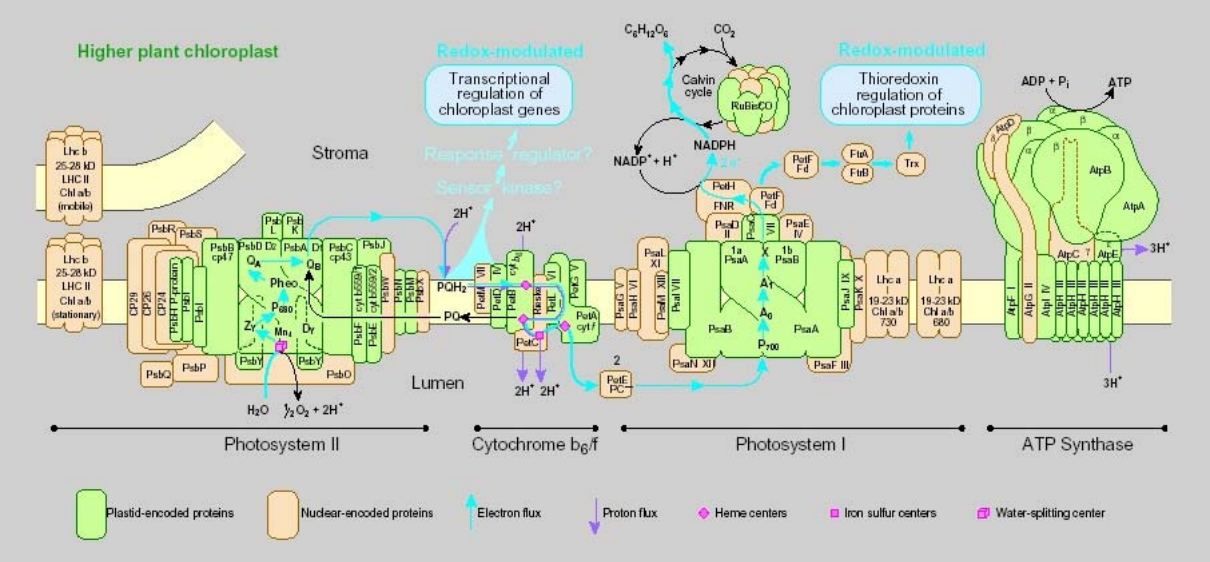


Figure 2. Scheme of the thylakoid membrane system (concept for scheme kindly provided by Prof. Herrmann). The thylakoid membrane system contains four major complexes stoichiometrically related to each other. A fifth minor abundant complex, the NADH-dehydrogenase complex is not shown. Each complex consists of nuclear (yellow) and plastid encoded (green) components.

1.3.1. Synthesis and assembly of the cytochrome *b*$_6$*f* complex

The cytochrome *b*$_6$*f* complex is the simplest of the thylakoid membrane protein complexes in terms of polypeptide composition. It consists of at least 8 polypeptide species (encoded by the genes *petA*, *petB*, *petC*, *petD*, *petG*, *petL*, *petM* and *petN*) of which in vascular plants six (genes: *petA*, *petB*, *petD*, *petG*, *petL* and *petN*) are encoded in the plastid genome. Nuclear-encoded genes *petC* (coding for the Rieske protein) and *petM* (coding for subunit VII) are transcribed and translated from individual mRNAs in the nucleo–cytoplasmatic compartment. The respective polypeptides are imported into plastids via N-terminal targeting signals. Plastid-encoded *pet* genes are scattered in four transcription units: *petA* (coding for cytochrome *f*) is transcribed as part of a tricistronic *ycf4/cemA/petA* message, whereas *petB* and *petD* (coding for cytochrome *b*$_6$ and subunit IV, respectively) are transcribed as part of the *psbB* operon (*psbB/psbT/psbH/petB/petD*). Both, *petB* and *petD*, contain group II introns which have to be removed by splicing as a prerequisite for expression of the encoded polypeptide. In addition, *petB* of tobacco and other vascular plants contains an editing site, at which a proline codon (CCA) has to be

posttranscriptionally changed to a leucine codon (CUA) by C-to-U editing to encode a functional cytochrome b_6 polypeptide (Freyer et al., 1993; Zito et al., 1997). *petL* and *petG* (coding for subunit VII and V, respectively) are transcribed as a dicistronic message whereas *petN* (encoding subunit VI polypeptide) encodes a probably monocistronic mRNA (see Fig.1).

Various mutants impairing assembly of the cytochrome b_6f complex of the green alga *Chlamydomonas reinhardtii* and higher plants have been investigated (Lemaire et al., 1986; Bruce and Malkin, 1991; Voelker et al., 1995; Monde et al., 2000a,b). It has been shown that deletions of individual plastid-encoded *pet* genes result in a dramatic, by a factor ten or more, decrease of the accumulation of the remaining subunits (Choquet and Vallon, 2000). Two types of mechanism have been suggested to be responsible for the concerted accumulation of cytochrome b_6f subunits: (i) degradation of unassembled polypeptides, and (ii) regulation of translation by assembly (Choquet and Vallon, 2000). In *Chlamydomonas*, for instance, the half-life of subunit IV and cytochrome b_6 is dramatically decreased upon deletion of another subunit, whereas the synthesis rates remained unchanged. This suggests that unassembled polypeptides are rapidly degraded. Different from that, in the absence of its assembly partners, the synthesis rate of cytochrome *f* is decreased to approximately 10% of wild-type levels. The low amount of cytochrome *f* synthesised, however, is inserted into the thylakoid membrane and is as stable as in wild-type. This indicates a hierarchical organisation of the expression of cytochrome b_6f subunits, which has been termed as control by epistasy of synthesis (CES; Wollman et al., 1995). According to this, cytochrome *f* requires the presence of its dominant assemby partners to be synthesised at wild-type rates. Mechanistically, the C-terminal domain of cytochrome *f* seems to be involved in autoregulation of *petA* mRNA translation. Recent studies on thylakoid membrane complexes in *Chlamydomonas reinhardtii* showed that CES regulation is a general key feature in the assembly of photosystem I and II and the ATP synthase as well (F.A Wollman, personal communication).

1.4. Goals of this study

Plastid gene organisation maintains characteristics typical of its prokaryotic ancestry. The regulation of plastid gene expression however strongly deviates from that one its free-living cyanobacteria relatives. This intriguing complication of plastid gene expression characteristics is the result of an integration process of the cellular subgenomes that introduced eukaryotic traits into the formerly prokaryotic compartment (Herrmann et al. 1997). An interesting example for this process is the transcription system, which consists of both prokaryotic (PEP) and eukaryotic (NEP) RNA polymerases. In order to understand, how far the transcriptional apparatus within plastids was adapted to nuclear needs, three approaches have been undertaken.

Firstly, the tobacco homologues of the NEP-enzymes known from *Arabidopsis* were determined and characterised.

Secondly, an extensive transcript analysis for all plastid operons was carried out with wild-type and PEP-lacking material in order to assess the contribution of the two systems to transcription. In order to rapidly screen this plant material, an array-based technique was established. As no arrays for the plastid chromosome had been described, the preparation of the filters, optimisation of hybridisation conditions and probe preparation together with the use of the proper controls was one of the main challenges of this work, in particular as macroarrays are interesting for various other applications.

Thirdly, to assess the regulation of PEP expression, an *in vivo* analysis of the promoter for the *rpoB* operon was performed. According to data from other groups, this operon is controlled by NEP, which – in addition to other data – led to the suggestion that PEP is switched on by NEP (Hajduckiewicz et al., 1997). To test this, point mutations of the described NEP *rpoB* promoter (Liere and Maliga; 1999) should be prepared and introduced into the plastid compartment by particle transformation. Insight into the regulation of PEP is crucial in order to understand the interplay of the two RNA polymerase systems.

Finally, in addition to studies on plastid transcription, later steps in plastid gene expression were examined in the course of this work as well. As a model to study assembly of a multisubunit complex, the cytochrome b_6/f complex was analysed by a gene-disruption approach. Deletion and/or insertion mutants of all plastid encoded subunits of the complex were prepared in order to evaluate the assembly strategy for the holocomplex.

2. Materials

2.1. Chemicals and enzymes

Chemicals used in this study that are not listed in the table below were purchased from: Amersham-Pharmacia (Freiburg, Germany), Roche GmbH (Basel, Switzerland), Invitrogene (Karlsruhe, Germany), Biomol (Hamburg, Germany), New England Biolabs (Bad Schwalbach, Germany) or MBI Fermentas (Vilnius, Lithuania).

acrylamid powder 2x	Serva; Heidelberg, Germany
agar, purified	Sigma, Heidelberg; Germany
agarose	BioWhittaker Molecular apl.; Rockland,USA
ammoniumpersulfate	Sigma; Heidelberg; Germany
amoniumsulfate	Roth, Karlsruhe; Germany
6-aminocapronic acid	Sigma; Heidelberg; Germany
ampicillin	Serva, Heidelberg; Germany
Bacto Agar	ICN, Biomedicals Inc.; Ohio, USA
Bacto Tryptone	USB, Cleveland; USA
BAP	Sigma-Aldrich; St. Luis; USA
bis-Tris	Roth, Karlsruhe; Germany
bromphenolblue	Merck, Darmstadt; Germany
beef extract powder	Sigma, Heidelberg; Germany
BSA, fraction V	Roth, Karlsruhe; Germany
casein hydrolysate	Merck; Darmstadt; Germany
tri-chloracidic acid	Roth, Karlsruhe; Germany
coomassie brilliant blue, R250	Merck; Darmstadt; Germany
DEAE-cellulose membrane	Schleicher and Schuell; Dassel, Germany
1,4-dithiothreitol	Roth, Karlsruhe; Germany
ethidiumbromide	Roth, Karlsruhe; Germany
formaldehyd	Roth, Karlsruhe; Germany
formamide	Sigma, Steinheim; Germany
glycerol	Roth, Karlsruhe; Germany
hexadecyltrimethyl-amonium bromide	Sigma; Heidelberg; Germany
HEPES	Roth, Karlsruhe; Germany
Hybond-N-membrane	Amersham-Buchler; Braunschweig; Germany
IPTG	Sigma-Aldrich; St. Luis; USA
MES	Serva; Heidelberg; Germany
β-mercaptoethanol	Sigma, Heidelberg; Germany
mannitol	Roth, Karlsruhe; Germany
Miracloth-filter	Calbiochem; Merck, Darmstadt; Gremany

MOPS	Roth, Karlsruhe; Germany
NAA	Sigma-Aldrich; St. Luis; USA
di-sodium hydrogenphosphate	Merck, Darmstadt; Germany
sodium dodecylsulfate	Serva, Heidelberg; Germany
sodium chlorid	Roth, Karlsruhe; Germany
sorbitol	Roth, Karlsruhe; Germany
spectinomycin	Sigma-Aldrich, St. Luis; USA
percoll	Pharmacia; Uppsala; Sweden
phenol/chloroform	Roth, Karlsruhe; Germany
Polyethyleneglycol 6000	Serva; Heidelberg; Germany
Ponceau S	Serva; St. Luis; USA
rifampicin	Fluka; Steinhein; Germany
Qiagen colums	Qiagen; Hilden; Germany
TEMED	Serva; Heidelberg; Germany
Tris	Roth, Karlsruhe; Germany
Tris-sodium citrate 2-hydrate	Roth, Karlsruhe; Germany
Triton X-100	Roth; Karlsruhe; Germany
Tween 20	Applichem, Darmstadt; Germany
X-gal	Sigma-Aldrich, St. Luis; USA
xylenecyanol	Serva; Heidelberg; Germany
yeast extract	Merck, Darmstadt, Germany

2.2 Length and weight standards

A 1 kb DNA length standard 54 to 8000 bp was purchased from Invitrogene, Karlsruhe. In addition, λ DNA restricted with *EcoRI* and *HindIII* yielding fragments between 564 to 21226 bp was used. An RNA molecular weight standard (0.24-9.5 kb) was purchased from Invitrogene, Karlsruhe. As a standard for the determination of the molecular weight of proteins, the marker SDS-7 (Pharmacia) was used.

2.3. Antibodies

Polyclonal goat anti-rabbit IgGs conjugated to alkaline phosphatase and a goat anti-rabbit IgG conjugated to peroxidase were obtained from Sigma Chemical Company. Further antibodies were produced in the laboratory of Prof. R.G. Herrmann, with exception of αPetL which was kindly provided by Prof. J.D. Rochaix (Geneva, Switzerland).

2.4. Bacterial strains

E. coli DH5α	Hanahan (1983)
E. coli Y1090	Young and Davis (1983)
E. coli CJ236	Invitrogen (Karlsruhe, Germany)

2.5. Plasmids

pBlueScript (SK^- and SK^+)	Stratagene (La Jolla, USA)
pPCR-TA	Invitrogen (Karlsruhe, Germany)

2.6. Oligonucleotides

All oligonucleotides used for macroarray preparations, cloning, or sequence analyses have been obtained from MWG-Biotech (Ebersberg, Germany). The list of primers is given in an additional table at the end of this thesis.

2.7. Plant growth media

B5 medium was prepared according to Gamborg et al. (1968), protoplast culture media (PCN and PIN) according to Koop and Kofer (1995). RMOP medium for shoot regeneration was prepared according to Svab and Maliga (1993).

Table 2. Composition of plant growth media (adapted from Brunner, 1997).

	mg/l	B5	F-PCN	F-PIN	RMOP
macro salts	NH_4NO_3				1650
	KNO_3	2500	1012	1012	1900
	$CaCl_2$ $2H_2O$	150	440	440	440
	$MgSO_4$ $7H_2O$	250	370	370	370
	KH_2PO_4		170	170	170
	NaH_2PO_4 H_2O	150			
	$(NH_4)_2SO_4$	134			
	NH_4-Succinat		0.2 mM	0.2 mM	
micro	EDTA-Fe(III)Na	40	40	40	40
	KJ	0,75	0,75	0,75	0,83
	H_3BO_3	3	3	3	6,2

	$MnSO_4$ H_2O	10	10	10	22,3
	$ZnSO_4$ $7H_2O$	2	2	2	8,6
	Na_2MoO $2H_2O$	0,25	0,25	0,25	0,25
	$CuSO_4$ $5H_20$	0,025	0,025	0,025	0,025
	$CoCl_2$ $6H_2O$	0,025	0,025	0,025	0,025
vitamines	Inositol	100	200	200	100
	Pyridoxin-HCl	1	2	2	
	Thiamin-HCl	10	1	1	1
	Biotine		0,02	0,02	
	Nicotinic acid	1	2	2	
other	BAP		1	1	1
	NAA		0.1	0.1	0.1
	Polybuffer 74		10 ml	10 ml	
	Sucrose	20 000		130 000	30 000
	Glucose		85 000		
	pH (KOH)	5,7	5,8	5,8	5,8
	Osmolarity		550 (mosm/l)	550 (mosm/l)	
	Agar	7000			8000

2.8. Plant material

For plant transformation experiments as well as for biochemical and biophysical analyses tobacco plants (*Nicotiana tabacum* L. cv. Petit Havana) were grown axenically under standard light conditions (100 µE m^{-2} s^{-1}) or under lower (ca. 4 - 10 µE m^{-2} s^{-1}) light regimes generally at 20°C - 25°C, and 16/8 h light/dark cycles (Osram L85W/25 Universal White fluorescent lamps). For phenotypic comparison, plants were grown in soil under greenhouse conditions.

2.9. Media for growth of bacteria

All values refer to 1 liter of medium:

LB medium	10 g	Bactotryptone
	5 g	yeast extract
	5 g	NaCl

Plates used for growth of bacteria were prepared by adding 15 g Agar (Difco, Kansas City, USA). All media containing antibiotics were prepared by adding the antibiotic from the stock solutions after sterilisation and cooling down to approx. 50 ^oC.

2.10. Antibiotic stock solutions

ampicilline	40 mg/ml	dissolved in water
chloramphenicol	40 mg/ml	dissolved in 70% ethanol

2.11. General buffers and solutions

TE Buffer	10 mM	Tris/HCl
pH 8.0	1.0 mM	EDTA
10xTBE Buffer	1.0 M	Tris
	0.5 M	boric acid
	20 mM	EDTA
50xTAE Buffer	2.0 M	Tris/HCl
pH = 7.4-7.6	1.0 M	sodium acetate
	20 mM	EDTA
10xMOPS Buffer	200 mM	MOPS
pH 7.0	50 mM	NaOAc
	1 mM	Na_2EDTA
10xSDS-PAGE Buffer	250 mM	Tris/HCl
pH 8.3	14.4% (w/v)	Glycin
	1% (w/v)	SDS
20xSSC	3 M	NaCl
pH 7.2	0.3 M	tri-sodium citrate
10xPBS Buffer	750 mM	NaCl
pH 7.5	30 mM	KCl
	45 mM	Na_2HPO_4
	15 mM	KH_2PO_4
Denaturation Solution	1.5 M	NaCl
	0.5 M	NaOH
Neutralisation Solution	1.5 M	NaCl
pH 7.0	0.5 M	Tris/HCl

3. Methods

I. Manipulation of Nucleic Acids

3.1. Isolation of nucleic acids

3.1.1. Small-scale plasmid isolation from *E. coli* (Miniprep)

3 ml LB/Amp-Medium (100 µg/ml ampicillin) was inoculated with a single colony and incubated overnight at 37°C with constant agitation. Cultures were transferred into 2 ml Eppendorf tubes and cells were pelleted by centrifugation (14.000 g, 1 min, RT). Plasmids were isolated from bacterial cells using anion exchange columns (Qiagen) according to the manufacturer's protocol. DNA was eluted from the columns by addition of 50 µl 10 mM Tris/HCl, pH 8.0, with subsequent centrifugation (14.000 g, 2 min, RT).

3.1.2. Isolation of genomic DNA

Isolation of DNA from plants was performed according to Doyle and Doyle (1990) using CTAB as detergent. The leaf material (approx. 500 mg) was ground in liquid nitrogen. DNA extraction was performed using 400 µl of the CTAB-buffer and continued grinding. Afterwards, the slurry was incubated at 65^{o}C for 30 - 60 min. The reaction was centrifuged to remove cellular debris (12.000 g; 10 min) and DNA precipitated from the supernatant with isopropanol (4^{o}C; 30 min;18.000 g).

CTAB-buffer

2%	CTAB
1.4 M	NaCl
20 mM	EDTA, pH 8.0
100 mM	Tris/HCl, pH 8.0
100 mM	β-mercaptoethanol (added before use)

3.1.3. Isolation of RNA

Total cellular RNA was isolated from plant tissue using TRIzol reagent (Invitrogene, Karsruhe, Germany) following the supplier's protocol. The concentration of isolated RNA was checked by spectroscopic measurement (in Section 2.1.5). As a quality control of isolated RNA, aliquots (2 µg) were fractionated on formaldehyde-containing denaturing gels with ethidiumbromide in the RNA sample to visualise rRNAs.

3.1.4. Elution of nucleic acids from agarose gels

DNA-fragments were recovered from agarose gels by electro-elution onto DEAE membranes using NET Buffer (Sambrook et al., 1989).

NET Buffer

1.0 M	NaCl
0.1 M	Na_2EDTA
20 mM	Tris/HCl, pH 8.0

3.1.5. Determination of nucleic acid concentrations

DNA and RNA concentrations were determined spectroscopically using an Amersham-Pharmacia (Ultraspec 3000) spectrometer. The absolute volume necessary for measurement was 100 µl. Concentrations were determined by measuring the absorbances at 260, 280 and 320 nm. A ratio of A260/A280 between 1.8 and 2 indicated a sufficient purity of the nucleic acid preparation.

3.2. Enzymatic modifications of nucleic acids

3.2.1. Restriction analysis of DNA

For restriction, the DNA was incubated with twice the recommended amount of appropriate enzymes in the recommended buffer for 2 h. Restriction was terminated by addition of sample buffer and the sample was applied onto an agarose gel. If the conditions for two enzymes were incompatible with each other, the DNA was digested successively with the respective enzymes. The DNA was purified between the two digestion assays using the rapid purification kit (Life Technologies).

3.2.2. Ligation of DNA fragments

Ligation of DNA fragments was performed by mixing 50 ng vector DNA with a fivefold molar excess of insert DNA. 1 µl of T4 DNA ligase and 2 µl of 10x ligation buffer were added and the reaction mix was brought to a final volume of 20 µl. The reaction was incubated either for 2 h at RT or overnight at 16°C. The assay was used directly for transformation of *E. coli* cells without any further purification.

3.2.3. Dephosphorylation of plasmid DNA

After restriction, plasmid DNA was purified by the gel purification method (described in Section 3.1.4.). Then SAP buffer (Boehringer Ingelheim) and 1 U SAP (shrimps alkaline phosphatase) per 100 ng plasmid DNA were added and the reaction was incubated at 37°C for 2 h and terminated by incubation at 70°C for 10 min. The plasmid DNA was used for ligation without further purification.

3.2.4. RNAse A treatment of DNA preparations

RNAse A treatment of DNA preparations to eliminate RNA contamination was performed by adding 1 µg/ml of RNAse A. The reaction was incubated for 30 min at 37^{o}C with a subsequent phenol/chlorophorm purification step in order to remove the enzyme.

3.2.5. Synthesis of cDNA

RNA preparations were treated with RNase-free DNase I, extracted with phenol-chloroform and precipitated with 2 volumes of ethanol in the presence of 0.3 M sodium acetate, pH 4.8. Aliquots of 3 µg plastid RNA were reverse transcribed with Moloney murine leukemia virus RNase H-free reverse transcriptase (Superscript^{TM}, GIBCO/BRL, USA). The reactions were either primed with random hexanucleotides or sequence-specific oligonucleotides. The reaction was terminated by addition of 0.2 volumes of 0.5 M EDTA, pH 8.0. To remove RNA, the product mixture was ethanol precipitated, and the pellet was incubated in 0.3 M NaOH, 5 mM EDTA, pH 8.0, for 30 min at 65°C. The sample was then neutralized by addition of 1.2 volumes of 1 M Tris/HCl, pH 7.5.

For preparation of labelled cDNA probes for macroarray hybridisation, reverse transcription was carried out in the presence of α-^{32}P-dCTP. As primers, either random hexaoligonucleotides or a mixture of 129 twenty-base oligodesoxynucleotides (8 pmol each), each complementary to the coding strand of one of the plastid-encoded genes as well as of several open reading frames were used.

3.2.6. Primer extension analysis

Primer extension reactions were carried out with 50 µg of total leaf RNA with Superscript^{TM} reverse transcriptase (GIBCO/BRL, USA) and a fluorochrome-labeled primer (5'-IRD700-TGTATCTGATTAAATCCAGGTATTG-3'; MWG Biotech, Ebersberg, Germany) annealing to the region between +55 and +31 relative to the first nucleotide of the *rpoB* coding sequence. DNA sequences were generated by using the same primer and the Thermosequenase kit (Amersham Pharmacia Biotech, Freiburg, Germany). Products were analysed with the LI-COR 4200IR^2 two-laser system (MWG Biotech, Ebersberg, Germany).

3.2.7. 5' and 3' RACE reactions

RACE PCRs were carried out using SMART RACE (Clontech, Heidelberg, Germany) as well as GeneRacer (Invitrogen, San Diego, CA, USA) protocols, following the manufacturers instructions.

3.2.8. Polymerase chain reaction (PCR)

Amplification of DNA fragments was performed in a 50 µl reaction mixture with thin-walled PCR tubes in a Hybaid PCR cycler. The following reaction mixture was used:

template: 2 - 10 ng
primer 1 (10 pM): 1 µl
primer 2 (10 pM): 1 µl
nucleotides (dNTPs; 20 mM): 1 µl
PCR buffer (10 x): 5 µl
Taq polymerase: 2.5 U
H_2O ad 50 µl

The PCR was performed with the following steps, if not otherwise stated:
1) Initial denaturing at 94°C for 2 min,
2) denaturing at 94°C for 0.5 min,
3) annealing at Tm-4°C for 0.5 min,
4) synthesis at 72°C for approx. 1 min/1kb DNA,
5) termination at 72°C for 5 min,
6) cooling to 4°C.

The amplification procedure (steps 2 - 4) was repeated 30 times.

The melting temperature of the primers depends on their GC content and was calculated by the following formula:

$Tm = n\ (G+C) \times 4\ ^{o}C + (A+T) \times 2\ ^{o}C$

If the two primers chosen had different melting temperatures, the lower one was used. The quality of PCR products was monitored by gel electrophoresis.

3.2.9. Sequencing of DNA

Nucleotide sequences were determined by the dideoxy chain termination method (Sanger et al., 1997) using gene-specific primers with an ABI 377 sequencer (Applied Biosystems, USA). Sequences were evaluated and assembled using Sequencher 3.0 (Gene Codes Corporation, USA). Sequences were aligned using the BioEdit sequence alignment editor (North Carolina State University).

3.2.10. Gelelectrophoresis of nucleic acids

For electrophoresis of DNA fragments agarose gels from 0.5% to 2% in 1x TBE were used. For electrophoresis of *in vitro* transcripts used in macroarray studies 4.5% polyacrylamide/urea gels were chosen. For electrophoresis of RNA 1 - 1.2% denaturing formaldehyde agarose gels were employed (Sambrook et al., 1989).

3.2.11. Southern analysis of DNA

For Southern analysis 5 µg of total DNA was restricted with the desired enzyme and electrophoretically separated in an agarose gel. DNA transfer from the gel onto nitrocellulose membranes was performed by capillary blotting as described in Sambrock et al. (1989). As a quality and loading control, the membrane was stained with methylene blue solution for 30 sec and destained with deionized water for 10 min.

3.2.12. Northern analysis of RNA

5 µg aliquots of RNA from either isolated plastids or leaf material were electrophoresed in 1% agarose gels containing 1x MOPS buffer and 2%

formaldehyde, and transferred onto nylon membranes by capillary blotting. The blot is blotted as described in Sambrook et al., 1989.

Denaturing formaldehyde-containing RNA gel (for 120 ml)

1% (w/v)	agarose
90 ml	water
12 ml	10x MOPS, pH 7.0
20 ml	formaldehyde (>37%) (Roth, Karsruhe, Germany)

Running buffer (for 500 ml)

50 ml	10x MOPS, pH=7.0
50 ml	formaldehyde (>37%) (Roth, Karsruhe, Germany)
400 ml	water

Sample buffer

1.2 ml	formamide (deion., Sigma-Aldrich, St. Louis, USA)
0.4 ml	formaldehyde (>37%) (Roth, Karsruhe, Germany)
0.24 ml	10x MOPS, pH 7.0

with addition of 1 µl of 100x Bromphenol-Blue solution (0.1 mg of Bromphenol-Blue solubilized in 1 ml water).

Methylene-blue staining solution (for 100 ml)

0.03% (w/v)	methylene blue
0.3 M	sodium acetate, pH 5.2

3.2.13. RNA ligation

Ligation of oligonucleotides to transcripts was used to detect 5' or 3' ends of distinct mRNA species. The principle of the method is based on the catalytic reaction of T4 RNA ligase that forms intramolecular phosphodiester bonds between 5' phosphate and 3' hydroxyl termini of single-stranded RNA. Its minimal substrate is a 3' ribonucleoside and 5' phosphate group. For RNA ligation DNAse I-digested mRNA was used, to keep the reaction specific. For low expressed genes, 50 µg of total plastid RNA was used in a reaction of 50 µl volume. The ligation reaction was perfomed over night at 4 - 10 ^{o}C, according to the supplier.

3.2.14. Preparation of radiolabeled probes

3.2.14.1. Radioactive labeling of PCR products

PCR products (100 ng DNA) were radiolabeled by random priming with Klenow enzyme and ^{32}P-dATP or ^{32}P-dCTP according to Sambrook et al. (1989). For this assay, the Random Primed DNA Labeling Kit from Roche (Basel) was used following the manufacturer's protocol.

3.2.14.2. Radioactive labeling of oligodesoxynucleotides

The 5' hydroxy group of synthetic oligonucleotides (primers) was radioactively labeled by 30 min incubation at 37°C with T4 polynucleotide kinase (New England Biolabs, Bad Schwalbach). Residual radioactive mononucleotides were removed by ethanol precipitation.

Assay mixture (25 µl)

0.5 µg oligodesoxynucleotides
2.5 µl 10x PNK buffer
7 µl [γ-^{32}P]ATP
1 µl T4 polynucleotide kinase

3.2.14.3. Preparation of radiolabeled RNA probes

To obtain a probe for strand-specific hybridisation, PCR with a single primer containing 5' the T7 promoter sequence was performed. Approximately 100 ng of the PCR product was used for *in vitro* transcription in the presence of [γ-^{32}P]ATP. This resulted in a radiolabeled antisense transcript for the desired RNA.

Radioactive *in vitro* transcription mixture (20 µl)

2 µl	10x transcription buffer (NEB)
2 µl	100 mM DTT
0.5 µl	RNasin (RNase inhibitor, Promega)
4 µl	rNTP mix-UTP (2.5 mM per base)
2.4 µl	100 µM UTP
x µl	PCR product (100 ng)
5 µl	[α-^{32}P]-UTP
1 µl	T7 RNA polymerase (20 U/µl)

The transcription reaction was carried at 37°C for 2 h. Subsequently, 1 µl of DNase I was added to the mix and the sample was incubated for 15 min at 37°C. After incubation, 80 µl of sterile water was added and two cycles of chloroform extraction were performed. RNA was precipitated from the aqueous phase by adding 2.5 volumes of 100% ethanol and 0.1 volume of Na acetate, incubated for at least 30 min at –20°C. The RNA pellet was washed once with 70% ethanol, dried and dissolved in RNAse-free water. Before hybridization the radioactive RNA sample was denatured for 5 min at 65°C.

3.3. Hybridization procedure

Both, nucleic acids on Northern and Southern membranes were fixed by UV crosslinking (2 x Autocrosslink on "UV-Stratalinker^{TM} 2400", Stratagene) and prehybridized at 65°C for at least 2 h in Church buffer. Hybridization was performed by incubation of the membrane with the denatured radiolabeled probe in 10 ml of Church buffer overnight at 65°C.

Afterwards, the filter was washed with a series of solutions containing decreasing amounts of salts (1x to 0.05x SSC, 0.1 SDS) in order to increase hybridization stringency and thereby eliminating non-specific probe binding. Finally, the filters were sealed in a plastic foil and exposed to X-ray film (Biomax film, Fujifilm, Belgium). Alternatively, autoradiographs for both RNA and DNA analyses were evaluated using Fuji Bio Imaging plates type BASIII, a Fuji Bio Imaging analyzer, the BAS200025 software package (Fuji) and the TINA20 software package v2.08 beta (Raytest, Spröckhovel/Germany).

3.4. Dot-blot DNA/RNA hybridisation analysis

In oder to detect differences between single-stranded and double-stranded probes and to test the sensitivity of macro arrays, dot-blot analysis was performed. 1 - 5 µg of total RNA or 7 µg of *in vitro* transcripts were mixed with RNA sample buffer and denatured at 65^{o}C for 5 - 8 min. The samples are dotted onto pre-wet (in 5x SSC) positively charged nylon membranes (Hybond N+, Amersham, Freiburg, Germany). The soaking effect is supported by using Whatman paper below the membrane. The samples were fixed by UV crosslinking and hybridised.

II. Array Preparation

3.5. Nylon filter array preparation

Amplicons were arrayed on 7.8 x 11.9 cm positively charged nylon filters (Hybond-N+, Amersham Pharmacia Biotech, Freiburg) using a 96-pin tool (0.4 mm pins) with a BioGrid spotting device (BioRobotics, UK). Assuming that a pin of 0.4 mm diameter transfers 20 nl of liquid per spotting event three different dilutions of each amplicon (6.25, 25 and 100 ng/µl, respectively) were prepared. Each sample was spotted tenfold to the same position in order to obtain final DNA quantities of 1.25, 5 or 20 ng per sample on different spots. Each sample was arrayed in duplicate using a spotting density of 3 x 3. As a negative control pBluescript vector DNA was spotted onto the 8x12 cm filters. Prior to hybridisation the arrayed DNA was UV crosslinked to the filter.

3.6. End-labeling of plastid transcripts

5´ ends of DNase I-treated RNA (6 µg) were labeled using 5 units T4 polynucleotide kinase in the presence of 30 µCi of [γ-^{32}P]ATP in a final volume of 20 µl according to the supplier's protocol (New England Biolabs, USA). The reaction was stopped by heat inactivation at 65°C for 20 min and remaining nucleotides were removed by the use of columns as described for run-on transcription assays (see above).

3.7. Run-on transcription assays

Run-on transcription assays with 2 x 10^7 lysed plastids were carried out in a 100 µl volume in the presence of heparin as described in Klein and Mullet (1990). Unincorporated nucleotides were removed from the reaction assay by MicroSpin S-200 HR Columns (Amersham Pharmacia Biotech, Freiburg, Germany) following the manufacturer's instructions. Incorporation of α-^{32}P-UTP into elongating transcripts was determined as described by Hallick *et al.* (1976) with aliquots spotted onto DE81 filters (Whatman, Maidstone, UK). The reaction was applied directly as a probe to hybridise macroarrays.

3.8. Array hybridisation

Arrays were prehybridised for 2 h at 65°C in ‚Church buffer' (0.25 M Na_2HPO_4, 7% SDS). Hybridisation of ^{32}P-labeled run-on transcripts, ^{32}P-end-labeled RNAs and ^{32}P-labeled cDNAs, respectively, was performed under the same conditions over night.

Arrays were successively washed at 65°C for 10 min in 2 x SSC, 0.1% SDS, 10 min at 65°C in 1 x SSC, 0.1% SDS, 10 min at 65°C in 0.5 x SSC, 0.1% SDS followed by a final wash at 65°C in 0.2 x SSC, 0.1% SDS for 10 min.

3.9. Image analysis

Radioactive signals were detected using a Fujifilm BAS 1500 Phosphorimager (Fuji, Japan). Images were directly imported into the TINA (raytest, Germany) or ArrayVision software (Imaging Research Inc., Canada) and analysed. In the case of arrays intensities of duplicate spots representing the same gene were averaged. Background correction was done against the negative control (pBluescript).

III. Manipulation of proteins

3.10. Extraction of leaf proteins

Plant tissue was ground in liquid nitrogen. Before thawing, 0.5 ml of the homogenisation solution was added under continuous grinding. The resulting slurry was frozen and stored at –70°C.

Homogenisation buffer

40 mM	β-mercaptoethanol
10%	sucrose
100 mM	Tris/HCl, pH 7.2
5 mM	EDTA
5 mM	EGTA
2 mM	PMSF

3.11. One-dimensional SDS-polyacrylamide gel electrophoresis

Separation of proteins was performed by means of the discontinuous SDS-polyacrylamide gel electrophoresis (SDS-PAGE).

Running gel: height 4.5 cm, thickness 1 mm
10% - 12% acrylamide solution
Stacking gel: height 0.8 cm, thickness 1 mm
4% - 5% (v/v) acrylamide solution

15-well combs

After complete polymerization of the gel, the electrophoresis chamber was assembled. Up to 35 µl sample was loaded per lane and the gel was run at a constant voltage of 80 V for 10 min and then at 140 V until the Bromphenol Blue had reached the end of the gel. Gels were then either stained or subjected to Western analyses.

3.12. Silver staining of protein gels [Heukeshoven and Dernick (1988)]

After SDS-PAGE, gels were fixed and incubated with freshly prepared thiosulphite in acetate/ethanol solution for 30 min at RT with constant agitation. Gels were washed intensively for at least three times for 5 min and subsequently stained for 20 min. Afterwards, gels were developed with sodium carbonate and the reaction was stopped with glycine buffer.

Fixation solution		Thiosulfite solution	
40%	ethanol	0.3%	Na_2SO_3 x 5x H_2O
10%	acetic acid	30%	ethanol
		0.83 M	sodium acetate

Silvernitrate solution		Developing Solution	
0.1%	$AgNO_3$	3%	Na_2CO_3
0.05%	formaldehyde	0.025%	formaldehyde

Stop solution

1%	Glycine

3.13. Coomassie Blue R-250 staining of protein gels

After SDS-PAGE, the gels were stained in Coomassie staining solution (at room temperature for 1h) with constant agitation. The gels were then incubated in destaining solution until the background of the gel appeared nearly transparent.

Coomassie Blue R-250 staining solution

45% (v/v)	methanol
9% (v/v)	acetic acid
0.2% (w/v)	Coomassie Blue R-250 (Sigma-Aldrich GmbH)

Destaining solution

20% (v/v)	methanol
7% (v/v)	acetic acid

3.14. Electrophoretic transfer of proteins onto nitrocellulose membranes (Western analysis)

Proteins were transferred from SDS-gels onto nitrocellulose membranes (Protran Nitrocellulose, Schleicher & Schüll) using a Semi-dry blotting apparatus (Owl, USA). After equilibration of SDS gels in blot buffer for 5 min, the blotting sandwich was assembled as described in the manufacturer's protocol. Proteins were transferred electrophoretically in blot buffer at a constant voltage (85 V for 120 min or 35 V overnight at 4°C).

3.15. Immunological probing of proteins on nitrocellulose membranes

After electrophoretic transfer, the membranes were removed from the sandwiches and placed, protein-binding site up, in glass vessels. Membranes were washed once in TBS (Tris blocking solution without Tween 20) and incubated in blocking buffer for 1 h at room temperature. Afterwards, the primary antibody was added in an appropriate dilution either for 2 h at room temperature or overnight at 4°C. The primary antibody was removed and membranes were washed five times for 5 min with TBST (Tris blocking solution with Tween 20 in concentrations of 0.1% – 0.5%). The secondary antibody/peroxidase conjugate was applied for 2 h at RT. Membranes were washed again five times for 5 min with TBST and immunoreactive bands were visualized using the enhanced chemiluminescence detection system (ECL, Amersham, Braunschweig).

3.16. Immunological detection using enhanced chemiluminescence (ECL)

The immunocomplex composed of nitrocellulose membrane-bound protein, primary antibody and secondary antibody coupled with horse-raddish peroxidase (HRP) was detected using the enhanced chemiluminescence detection system (Orosz et al., 1996). The membrane was soaked for 1 min in detection solution (1:1 mixture of solutions I and II). The solution was removed and the blot was placed between two Saran wrap foils. The membrane was exposed to X-ray film (Biomax-MR, Kodak) for varying periods.

Solution I	Solution II
2.5 mM Luminol	5.4 mM hydrogen-peroxyde
0.4 mM coomaric acid	0.1 M Tris/HCl, pH 8.5
0.1 M Tris/HCl, pH 8.5	

IV. Bacterial transformation

3.17. Preparation of competent bacterial cells

E. coli DH5α cells were streaked on LB-plates containing the appropriate antibiotics and incubated overnight at 37°C. Single colonies were picked and used for inoculation of 10 ml of an overnight culture. 1 ml of the overnight culture was added to 100 ml of pre-warmed LB broth containing antibiotics and shaken until an absorbance of 0.5 at 600 nm was reached (approximately 90 – 120 min). The culture was chilled on ice, transferred to sterile round-bottom tubes and centrifuged at low speed (4000 x g, 5 min, 4°C). The supernatants were discarded and the cells resuspended in ice-cold KP1 buffer (30 ml for a 100 ml culture). The suspension was kept on ice for an additional 90 min. Then, the cells were collected by centrifugation (4000 x g, 5 min, 4°C), the supernatant was discarded again and the cells resuspended in 4 ml ice-cold KP2 buffer. Aliquots of 100 µl were prepared, frozen in dry ice-ethanol mix, and stored at –80°C.

	KP1 BUFFER	**KP2 BUFFER**
potassium acetate	30 mM	
$MnCl_2$	50 mM	
KCl	100 mM	10 mM
$CaCl_2$	10 mM	75 mM
MOPS, pH=7.0		10 mM
glycerol	15%	15%

3.18. Transformation of bacteria

To 100 µl of competent DH5α cells, either 50 - 100 ng of plasmid DNA or 20 µl of ligation mixture were added and incubated for 30 min on ice. After a heat shock (2 min, 42°C) and successive incubation on ice (3 min), 800 µl LB-medium were added

and the bacteria incubated at 37°C for 30 min. The cells were then centrifuged (10000 x g, 1 min, RT) and the supernatant removed. Cells were resuspended in 100 µl of LB medium and plated onto LB plates containing the appropriate antibiotics. Plates were incubated overnight at 37°C.

V. Manipulations of plant cells and organelles

3.19. Seed sterilisation

Seeds were washed in a 6% sodium hypochlorite solution (Ruel-Chemie, Karlsfeld, Germany) for 1 min. The hypochlorite solution was replaced by 70% ethanol (1 min). The ethanol was immediately removed and replaced by sterile water. Several washing steps with sterile water followed.

3.20. Isolation of plastids (according to Müller and Eichacker, 1999)

Chloroplasts were isolated from 3 week-old plant tissue culture growing on modified VBW medium (Aviv and Galun, 1985). Approximately 50 g of leaf material was used for homogenisation in 250 ml of medium A (ice-cold) in a Waring blender. The homogenate was filtered through Miracloth filter (Calbiochem, Darmstadt, Germany) and centrifuged for 2 min at 2545 g.

The chloroplast pellet was gently resuspended in medium A and centrifuged again for 30 sec at 5725x g and 4 ^{o}C. The pellet was then resuspended in approx. 500 µl of medium B and loaded to Percoll gradients (80%/40% for wild-type plastids; 80%/30% for PEP-deficient plastids). The gradients were centrifuged for 20 min, at 7800 g and 4^{o}C. Intact chloroplasts accumulate at the phase between the two Percoll layers of different density. Chloroplasts isolated from gradients were diluted 1:3 in medium B and centrifuged for 5 min, 5725x g and 4^{o}C. Plastids are resuspended in 1 ml of medium B and counted in a Thoma chamber. For run-on analysis 2 x 10^6 plastids/µl were used.

<u>Medium A</u>

4 mM	ascorbate
1.2 mM	$MnCl_2$
0.8 mM	$MgCl_2$

4 mM	EDTA
1 mM	KH_2PO_4
4 mM	DTT
0.2% (w/v)	BSA
0.1% (w/v)	PVP-10
25 mM	MES
25 mM	HEPES, pH 6.8 (KOH)

Medium B

0.33 M	sorbitol
50 mM	HEPES, pH 7.6

80% Percoll solution, pH 7.6 (KOH)

80% (w/v)	Percoll
0.33 M	sorbitol
50 mM	HEPES

30% Percoll solution, pH 7.6 (KOH)

30% (w/v)	Percoll
0.33 M	sorbitol
50 mM	HEPES

3.21. Fractionation of chloroplasts into stroma and thylakoid membranes

Freshly prepared chloroplasts were osmotically ruptured with TMK buffer (100 µg chlorophyll /0.5 ml TMK buffer). Chloroplasts were lysed for 10 min in darkness on ice and centrifuged for 3 - 5 min at 2000 x g. The supernatant (stroma fraction) was collected, the pellet (thylakoids) was washed twice by centrifugation in TMK buffer. The final pellet was resuspended in TMK buffer. Aliquots of the thylakoids were stored at –70°C.

TMK Buffer

10 mM	Tris/HCl, pH 6.8
10 mM	$MgCl_2$
20 mM	KCl

3.22. Isolation of the major thylakoid protein complexes

Thylakoid protein complexes, equivalent to 3 x 10^8 chloroplasts were solubilized for 10 min in ice by 1.5% (w/v) β-dodecylmaltoside (final concentration) as described in Müller et al. (1999). The thylakoid lysate was loaded onto linear 0.1 - 1.0 M sucrose gradients and centrifuged for 16.5 h at 36 000 x g and 4°C (Beckman rotor SP 40 Ti). The sucrose gradients were top-to-bottom fractionated into 35 fractions of 300 µl each using an ISCO 640 gradient fractionator (Instrumentation Specialties Company, USA). For further analyses individual gradient fractions were precipitated with 10% (w/v) trichloroacetic acid and washed with 100% acetone.

3.23. Plastid transformation

The biolistic transformation technique (Boynton et al., 1988; Svab et al., 1990; Svab et al., 1993) was applied. 60 mg of gold particles (0.6 µm) were suspended in 220 µl sterile water and mixed with 25 µl of purified plasmid DNA (1 µg/µl), 250 µl of 2.5 M $CaCl_2$, 50 µl spermidin (0.1 M), followed by 2x washing of the DNA-coated gold particles in 100% ethanol (p.a.), and final suspension of the gold particles in 72 µl of 100% ethanol. The gold particles were delivered into plant cells using Particle Gun - PDS 1000/He, Bio-Rad (Hercules, USA).

<u>Parameters used with the Particle Gun Device</u>

helium pressure	1100 psi
rupture discs	900 psi
distance rupture disc/macrocarrier	8 - 10 mm
distance macrocarrier/stopping screen	10 mm
distance stopping plate/table	7 cm
vacuum (26 – 27 inches Hg)	0.85 bar

For each plasmid construct 20 leaves were shot and dissected 48 h after particle bombardment into 3 x 3 mm segments that were kept in sterile Petri dishes under non-selective conditions for 3 days, and then under selective conditions.

3.24. Selection and regeneration of the mutants

Selection with 500 mg/liter of spectinomycin started three days after transformation. Green colonies began to appear and grow vigorously after approximately 3 weeks, while untransformed cell lines bleached and showed impaired growth. Resistant calluses were transferred onto RMOP medium and cultivated in Petri dishes until shoot formation. For further segregation transplastomic lines were transferred into Petri dishes containing fresh medium at 3 - 4 week intervals. After several rounds of segregation, small green shoots were transferred to 750 ml glass jars containing B5 agar with antibiotic, and grown for 4 - 6 weeks until use.

3.25. Isolation of protoplasts and transformation for transient expression (according to Bayley et al., 1992)

This method was used to characterise the import of the nuclear encoded T3/T7-phage type RNA-polymerases (method modified according to Bayley et al., 1992).

Axenically grown wild-type tobacco leaves (3 weeks old) or leaves of PEP-deficient material (4 - 5 weeks old after regeneration) were harvested (5 - 8 leaves) and incubated with 10 - 15 ml K3AS-medium containing 0.25% Cellulase R10 (Serva) and 0.125% Macerozyme R10 (Serva). The leaves were cut into small stripes (1 mm) and incubated overnight at RT gently shaking in a shaded box. To facilitate release of the protoplasts, the suspension was gently pipetted up and down using a wide-bored pipet. The suspension was filtered through a sterile 100 mesh filter and washed with 5 ml of K3AS medium.

Subsequently, protoplasts were centrifuged at 200x g for 20 min without brake. Intact protoplasts floated and were removed by gently pipetting into a new vessel, then diluted by slowly adding 4 volumes of W5 solution. The protoplasts were again centriguged for 5 min at 200x g, this time with brake. The pellet was resuspended in 40 ml of W5 medium and centrifuged again for 5 min at 200 g with brake. Finally, protoplasts were resuspended in 10 ml of 1x MaMg medium. Cells were counted in a Thoma-chamber. 1 - 1.5 x 10^6 protoplasts were used per transformation. The protoplasts were diluted in 10 ml 1x MaMg solution, centrifuged for 5 min at 200x g with brake and resuspended in 0.5 ml of 1x MaMg medium. 60 µg of the transforming DNA diluted in pure water was added to the sample and mixed gently. After at least 2

min of incubation at room temperature 0.5 ml of 40% PEG solution were added and gently mixed. The suspension was incubated for 25 min at room temperature and then diluted with 5 ml of W5 medium, and then centrifuged for 5 min at 200x g. The pellet was resuspended in K3AM solution to approximately 10^5 protoplasts/ml and incubated in the dark. The cells were harvested after 16 - 24 h by centrifugation for 5 min at 200x g, resuspended in a few µl of K3AM medium and analysed microscopically.

K3AS medium (for 100 ml), pH 5.8 (KOH), after (

Murashige and Skoog-salts + vitamins	10 ml of 10x stock solution
$CaCl_2x2H_2O$	0.046 g
sucrose	13.7 g

W5 medium (for 500 ml), pH 5.8 (KOH)

NaCl	4.5 g
$CaCl_2 \times 2H_2O$	9.18 g
KCl	0.185 g
glucose	0.45 g

2x MaMg medium (for 50 ml), pH 5.8 (KOH)

mannitol	7.3 g
$MgCl_2 \times 6H_2O$	0.305 g
MES	0.2 g

K3AM medium (for 100 ml), pH 5.8 (KOH)

Murashige and Skoog salts +vitamins	10 ml of the 10x stock solution
$CaCl_2x2H_2O$	0.046 g
sucrose	3 g
mannitol	7.28 g

40% PEG solution (for 10ml), pH 5.8 (KOH)

4 g of PEG6000 was disolved in 5 ml of 2x MaMg medium and subsequently adjusted up to 10 ml with water. All solutions were filter sterilized.

4. Results

4.1. Identification of tobacco NEP genes

The transcription machinery of plastids and mitochondria is amazingly complex. Based on biochemical data, the existence of a nuclear coded, plastid targeted RNA polymerase has been suspected quite early (Bogorad et al., 1991; Igloi and Kössel., 1993; Hess et al., 1994). Later, in *Arabidopsis thaliana,* three nuclear genes for two closely related isoforms of a T3/T7 bacteriophage-like DNA-dependent RNA polymerase designated RpoT;1-3 have been described, the gene products of which are posttranslationally imported into mitochondria or chloroplasts, or both organelles (Hedtke et al., 1997, Hedtke et al., 2000). Homologues of these genes have been found in a broad variety of other taxa (Cermakian et al., 1996). A recent report by Bligny et al. (2000) suggested the existence of a further, non-phage-type nuclear-encoded enzyme in spinach plastids.

In the course of this part of my work, which was carried out in co-operation with Boris Hedtke (Berlin), the RpoT gene family of tobacco has been analysed for three principal reasons. First, the plastid genome of tobacco is amenable to genetic manipulation (Bock., 2001). In this way transplastomic lines were generated that lack PEP subunits (Allison et al., 1996; De Santis-Maciossek et al., 1999; Serino and Maliga., 1998). These mutants allow to study the integration of the eubacterial PEP enzyme with the nuclear NEP enzymes. Second, tobacco NEP promoters are well characterized (reviewed in Liere and Maliga., 1999; Weihe and Börner., 1999) as is the plastid transcript pattern of that plant (Krause et al., 2000). Third, *Nicotiana tabacum* is a natural amphidiploid species (2n = 4x = 48), proposed to have originated from a hybridization event between *N. sylvestris* and *N. tomentosiformis* (Kawashima et al., 1976). Studies on synthetic and natural allopolyploids indicated that the majority of parental orthologous genes remained functional (reviewed by Comai., 2000; Wendel., 2000). The fate of nuclear genes encoding organellar proteins in allotetraploids is of particular interest since plastids and mitochondria originate generally from only one of the parental species that served as maternal parent in the original cross. In tobacco, mitochondria and plastids are derived from *N. sylvestris*, whereas the nucleus comprises the genomes of both parental species, *N. sylvestris* and *N. tomentosiformis* (Bland et al., 1985; Olmstead and Palmer., 1991).

Tobacco *rpoT* genes were isolated based on the *Arabidopsis thaliana* homologues. An alignment of the three NEP enzymes from *Arabidopsis* shows an increase in homology from the N- to the C-terminus (Hedke et al., 1997). The conserved regions were used for deduction of degenerate primers. With these and gene-specific primers a large proportions of tobacco *rpoT* genes were amplified from *Nicotiana tabacum* complete genomic DNA and cloned. In addition, clones from a genomic lambda bank were isolated using PCR-derived probes. Altogether, fragments of six genes were isolated, consistent with the expectation that the diploid *Arabidopsis*-like set of three *rpoT* genes is duplicated in tobacco as a consequence of allotetraploidy. The gene pairs exhibited 96.7 - 98.3% identity within coding regions, whereas the conservation amounts to 88.9 - 92.3% in introns (Fig. 3). The six *rpoT* gene fragments were deposited in the EMBL database under the accession numbers AJ302018, AJ302019, AJ302020, AJ416569, AJ416569, and AJ416570.

Using the same set of primers, fragments of *rpoT* genes of the parental diploid species *N. tomentosiformis* and *N. sylvestris* were isolated as well. A comparison of these fragments with the genomic sequences obtained for tobacco allowed to clearly assign each of the six *N. tabacum rpoT* genes to either the *N. sylvestris* or the *N. tomentosiformis* parent (Fig. 3).

4.1.1. Characterisation and analysis of *Nicotiana rpoT* cDNAs

By applying RT-PCR, the expression of all 6 genes was verified and the existence and processing of the complete set of exons and introns known from *Arabidopsis thaliana* was confirmed. These data showed that all six genes are active in tobacco. Next, 5'- and 3'-RACE was carried out in order to obtain full-length cDNAs for the six tobacco *rpoT* genes. However, sequence identity and lack of complete genomic *rpoT* sequences made it difficult to assign unambiguously *N. tabacum* RACE products to either the *N. sylvestris*- or the *N. tomentosiformis*-derived genes. This situation was further complicated when subcloning revealed that even individual tobacco amplicons could contain a mixture of *syl*- as well as *tom*-specific cDNA segments, obviously because of template switching during PCR.

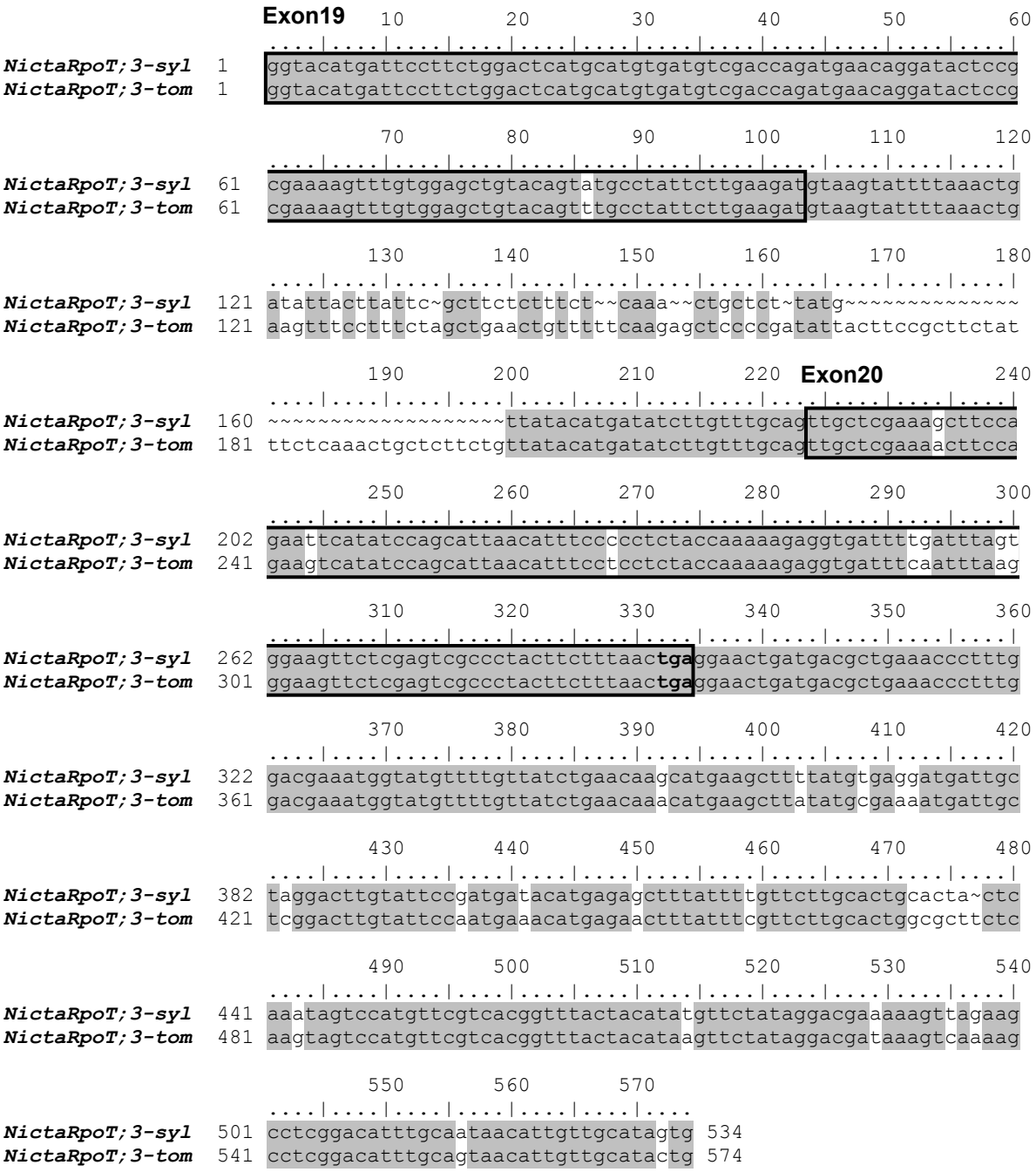


Figure 3. Alignment of *rpoT;3* sequence intervals isolated from *Nicotiana tabacum* genome and their assignment to *N. sylvestris* (*NictaRpoT;3-syl*) or *N. tomentosiformis* (*NictaRpoT;3-tom*) by their similarity. Higher sequence similarity is observed within coding regions (boxed regions), whereas decrease of conserved regions is more evident within introns (region between boxed exon 19 and exon 20). Sequence alignment after stop codon (bold print) of *rpoT;3* gene shows the highly conserved 3'-UTR. Identical bases are underlayed with gray boxes, whereas the wave slashes indicate sequence gaps. Residue numbers related to the outlined segments are given at the left.

To separate and better distinguish the differences between the orthologous, *rpoT* transcripts of the maternal line of *N. tabacum, N. sylvestris*, have been isolated by RT-PCR. In this case, for each PCR only a single product with an unambiguous sequence was obtained and full-length cDNA clones were prepared for all three *rpoT* transcripts. These clones exhibited 99.6 to 100% identity to the tobacco sequences

and hence could be used to assign tobacco RACE-products to either the *N. tomentosiformis* or *N. sylvestris* ancestral gene. Sequencing of small sections of *rpoT* cDNAs from *N. tomentosiformis* further substantiated this assignment (Boris Hedtke, Berlin). Sequence aligment of *N. tomentosiformis* gene products derived from tobacco mRNA (designated as *RpoT;1-tom*, *RpoT;2-tom* and *RpoT;3-tom*) together with those of *Nicotiana sylvestris* (designated *RpoT;1-syl, RpoT;2-syl* and *RpoT;3-syl*) shows increasing sequence similarity from the N- to the C-terminus (Fig. 4).

```
           ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
RpoT;1-tom 1    -------MWRYISKHAYSRK-----FRNSHDSALLGFSQYSSSFGKTRPLQCLCEESTTHPNLGLSQNSIFSRISRKVRH 68
RpoT;1-syl 1    -------MWRYISKHAYSRK-----FRNSHDSALLGFSQYSSSFGKTRPLQCLCEESTTNPNLGLSQNSIFSRISRKVRH 68
RpoT;2-tom 1    MSSTKTPISLTIKLNQFTDKPTGLDINRYHNSPIMWRNIIKQLSSRTPQKLLFSSKNRTYSFLGFGQDSVFKDNT-KFRS 79
RpoT;2-syl 1    MSSTKTPISLTIKLNQFTDKPTGLDINPYHNSPIMWRNIIKQLSSRTPQKLLFSSKNRTYSFLGFGQDSIFKDNT-KFRS 79
RpoT;3-tom 1    ----------------------------MASTASYSPSPTSQWRTQKLPKRFNFYVIHNQEFGKLSQSSSLSTSSFPKTL 52
RpoT;3-syl 1    ----------------------------MASTASYSPSPTSQWRTQKLPKRFNFYVIHNQEFGKLSQSSSLPTSSFPKTL 52

           ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
RpoT;1-tom 69   LEGICEESSKNPHLGLSQNSTFSSVKGDFRICGKRGSGSLGRLRSYGSAAEAIVSTS---EED---IDEIQELIEEMDKE 142
RpoT;1-syl 69   LEGICEESSKNPHLGLSQNSLFSSVKGDFRVCGKRGSGSLGFLRSYGSAAEAIASTS---EED---IDEIQELIEEMNKE 142
RpoT;2-tom 80   LIPISCSNIVMGFQNLGEYLPGDEFLSRPLLKNQVNSNDFCCRKSYASVAEAVAVSSTDAEEDVSVVDEVQELLTELKKE 159
RpoT;2-syl 80   LIPISCSNIVMGFQNLGEYLPGDEFLSRPLIKNQVN-NNFCCRKSYASVAEAVAVSSTDAEEDVSVVDEVHELLTELKKE 158
RpoT;3-tom 53   KLPVIHMPINN-IQSQTTVCVSTDENLEELVNLQKIPNGFLNKESN----KRVFIQD---------PPWVSSLFMNSLFV 118
RpoT;3-syl 53   KLPVIQMPINNNIQSQTTVCVSTDENLEELVNLQKIANGVLTKESN----KRVFIQD---------PPWVSSLFMNSLFV 119

           ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
RpoT;1-tom 143  NEALKANLQPKQPKTIGGMGVGKYNFLRRRQIKVETEAWEEAAKEYQELLMDMCEQKLAPNLPYMKSLFLGWFEPLRDAI 222
RpoT;1-syl 143  NEALKTNLQPKQPKTIGGMGVGKYNLLRRRQIKVETEAWEEAAKEYQELLMDMCEQKLAPNLPYMKSLFLGWFEPLRDAI 222
RpoT;2-tom 160  EKKQFAFRRRKQRMLTSGMGHRKYQTLKRRQVKVETEAWEQAAKEYKELLFDMCEQKLAPNLPYVKSLFLGWFEPLRDKI 239
RpoT;2-syl 159  EKKQFAFRRRKQRMLTSGMGHRKYQTLKRRQVKVETEAWEQAAKEYKELLFDMCEQKLAPNLPYVKSLFLGWFEPLRDKI 238
RpoT;3-tom 119  RAKQVQGVRREFREIER---RRRYAMLRRRQIKAETEAWEQMVEEYRELEREMCEKKLAPNLPYVKKLLLGWFEPLRQAI 195
RpoT;3-syl 120  RAKQVQGVRREFREIER---RRRYAMLRRRQIKAETEAWEQMVEEYRELEREMCEKKLAPNLPYVKKLLLGWFEPLRQAI 196

           ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
RpoT;1-tom 223  AAEQKLCDEGKNRGAYAPFFDQLPAEMMAVITMHKLMGLLMTGG-GTGSARVVQAASYIGEAIEHEARIHRFLEKTKKSN 301
RpoT;1-syl 223  AAEQKLCDEGKNRGAYAPFFDQLPAEMMAVITMHKLMGLLMTGG-GTGSARVVQAASHIGEAIEHEARIHRFLEKTKKSN 301
RpoT;2-tom 240  AEEQELCSQGKSKAAYAKYLYQLPADMMAVITMHKLMGLLMTGG-DHGTARVVQAALVIGDAIEQEVRIHNFLEKTKKQK 318
RpoT;2-syl 239  AEEQELCSQGKSKAAYAKYFYQLPADMMAVITMHKLMGLLMTGG-DHGTARVVQAALVIGDAIEQEVRIHNFLEKTKKQK 317
RpoT;3-tom 196  EKEQNAETTVKHRAAFAPHIDSLPADKMAVIVMHKLMGLLMMGGKEERCVQVVQAAVQIGMAVENEVRIHNFLEKTKKLQ 275
RpoT;3-syl 197  EKEQNAETTVKHRAAFAPHIDSLPADKMAVIVMHKLMGLLMMGGKEERCVQVVQAAVQIGMAVENEVRIHNFLEKTKKLQ 276

           ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
RpoT;1-tom 302  ALSGDLEETPGDMMKERERLRKKVKILMKKQKLRQVRKIVKQQDDEKPWGQDNLVKVGCRLIQILMETAYIQPPNDQLDD 381
RpoT;1-syl 302  ALSGDLEDTPGDIMKERERLRKKVKILMKKQKLQQVRKIVKQQDDEKPWGQDNLVKVGCRLIQILMETAYIQPPNDQLDD 381
RpoT;2-tom 319  AEKDKQKEDGEHVTQEQEKLRKKVTNLMKKQKLRAVGQIVRRQDDSKPWGQDAKAKVGSRLIELLLQTAYIQPPANQLAV 398
RpoT;2-syl 318  AEKDKQKEDGEHVTQEQEKLRKKVTNLMKKQKLRAVGQIVRRQDDSKPWGQDARAKVGSRLIDLLLQTAYIQPPANQLAV 397
RpoT;3-tom 276  KHMTG-AQSQEDMSKETMILRKRVKSLIKRNRVVEVRKLMK-SEEPESWGRDTQAKLGCRLLELLTETAYVQRPVDQSAD 353
RpoT;3-syl 277  KHMTG-AQSQEDMSRETMILRKRVKSLIKRNRVVEVRKLMK-SEEPESWGRDTQAKLGCRLLELLTETAYVQPPVDQSAD 354

           ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
RpoT;1-tom 382  GPPDIRPAFVHTLKTV---ETMKGSRRYGVIQCDPLVRKGLDKTARHMVIPYMPMLVPPQSWLGYDKGGYLFLPSYIMRT 458
RpoT;1-syl 382  CPPDIRPAFVHTLKTV---ETMKGSRRYGVIQCDPLVRKGLDKTARHMVIPYMPMLVPPQSWLGYDKGAYLFLPSYIMRT 458
RpoT;2-tom 399  DPPDIRPAFLHSVRTVAK-ETKSASRRYGIIQCDELVFKGLERTARHMVIPYMPMLVPPVKWTGYDKGGHLYLPSYVMRT 477
RpoT;2-syl 398  DPPDIRPAFVHSVRTVAK-ETKSASRRYGIIQCDELVFKGLERTARHMVIPYMPMLVPPVKWTGYDKGGHLYLPSYVMRT 476
RpoT;3-tom 354  TPPDIRPAFRHVFKIATRDPGKNIVKKYGVIECDPLVVAGVDRTVKQMMIPYVPMLVPPKKWRGYDKGGYLFLPSYLMRT 433
RpoT;3-syl 355  TPPDIRPAFRHVFRIATRDPGKSIVKKYGVIECDPLVVAGVDRTVKQMMIPYVPMLVPPKKWRGYDKGGYLFLPSYLMRT 434

           ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
RpoT;1-tom 459  HGAKQQREAVKRVPKKQLEPVFQALDTLGNTKWRVNRKVLGIVNRIWASGGRLADLVDREDVPLPEAPDTEDEAEIRKWK 538
RpoT;1-syl 459  HGAKQQREAVKRVPKKQLEPVFQALDTLGNTKWRLNRKVLGIVDRIWASGGRLADLVDREDVPLPEEPDAEDEAQIRKWK 538
RpoT;2-tom 478  HGARQQREAVKRASRNQLQPVFEALDTLGSTKWRINKRVLSVIDRIWAGGGRLADLVDRDDAPLPEEPDTEDEALRTKWR 557
RpoT;2-syl 477  HGARQQREAVKRASRNQLQPVFEALDTLGNTKWRINKRVLSVVDRIWAGGGRLADLVDRDDAPLPEEPDTEDEALRTKWR 556
RpoT;3-tom 434  HGSRRQQDAVRGVPTKQMQQVYEALDTLGSTKWRVNKRILNVVESIWAGGGNIAGLVDRKDVPIPELSNSDDIMEVKKWK 513
RpoT;3-syl 435  HGSRRQQDAVRSVPTKQMQQVYEALDTLGSTKWRVNKRILSVVESIWAGGGNIAGLVDRKDVPIPELH-SDDIMEVKKWK 513

           ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
RpoT;1-tom 539  WKVKGVKKENCERHSQRCDIELKLAVARKMKDEDGFYYPHNLDFRGRAYPMHPYLNHLGSDLCRGILEFAEGRPLGTSGL 618
RpoT;1-syl 539  WKVKGVKKENCERHSQRCDIELKLAVARKMKDEDGFYYPHNLDFRGRAYPMHPYLNHLGSDLCRGILEFAEGRPLGKSGL 618
RpoT;2-tom 558  WKVKSVKKENRERHSQRCDIELKLAVARKMKDEEGFFYPHNVDFRGRAYPMHPHLNHLGSDICRGVLVFAEGRPLGESGL 637
RpoT;2-syl 557  WKVKSVKKENRERHSQRCDIELKLAVARKMKDEESFFYPHNVDFRGRAYPMHPHLNHLGSDICRGVLEFAEGRPLGESGL 636
RpoT;3-tom 514  WKVRKSKKINQELHSQRCDTELKLSVARKLKDEEGFYYPHNLDFRGRAYPMHPHLNHLSSDLCRGILEFAEGRPLGKSGL 593
RpoT;3-syl 514  WRVRKSKKINQELHSQRCDTELKLSVARKLKDEEGFYYPHNLDFRGRAYPMHPHLNHLSSDLCRGILEFAEGRPLGKSGL 593
```

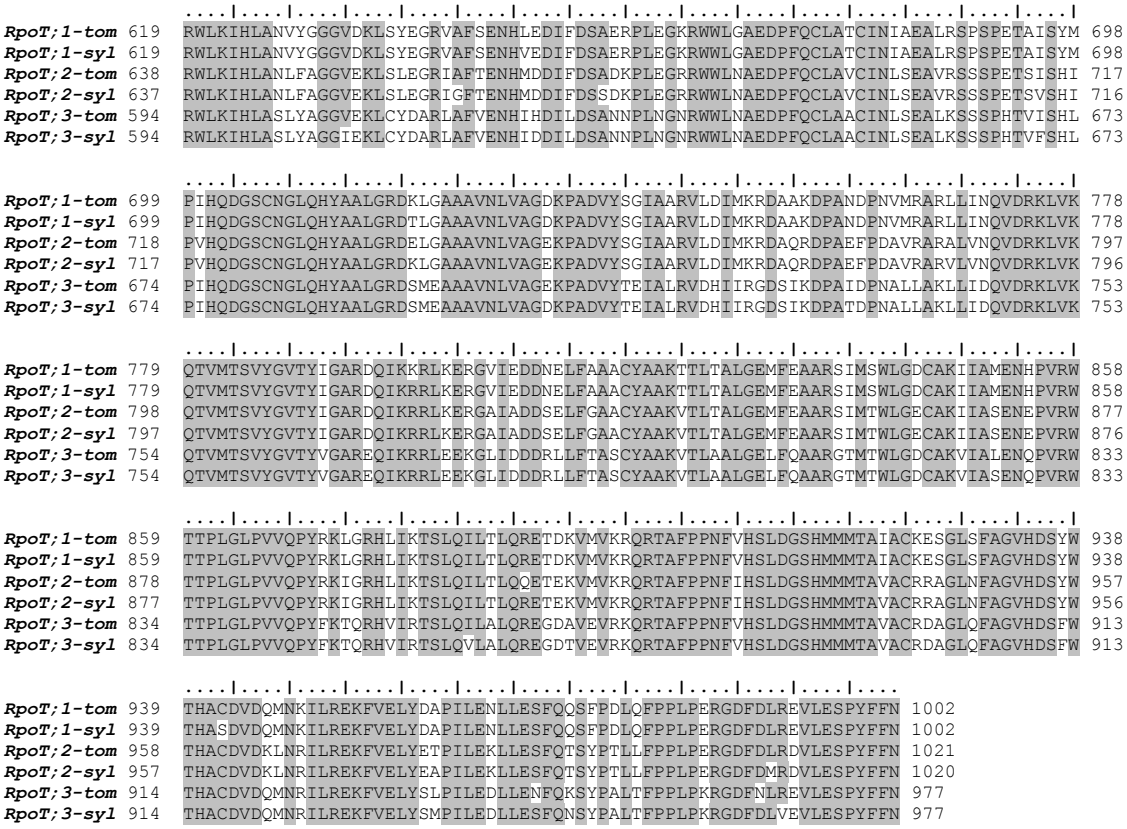


Figure 4. Protein alignment of six RpoT gene products. Amino acid sequence from each of the RpoT gene products, either *N. sylvestris* or *N. tomentosiformis*, is aligned via the ClustalW program (http://www.ebi.ac.uk/servicestmp/-24739.html). Amino acids underlayed in gray are identical. Residue numbers are given at the left and right of the alignment. Sequence gaps are indicated as dashes.

4.1.2. Subcellular localization of tobacco *rpoT* gene products.

Using full-length cDNAs, the subcellular localisation of tobacco RpoT gene products were predicted with the Predotar software (http://www.inra.fr/predotar/). According to this analysis, the RpoT;1 proteins are targeted to mitochondria just like their *Arabidopsis* orthologue (data not shown). Analysis of RpoT;2 at first sight did not yield any clear targeting prediction. However, when a second methionine 35 amino acids downstream was considered as the N-terminus of the mature protein, the protein was predicted to be targeted to the mitochondrium as well. The correspondence between prediction program and *in vivo* studies was shown in the case of RpoT;3 gene. When analysing the N-terminus of *Nicotiana tabacum* RpoT;3, no canonical translation start motif was found. Both *N. tabacum* 5'-RACE products lack the ATG present in *Arabidopsis* (Fig. 5).

Parallel transformation with control constructs expressing proteins targeted to the mitochondrium or the plastid, respectively, confirmed this prediction (Fig. 6 A, B and C, plasmids described in Small et al., 1998).

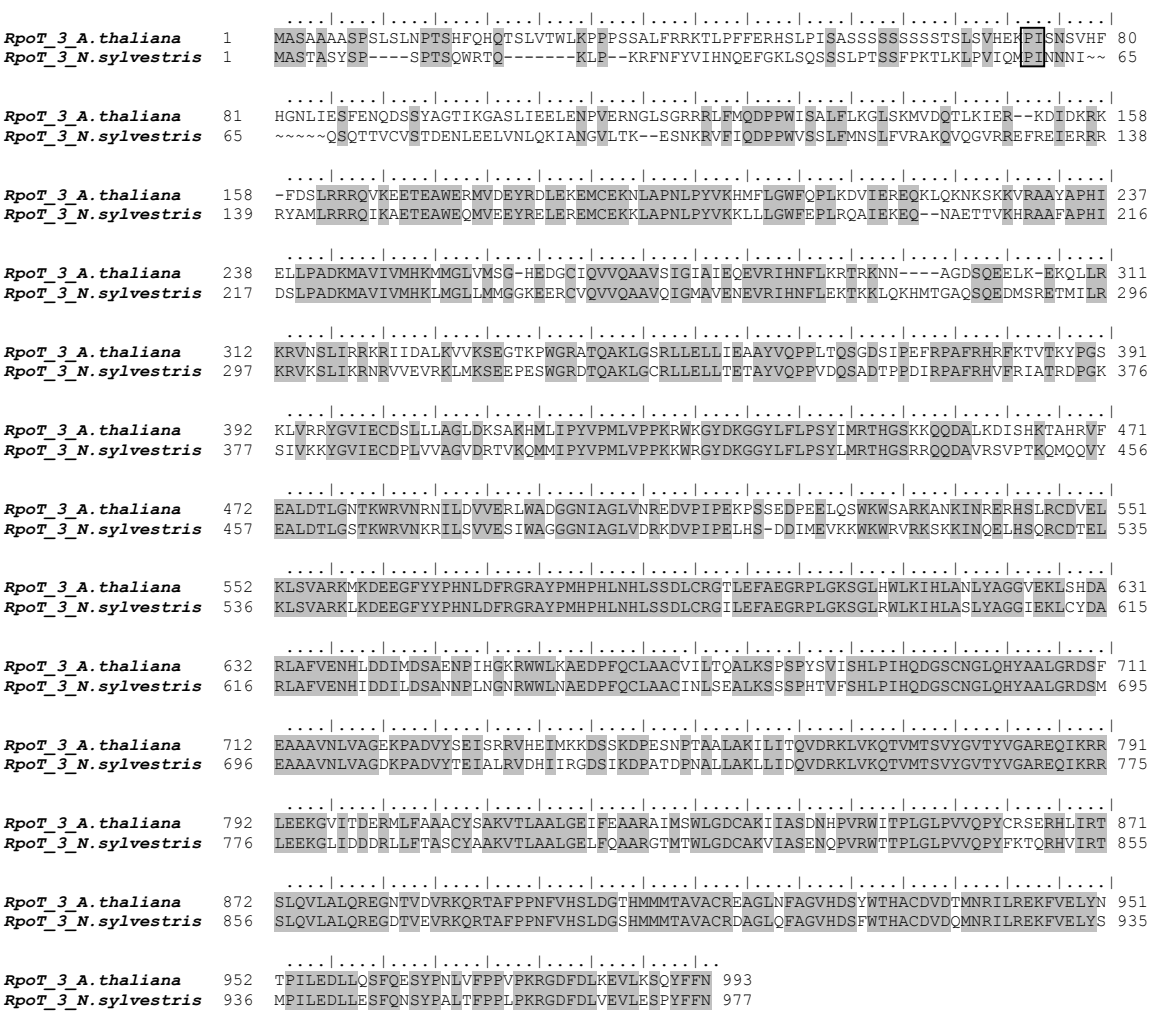


Figure 5. Alignment of RpoT;3 gene products. *Arabidopsis thaliana* and *Nicotiana sylvestris* RpoT;3 gene products are aligned using the ClustalW program (as in Fig. 3). Amino acids underlayed in gray are identical residues. Residue numbers are shown left and right of the alignment. Sequence gaps are indicated as dashes.

Import analyses of RpoT gene products were experimentally verified by applying an *in vivo* transient expression system. The N-terminal 153, and 135 amino acid residues respectively, of tobacco RpoT;1; RpoT;2 were fused to the jellyfish green fluorescence protein (Fig. 6 A; vector pOL-GFP; Akashi et al., 1998). Expression of the fusion construct within the tobacco organelles was supported by a 35S-CaMV-promoter. The expression of 35S-CaMV promoter was enhanced by a quadruple enhancer element in its 5'-UTR (Peeters et al., 1998). After PEG-mediated introduction of these plasmids into tobacco protoplasts, the GFP-signal was detected using fluorescence microscopy. In case of RpoT;1, the GFP signal is localized exclusively in the mitochondria (Fig. 6 D) and for RpoT;2, the signal is found in both organelles (Fig. 6 E).

The situation was clarified comparing the tobacco N-terminus with that of *N. sylvestris*. The *N. sylvestris* RpoT;3 product was lacking an appropriate ATG triplet as well, but stop codons in the 5' untranslated region of RpoT;3 excluded further upstream translation initiation sites. The methionine encoded by the most upstream in-frame ATG (MPINNI-motif boxed with PI in Fig. 5) aligns to amino acid residue 72 of *Arabidopsis* RpoT;3 (Fig. 5) and the derived amino terminus displayed neither plastid nor mitochondrial targeting properties (see Fig. 6 F). Hence, it seems possible that translation initiates not at an ATG codon for this message, which is further corroborated by two findings. First, a high degree of sequence conservation between *N. sylvestris*- and *N. tomentosiformis*-type cDNAs, which is typical for coding regions, starts about 180 nt upstream of the first ATG. Second, the deduced amino acid sequence of this conserved region is enriched in hydroxylated amino acids, but is virtually lacking acidic residues, thus exhibiting features of stroma-targeting chloroplast transit peptides (von Heijne et al., 1989). Analysis of this N-terminal sequence of RpoT;3 with the Predotar-algorithm indeed suggested that this protein is directed to the plastid, in accordance with the subcellular localisation of its *Arabidopsis* orthologue. Thus, a translational start at a non-ATG triplet was considered (Boris Hedtke, Berlin).

To test the hypothesis of a non-AUG initiation of translation, Boris Hedtke designed three different RpoT3-GFP fusion constructs, that contained the wild-type CTG as start codon, a construct with a CAC start codon as a negative control and a substitution of the wild-type CTG by a ATG as a positive control. After transformation, it was shown that both the CTG construct as well as the positive control directed GFP to plastids, whereas the CAC negative control resulted in a substantially weaker fluorescence, which was located exclusively in the cytosol (Hedtke et al., 2002).

In this study, a deletion of the upstream CTG and additional 59 5' terminal codons yielded a construct starting at “ATG59”, in order to test, whether this methionine could act as a start codon as well. The expression of this construct in tobacco protoplasts, did not result in organellar GFP fluorescence (Fig. 6 F), which suggests strongly that this ATG is not used for translation initiation *in vivo*.

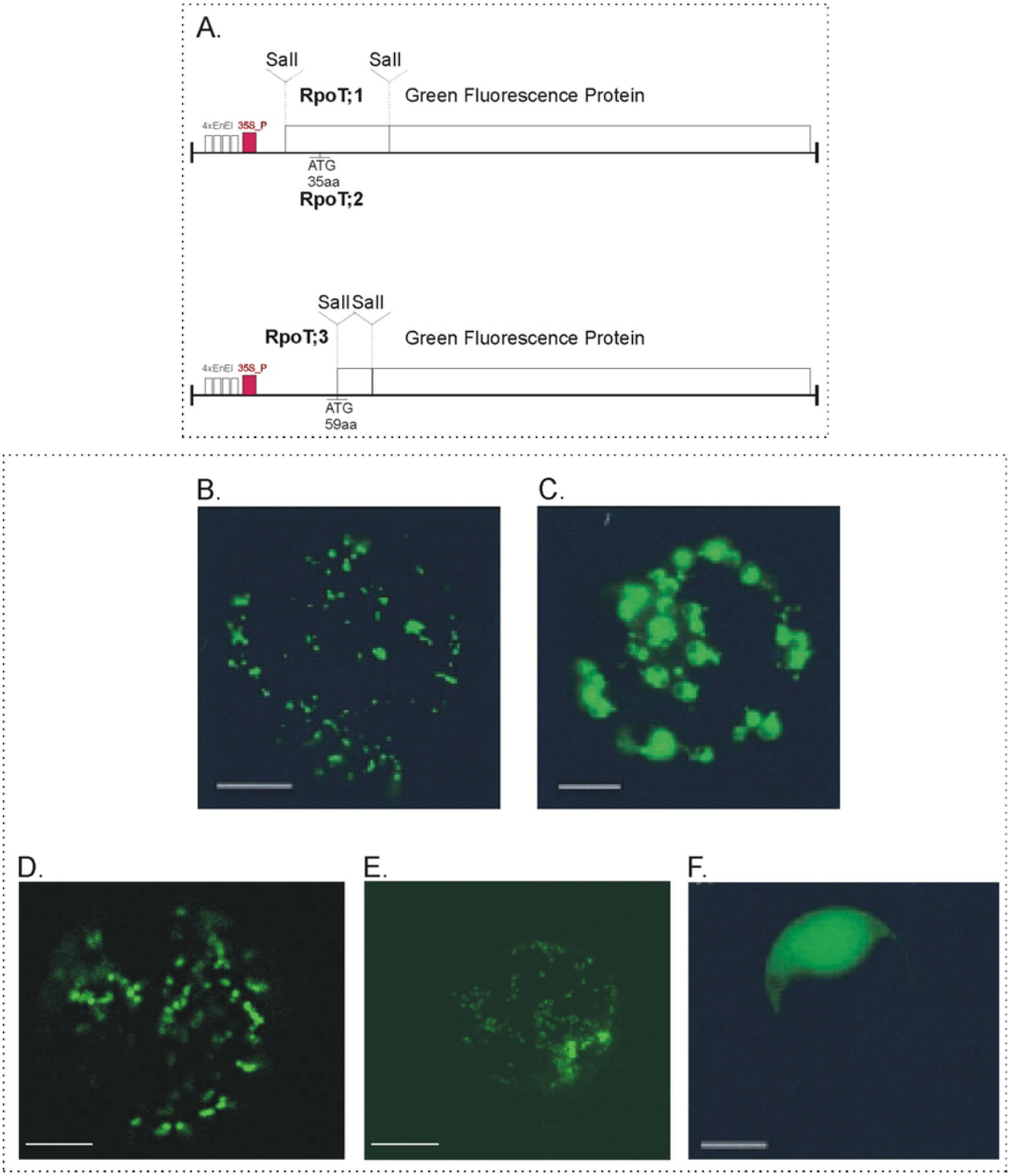


Figure 6. *In vivo* expression analyses of RpoT;1:GFP, RpoT;2:GFP and RpoT;3:GFP fusion constructs in tobacco protoplasts. A. Schematic overview of the constructs RpoT;1, RpoT;2 and RpoT;3 containing 153 and 135 amino acid residues of the N-terminal sequence of RpoT;1 and RpoT;2, respectively, fused in frame to GFP and a construct of 44 amino acid residues of the N-terminal part of RpoT;3 protein. 4xEnEI = 4 repetitive transcriptional enhancer elements; 35S_P = CaMV-35S promoter. (B and C) UV-microscopy of tobacco protoplasts transformed with control constructs CoxIV and RecA, respectively. Protoplasts were isolated from PEP-deficient material to avoid chloroplast autofluorescence. Control transformations using plasmids described previously (Peeters et al., 1998) confirm the identity of the respective organelles: The CoxIV derived construct directed GFP to mitochondria, whereas the RecA drived construct directed GFP to plastids. (D - F) UV-microscopy of tobacco protoplasts transformed with constructs of RpoT;1:GFP (D), RpoT;2:GFP (E) and RpoT;3:GFP (F). The scale bar denotes 10 µm.

4.2. Comparative analysis of plastid transcription profiles attributed to wild-type and PEP-deficient transcript machineries

4.2.1. Plastid macroarrays and their limits

The existence of at least two related transcription systems operating within plastids was documented more then ten years ago (see review Hess and Börner, 1998 and references therein). To untangle the relation of the nuclear and plastid RNA polymerases active in chloroplasts, comparisons of PEP-deficient material with wild-type material were performed in this study. To survey the complete transcriptional activity and transcript accumulation of a plant, a macroarray filter representing the gene content of the tobacco plastid chromosome was prepared. This is the first macroarray for a plastid organelle to date. Similar expression studies have been performed for the mitochondrial genome of *Arabidopsis* coding for only 57 genes (Giege et al., 2000).

DNA fragments representing the entirety of 118 genes and 11 open reading frames were amplified from total DNA. Those genes, containing introns where spotted as cDNA products in addition, representing a spliced form of the genes (*atpH*, *clpP*, *ndhA*, *ndhB*, *trnA(UGC), trnE(UUC), trnG(UCC), trnI(GAU), trnK(UUU), trnL(UAC), trnV(UAC), ycf3, rps12, rps16* and *rpl2*). This turned out to be crucial for intron-containing transfer RNAs, since DNA fragments with intervening sequences did not reliably detect spliced tRNAs under the chosen hybridisation conditions. Quality and amount of PCR products were checked by agarose gel electrophoresis, so that each of the amplicons could be adjusted to a distinct concentration (Fig. 7).

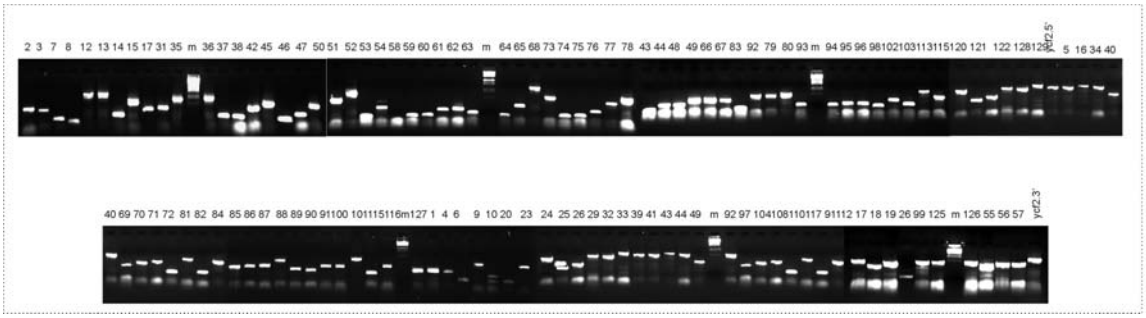


Figure 7. Amplificons of all tobacco plastid genes. The 129 PCR products amplified from the total tobacco genomic DNA were separated in a 2% agarose gel.

4.2.2. Technical advances of array studies

4.2.2.1. Specificity test

Using a Biogrid (BioRobotics, UK) robot device equipped with a 96-pin tool a total of 129 amplicons, including pBlueScript vector as a negative control, were arrayed on 8 x 12 cm positively charged nylon membranes (Hybond N+, Amersham-Pharmacia, Freiburg) according to the pattern presented in Figure 8.

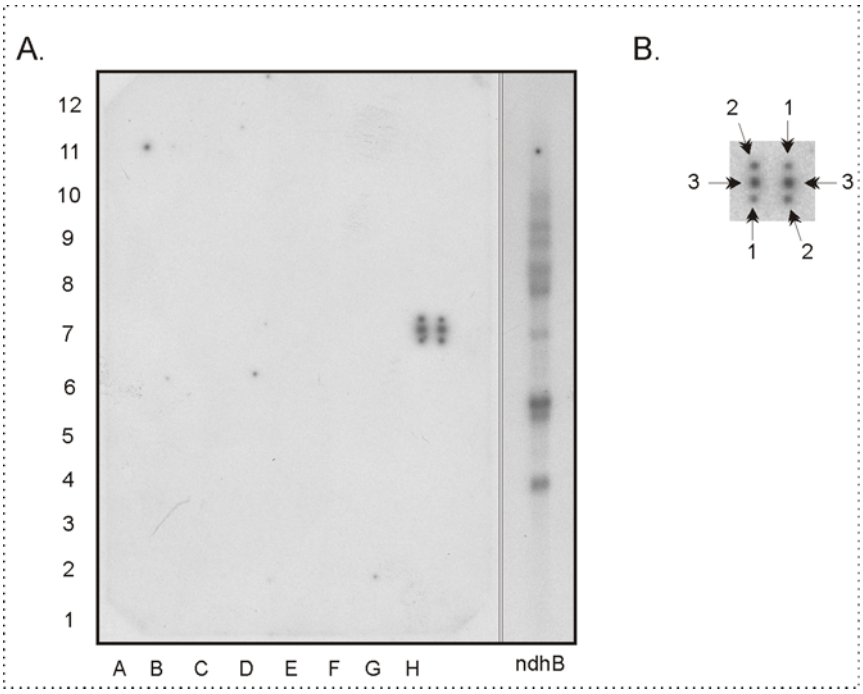


Figure 8. (A) Control of specificity of array hybridisation. An *in vitro* generated transcript specific for the *ndhB* gene was hybridised to an array membrane together with a Northern blot. Only spots related to *ndhB* were detected. The Northern blot shows the typical transcript pattern obtained for an *ndhB* specific probe. (B) Detail picture of dotted PCR product within one spot group. 1, 2 and 3 designate the positions where three different concentrations of the PCR probe are dotted (1.25, 5 and 20 ng, respectively). Duplicates of each concentration dotted on the filter is named as a double offset. Thus, each spot group, represented by 6 spots (as in B) represent one specific amplicon dotted in three different concentrations.

To increase the quantity of transferred DNA and to minimize eventual transfer differences, each sample was spotted by a 10-fold repetition to the same point onto the membrane. Each sample was finally arrayed in three different quantities (1.25, 5 and 20 ng, respectively) and in duplicates, causing double positives on probing (Fig. 8). Furthermore, the specificity of sample recognition was tested by hybridising arrays to radioactively labeled RNA probes synthesised by *in vitro* transcription of an individual PCR product. In each case, only the expected spot group was recognised (an example is shown in Fig. 8).

4.2.2.2. Sensitivity test

To test, whether sensitivity can be increased when applying more DNA to the filter, arrays prepared with 20 instead of 10 repetitive spotting events were hybridised in parallel to 10x-filters with equal amounts of end-labeled RNA. The resulting autoradiographs show that the sensitivity at only 10x application of the probes is not significantly lower than with twice the amount of PCR-products (Fig. 9).

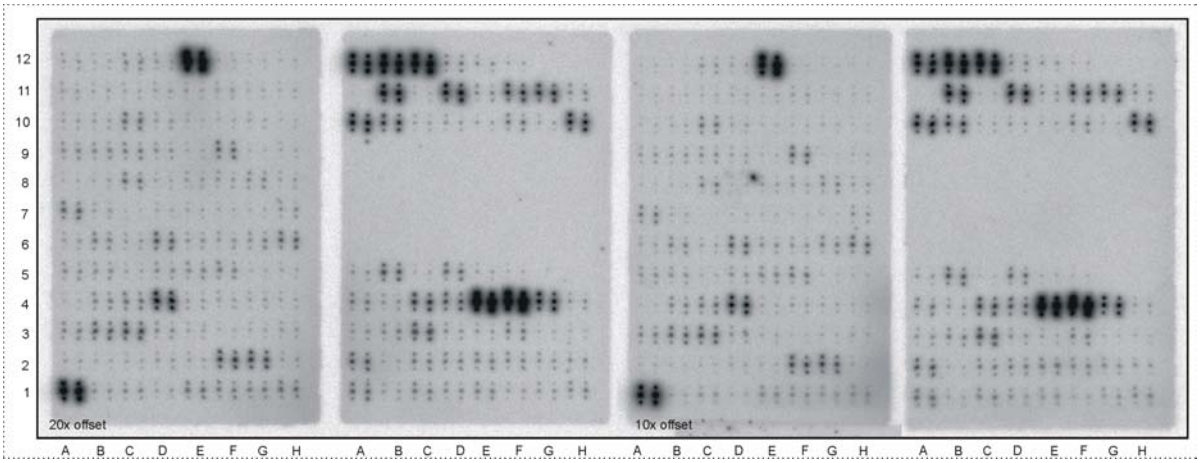


Figure 9. Comparison of 20x and 10x offset print. Filters with 20x and 10x repetitive spotting events per sample were prepared and hybridised with 5 µg end-labeled RNA from tobacco wild-type.

The uniformity and reproducibility of sample loading was verified by hybridising membranes to 5'-[γ^{32}P]-ATP-labeled oligonucleotides complementary to the T7-RNA-polymerase promoter sequence that is added to each PCR product as the result of 5' tagged primers utilised in the amplification reaction. Autoradiographs showed that arrays were uniform with respect to the amount of PCR products spotted (data not shown).

4.2.2.3. Single-stranded vs. double-stranded probes

Genes on plastid chromosomes are located on both DNA strands. Therefore, run-through transcription of genes encoded by the opposite DNA strand can lead to antisense RNAs that may contribute to transcript levels if probes include transcripts of both DNA-strands (e.g. run-on transcripts or end-labeled plastid RNA) or if probes are hybridised to double-stranded DNA fragments (e.g. gene-specific amplicons on arrays). The array approach presented here, as almost all previous work on transcript levels in wild-type and PEP-deficient tobacco plants (Allison et al., 1996; Hajdukiewicz et al., 1997; Serino and Maliga, 1998; Krause et al., 2000) and comparable work with other materials (Inada et al., 1996; Silhavy and Maliga,

1998a,b) is based on double-stranded probes, and hence, inherently does not allow to assess strand-specific transcription. To estimate the influence of antisense RNA when hybridising labeled transcripts to arrayed double-stranded DNA fragments, the reverse approach was performed, in which plastid RNA was dotted onto Nylon membranes and hybridised against surplus amounts of radioactively labeled single-standed antisense RNA (Fig. 10 A). In general, the results were consistent with those obtained with DNA-based filters. However, differences with respect to wild-type/mutant signal ratios were noted in a number of instances, of which some are illustrated by Figure 10. For *atpA*, *atpI*, *psbC*, *psaA*, *rps14*, *petB* and *petD*, which all showed decreased transcript levels in PEP-deficient plastids as determined by strand-specific probes, wild-type to mutant signal ratios turned out to be lower with the array approach (Fig. 10 B). The opposite was observed for *ycf3*, *rps18*, and *5' rps12*, which exhibited increased transcript levels in PEP-deficient plastids and a higher WT/rpo$^-$ ratio with the strand-specific approach (Fig. 10 B). These probe-dependent differences in wild-type/mutant transcript ratios are apparently caused by antisense RNA.

In a previous study, it was shown that the ratios between sense and antisense RNA can vary widely between individual genes but that the regional amounts of antisense transcripts in both wild-type and mutant leaf tissue were generaly comparable (Krause et al., 2000). The array data are consistent with these findings. The observed shifts in wild-type to mutant transcript ratios are most probably caused by comparable regional fractions of antisense transcripts in both materials which are detected with arrayed double-stranded DNA but not with strand-specific probes. If, in the case of *psaA* with a wild-type to mutant RNA ratio of approximately 2.58 to 1.0 an antisense fraction of 0.72 is added the ratio decreases to 1.92 (3.3 to 1.72) as measured by the array approach. On the other hand, in the case of *rps18*, addition of an antisense fraction of 0.6 leads to an increase of the ratio from 0.31 to 0.57 (cf. Fig. 10 A). Thus, quantitative comparison of transcript levels in plastids strongly depends on the probe-type applied, that is, single-stranded or double-stranded.

To investigate antisence transcripts in detail, plasmid (pBlueScrips II SK$^-$) containing a *petD* insertion (from nucleotides 79042 to 80267, according to tobacco plastid genome, Acc.Z00444) was cut with a single cutting enzyme (*SalI*) allowing to prepare

in vitro a sense-specific probe by using T7-RNA polymerase. Bands of approx. 4, 1 and 0.24 kb, covering partialy the *petD*-coding region show that antisense transcripts in the region of *petD* exist and somehow regulate *petD* gene expression in both, wild-type and PEP-lacking material. This was in accordance with data obtained, as shown in Fig. 10 A and B.

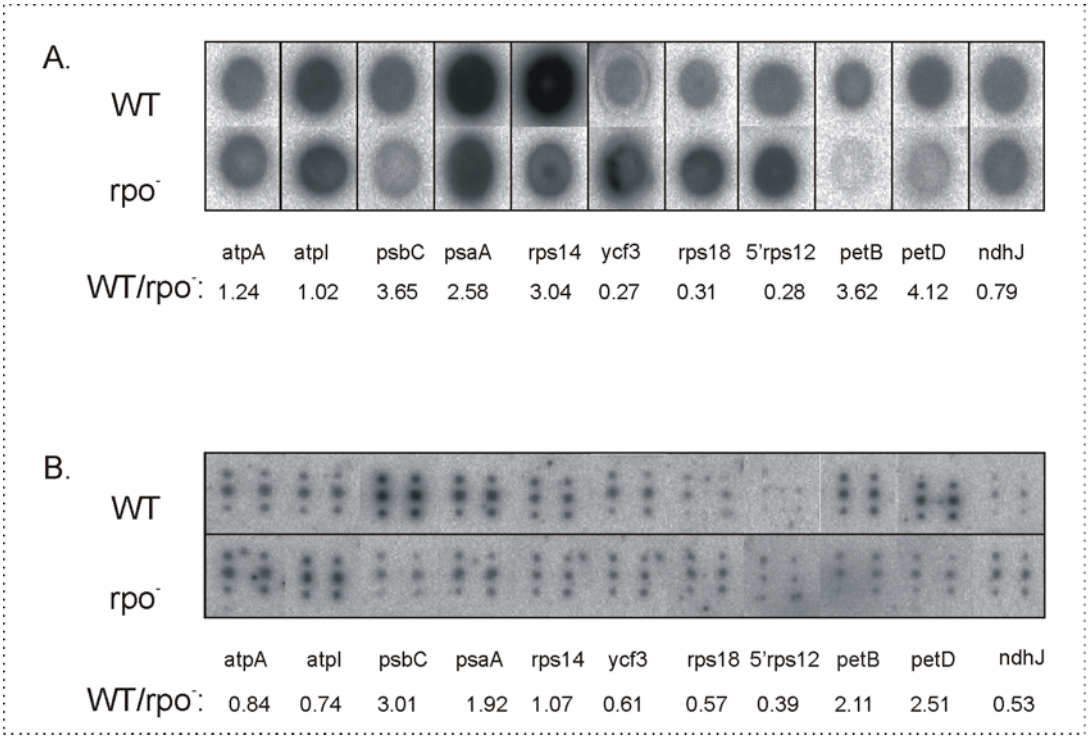


Figure 10. Quantitative comparison of transcript levels in plastids strongly depends on the probe type used for hybridisation. A. Total cellular RNA from wild-type and *rpoC*-deficient mutants was spotted onto a Nylon membrane and hybridised against gene-specific single-stranded RNA-probes. B. Double stranded PCR products arrayed onto the membrane and hybridised against end-labelled RNA from wild-type and *rpoC*-deficient material. Ratios of measured values (Tina20 program) between wild-type and *rpoC* mutant are given below each gene.

4.2.2.4. Probe size, blot and hybridisation reproducibility

In order to assess the influence of the length of arrayed DNA fragments on signal intensity, radioactively end-labelled plastid RNA from mutant material was hybridised to *psbA* PCR products of different length spotted onto a Nylon filter. Signal intensities turned out to be only marginally stronger (<10%) with longer DNA fragments. Such differences are negligible in a comparative analysis of expression profiles as presented here (Fig. 11 A). In order to estimate reproducibility of the expression pattern, hybridisation data from two experiments with wild-type samples of two independent RNA preparations were used to probe the membranes. Scatter plot analysis of these data verified that reproducibility is relatively high. Correlation

coefficients R^2 of 0.96 to 0.97 (see Fig. 11 B) mean that the results deviate by only 3 - 4%.

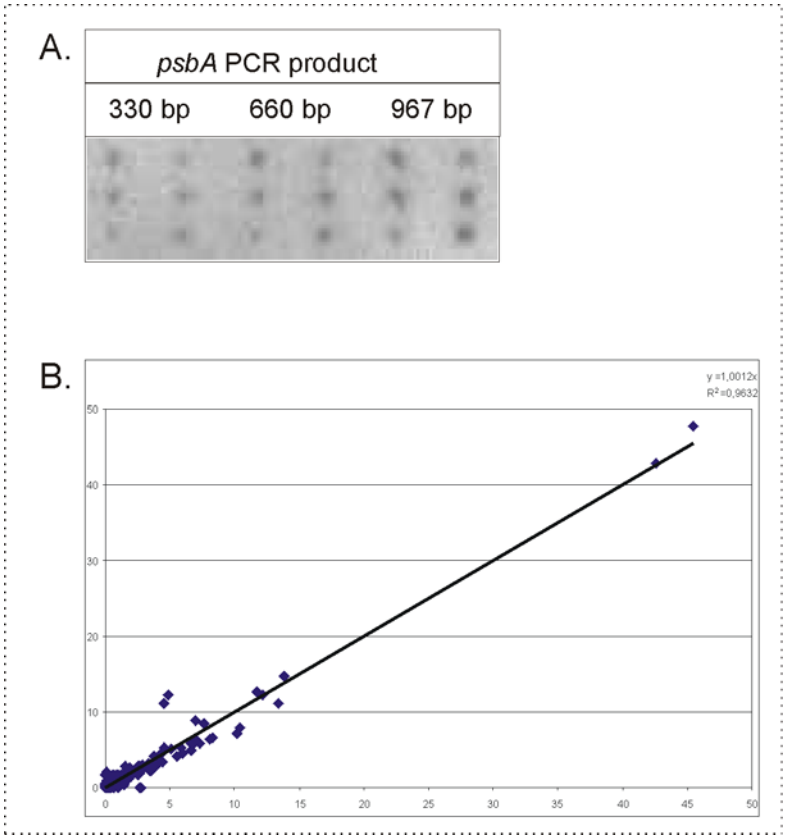


Figure 11. Influence of the lengths of the spotted PCR-product on signal intensity after hybridisation and test of reproducibility of hybridisation results. (A) Three overlapping PCR-products (the three products share one of the primers) of different length were arrayed on a membrane, which was hybridised with end-labelled RNA from *rpoC*-deletion mutants. (B) Scatter plot analysis of two independent hybridisation experiments. Results obtained from two independent hybridisation experiments where estimated for their signal intensitities with the Array visioner program. Signal intensities obtained from each spot group (six spots) where summarised and averaged by the number of spots resulting in an expression intensity of one related gene (Psl/mm^2). These signal intensities obtained from two independent experiments where directly compared resulting in the correlation slope taking coefficient R^2 of 0.96 to 0.97.

4.2.2.5. What is the best probe for hybridising macroarrays?

Steady-state levels of transcripts can be characterised by hybridizing macroarrays with either end-labeled total plastid RNA or with labeled cDNA prepared from plastid RNA.

The cDNA synthesis was performed by reverse transcription either priming with random hexa-(octa)-nucleotides or with a mixture of 129 gene-specific oligodesoxyribonucleotides, each complementary to the coding strand of one of the plastid genes or ORFs (Fig. 12). Although the use of gene-specific primers in reverse

transcription reactions leads to more pronounced signals compared to those derived with randomly primed cDNA, expression of a number of tRNA species could not be demonstrated. This was noted when the primer used in reverse transcription assay spanned the 3' genomic sequence of the tRNA. Posttranscriptional modification of the transfer RNAs by addition of the CCA motif at their 3' end (Gegenheimer et al., 1995) most probably prevents annealing of the gene-specific primer used in the reverse transcription reaction. Another problem is that for polycistronic messages a gene-specific primer leads to the production of a heterogenous cDNA population that covers the proximal cistrons more abundantly than the distal ones, because termination of reverse transcription is random and hence sequences close to the primer are transcribed more often than those more far away. These problems suggested that the use of RNA labeled with T4 polynucleotide kinase (PNK) is more adequate to study steady-state levels of transcript accumulation than labeled cDNA.

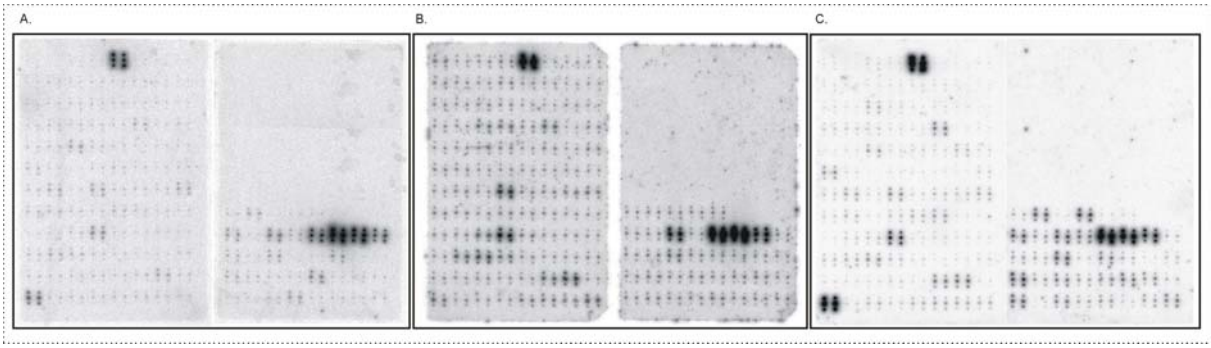


Figure 12. Generation of the probe influences hybridisation results. A complete set of PCR products for plastid-encoded genes was arrayed onto Nylon filters and hybridized with (A) wild-type tobacco cDNA generated by reverse transcription with random octaoligonucleotides as primers, (B) cDNA generated by reverse transcription with 129 gene-specific oligonucleotides as primers and (C) with 5' end-labelled total RNA (5 µg) from wild-type tobacco.

4.2.2.6. T4 Polynucleotide kinase specificity

The enzymatic activity of PNK leads to the transfer of the phosphate group from ATP to the 5' hydroxyl termini of RNA whereas the 5' triphosphate and 3' hydroxyl ends remain unaffected. This was experimentally verified by selective labeling of RNA oligonucleotides but not of *in vitro* transcripts (Fig. 13). Apparently, T4-PNK labels only 5' hydroxyl, but not 5' triphosphate groups. Hence, primary transcript ends with their free triphosphate are not labeled by this method.

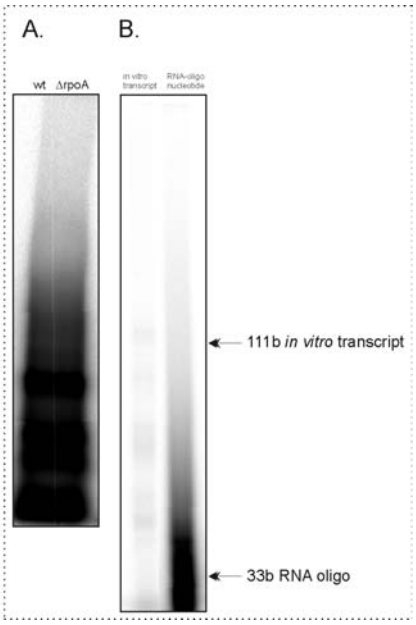


Figure 13. End-labelling of RNA with T4-polynucleotide kinase. (A) To investigate the activity of T4-PNK, 5 µg of end-labelled RNA from wild-type and *rpoA*-deficient material was treated with 10 U of T4-PNK. (B) End-labelling of the same amounts (0.641 nmol/µl) of an *in vitro* transcript (111 b of the gene *3' rps12*) and an RNA-oligonucleotide (33 b) with 10 U of T4-PNK. Whereas only minor amounts of the 0.65 nmol of the *in vitro* transcript are labelled, the oligonucleotide (0.6 nmol) seems to be extensively labelled. This shows that T4-PNK preferentially labels processed RNA molecules containing 5' OH groups.

4.3. Comparison of expression profiles of wild-type and PEP-deficient plastids

Changes in transcription rate of individual plastid genes do not necessarily lead to comparable changes in steady state transcript levels. To assess both, the activity of the different RNA polymerases as well as the fate of the synthesised transcripts, array-based profiles of transcription rates as well as of steady-state transcript levels were determined in this study.

For the first experiment, run-on transcript analyses were performed which take advantage of the observation that chloroplast lysates continue transcription in the presence of radioactive NTPs during a short time pulse so that isolated RNA from enriched plastid fractions giving a mirror of the amount of nascent transcripts, that is, the transcription rate (activity). This RNA was hybridized to macroarrays. For the second line of experiments, steady state levels of transcripts were characterised by hybridizing macroarrays with end-labeled total plastid RNA. In the former case, transcription of operons, genes and ORFs of the tobacco plastid chromosome was investigated by hybridising [α-^{32}P]UTP-labeled run-on transcripts from plastid lysates to arrays. Profiles of actual transcript levels, in turn, were studied by probing the arrays with total plastid RNA which had been 5' end-labeled with PNK in the presence of [γ-^{32}P]ATP.

The following sections provide a detailed comparison of transcription rates and transcript steady-state profiles of the complete gene set in wild-type and PEP-

deficient plastids, taking into account the above outlined technical aspects as well as the proposed classification of transcripts either to be synthesised by PEP, NEP or both enzymes (Hajdukiewicz et al., 1997). The study was complemented by Northern and Western analyses of selected sets of genes in order to trace the fate of transcripts synthesised in wild-type or mutant leaves in more detail.

4.3.1. Genes encoding photosynthesis-related components

Previous reports (Allison et al. 1996; Hajdukiewicz et al., 1997; Serino and Maliga, 1998) have shown that, in the absence of PEP, transcription initiation of tobacco plastid genes and operons is linked to promoters specifically used by NEP. Furthermore, Western analysis verified that plastid-encoded polypeptides involved in photosynthesis do not accumulate in plastids lacking PEP (DeSantis-Maciossek et al., 1999). Figure 14 outlines the arrangement of the genes of the tobacco plastid chromosome on the array. The arrays were then probed with either RNA fraction from run-on assay of lysated plastids or with end-labeled steady-state RNA from different sources. Aside of presenting the original autoradiographs, spots have been analysed with the Tina20 software, in order to quantify the results (Fig. 15).

Run-on transcription data show that all plastid genes are transcribed in a PEP-deficient background, though into different profiles compared to wild-type plastids (Fig. 14 B). This confirms that the absence of plastid-encoded photosynthesis-related gene products in mutants lacking PEP is not primarily based on selective transcription of the corresponding genes by this polymerase (cf. Krause et al., 2000).

4.3.1.1. Transcriptional activity in wild-type and Δ*rpoA* mutant

The genes *rbcL* (encoding the large subunit of Rubisco), *psbA, psbB, psbC, psbD, psbE, psbF*, *psbH*, *psbJ*, *psbL*, *psbM*, *psbT*, *psbN*, *psbZ* (*ycf9*, Swiatek *et al.*, 2001) (encoding subunits of the photosystem II assembly), *psaA* and *psaB* (encoding the two core proteins of photosystem I), *petB*, *petD*, *petL* and *petN* (encoding subunits of the cytochrome b_6f complex), *ndhG* and *ndhI* (encoding subunits of the NADH dehydrogenase) exhibit relative transcription rates that are significantly lower in mutant than in wild-type plastids, as expected. However, various genes encoding subunits of photosystem I (*psaC* and *psaI*), photosystem II (*psbI* and *psbK*), of the cytochrome b_6f complex (*petG*) and the NADH dehydrogenase (*ndhA*, *ndhC*, *ndhD*

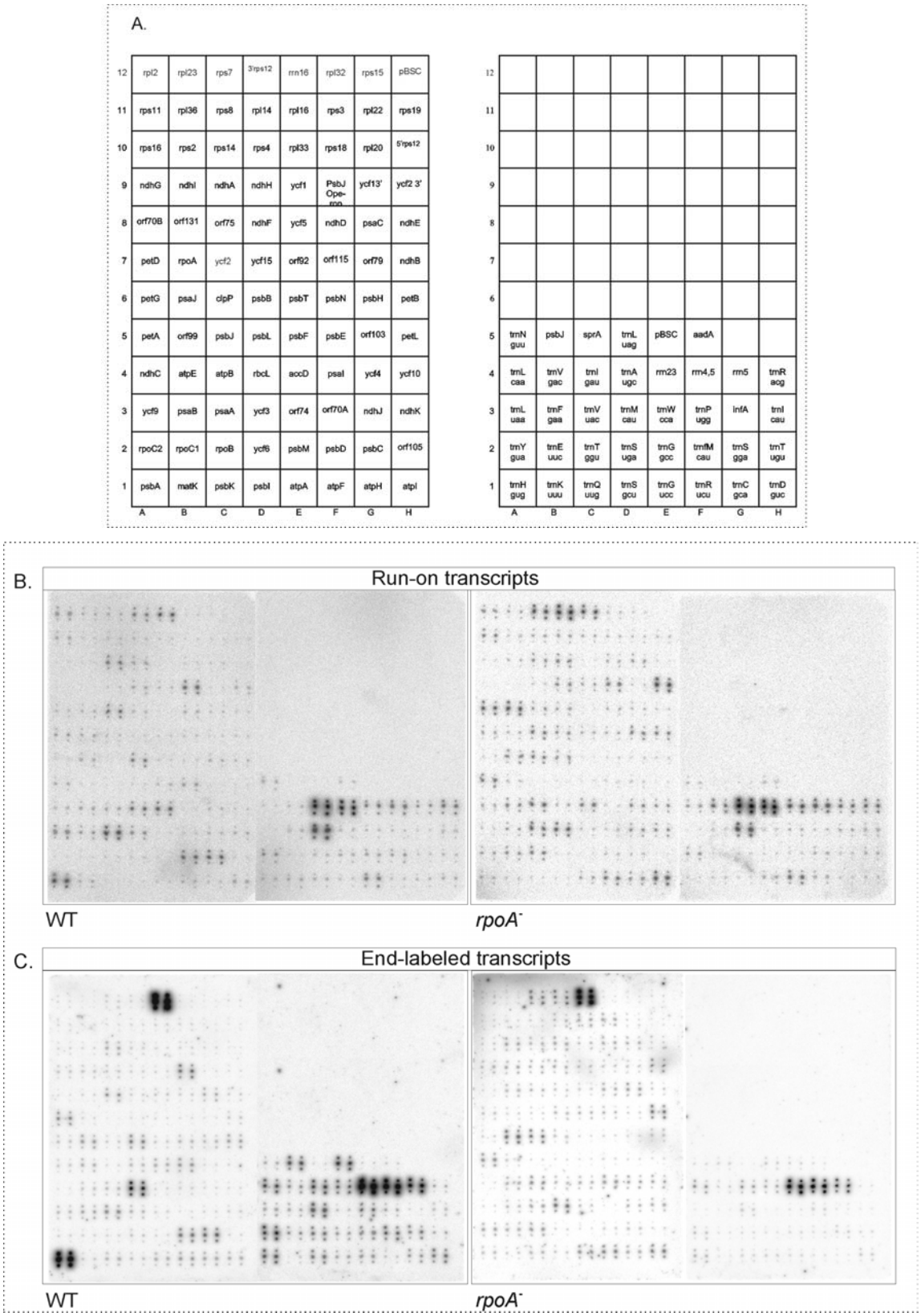
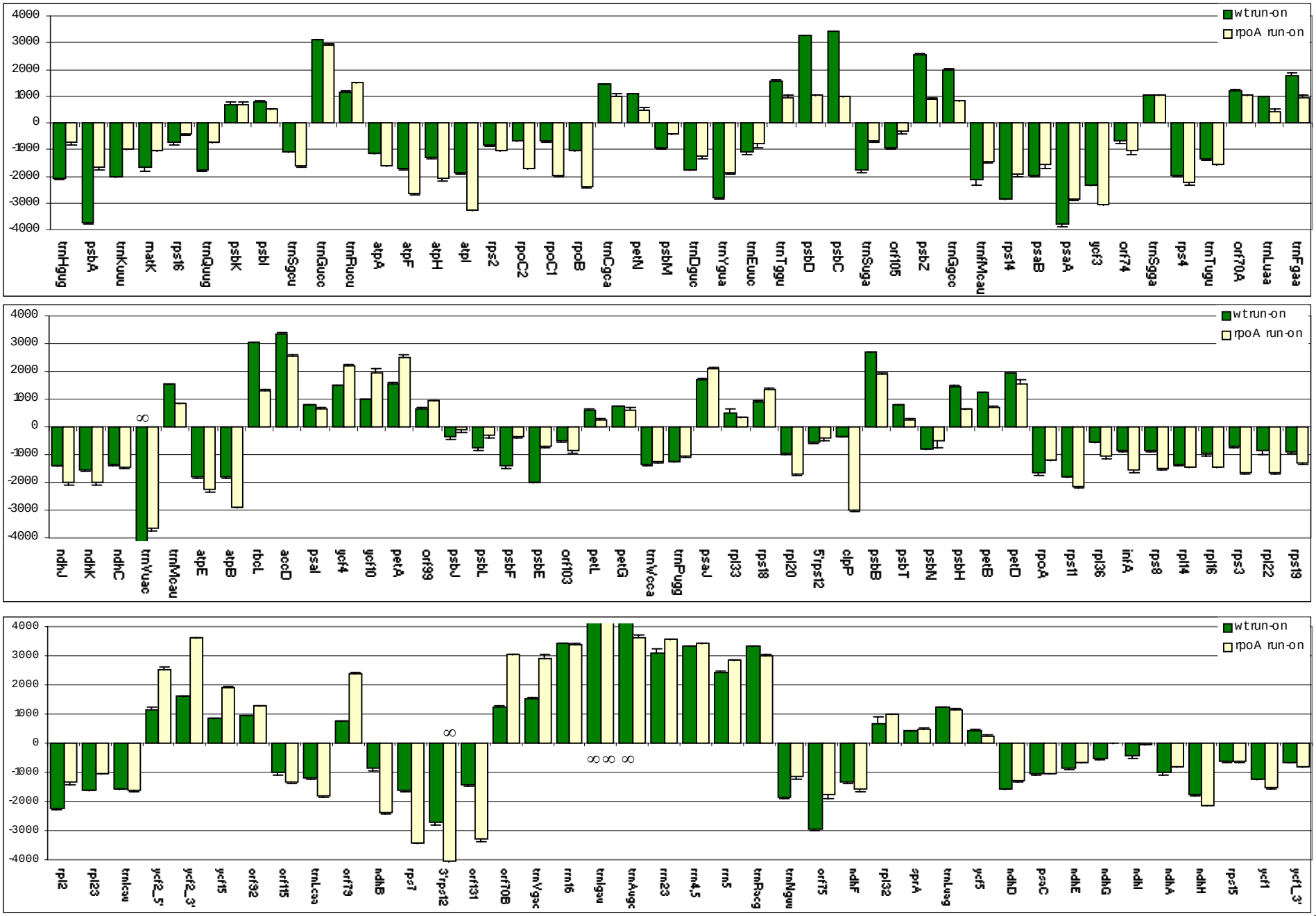


Figure 14. (A) Schematic presentation of positions of gene and ORF probes of the entire tobacco plastid chromosome on the array. (B and C) Autoradiograms of nylon filters representing all known genes and ORFs of the tobacco plastid chromosome probed in (B) with run-on transcripts and in (C) with 5' labeled RNA from wild-type (WT) and PEP-deficient (*rpoA*$^-$) plastids. Each DNA fragment was arrayed in three different quantities in duplicate using a spotting density of 3 x 3 resulting in the vertical pattern (left/right) 2, 3, 1 and 1, 3, 2, respectively, with 1, 2 and 3 representing pairs of spots with DNA quantities of 1.25, 5 and 20 ng, respectively (See Fig. 8). DNA fragments were arrayed according to the pattern depicted in (A).

A


wt run-on
rpoA run-on
4000
3000
2000
1000
0
-1000
-2000
-3000
-4000
trnHgug
psbA
trnKuuu
matK
rps16
trnQuug
psbK
psbI
trnSgcu
trnGucc
trnRucu
atpA
atpF
atpH
atpI
rps2
rpoC2
rpoC1
rpoB
trnCgca
petN
psbM
trnDguc
trnYgua
trnEuuc
trnTggu
psbD
psbC
trnSuga
orf105
psbZ
trnGgcc
trnfMcau
rps14
psaB
psaA
ycf3
orf74
trnSgga
rps4
trnTugu
orf70A
trnLuaa
trnFgaa
ndhJ
ndhK
ndhC
trnVuac
trnMcau
atpE
atpB
rbcL
accD
psaI
ycf4
ycf10
petA
orf99
psbJ
psbL
psbF
psbE
orf103
petL
petG
trnWcca
trnPugg
psaJ
rpl33
rps18
rpl20
5'rps12
clpP
psbB
psbT
psbN
psbH
petB
petD
rpoA
rps11
rpl36
infA
rps8
rpl14
rpl16
rps3
rpl22
rps19
rpl2
rpl23
trnIcau
ycf2_5'
ycf2_3'
ycf15
orf92
orf115
trnLcaa
orf79
ndhB
rps7
3'rps12
orf131
orf70B
trnVgac
rrn16
trnIgau
trnAugc
rrn23
rrn4,5
rrn5
trnRacg
trnNguu
orf75
ndhF
rpl32
sprA
trnLuag
ycf5
ndhD
psaC
ndhE
ndhG
ndhI
ndhA
ndhH
rps15
ycf1
ycf1_3'

B

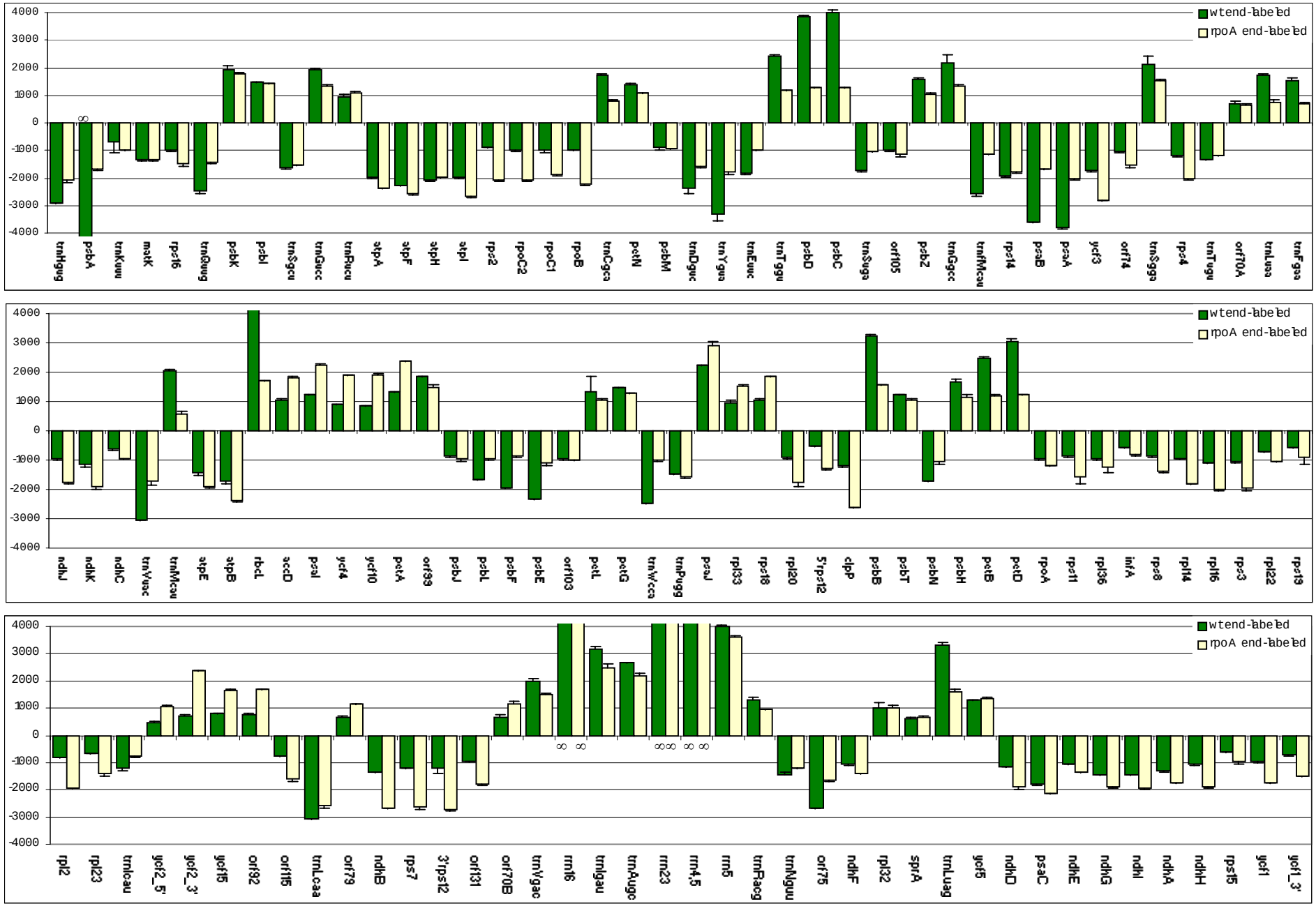

wtend-labeled
rpoA end-labeled
4000
3000
2000
1000
0
-1000
-2000
-3000
-4000
trnHgug psbA trnKuuu matK rps16 trnQuug psbK psbI psbG trnGucc trnRucu atpA atpF atpH atpI rps2 rpoC2 rpoC1 rpoB trnCgca petN psbM trnDguc trnYgua trnEuuc trnTggu psbD psbC trnSuga orf105 psbZ trnGgcc trnfMcau rps14 psaB psaA ycf3 orf74 trnSgga rps4 trnTugu orf70A trnUua trnFgaa
ndhJ ndhK ndhC trnVuac trnMcau atpE atpB rbcL accD psaI ycf4 ycf10 petA orf33 psbJ psbL psbF psbE orf103 petL petG trnWcca trnPugg psaJ rpl33 rps18 rpl20 5'rps12 clpP psbB psbT psbN psbH petB petD rpoA rps11 rpl36 infA rps8 rpl14 rpl16 rps3 rpl22 rps19
rpl2 rpl23 trnIcau ycf2_5' ycf2_3' ycf15 orf92 orf115 trnLcaa orf79 ndhB rps7 3'rps12 orf131 orf70B trnVgac rrn16 trnIgau trnAugc rrn23 rrn4.5 rrn5 trnRacg trnNguu orf75 ndhF rpl32 sprA trnLuag ycf5 ndhD psaC ndhE ndhG ndhI ndhA ndhH rps15 ycf1 ycf1_3'

Figure 15. (A) Histograms indicating the relative transcription rates (first three tables) and steady state RNA quantities (second three tables (B)) from plastids of wild-type (green bars) and PEP deficient (Δ*rpoA*) mutants (white bars). In order to compare data from Δ*rpoA* and wild-type material, equal numbers of plastids were used for run-ons. 16S *rRNA* transcription rates in both types are known to be identical. This is in accordance with our data (value of 3400 pixel/mm^2 in both plastid types) and hence serves as a positive control. For steady-state measurements, *atpI* was used as a positive control. For this gene, it was shown earlier, that transcript amounts in wild-type are 0.8 times those of Δ*rpo* tissue, consistent with our data (Allison et al., 1996). The values overexposed and designated with infinite sign are values which were exclueded from the evaluation true their high expression levels. The values indicated at the x-axis are arbitrary units obtained via evaluation in TINA20 sofware.

and *ndhE*) turned out to be transcribed at almost similar rates in both, PEP-deficient mutants and wild-type. Even more astonishingly, various genes, also coding for subunits of thylakoid assemblies, notably *atpA, atpB, atpE, atpF, atpH, atpI* (specifying constituents of the thylakoid-located ATP synthase), *psaJ, petA*, *ndhB*, *ndhF*, *ndhH*, *ndhJ and ndhK* are transcribed at even higher rates in mutant than in wild-type plastids. The same holds for the reading frames *ycf3*, *ycf4* and *ycf10* which have been shown to encode proteins with photosynthesis-related functions (Figs. 14 B and 15 A; Boudreau et al., 1997; Rolland et al., 1997; Ruf et al., 1997). It should be noted that individual genes of one operon can show quite different signal intensities (e.g. *psbE*, *psbF*, *psbL* and *psbJ* or *rpoB*, *rpoC1* and *rpoC2*) with decreasing values from the first to the last gene of the operon. Most probably, this can be attributed to the lower number of longer transcripts generated in run-on assays, possibly also *in vivo*.

4.3.1.2. Transcript levels in wild-type and ∆*rpoA* mutant

When probing arrays with end-labeled RNAs, in many cases hybridising transcript quantities correlate with findings of run-on analyses (see Figs. 14 B and C, 15 A and B). Operons or genes with high transcription rates generally lead to higher transcript levels and *vice versa*. However, in a number of cases, concerning both photosynthesis and non-photosynthesis-related genes (see below), no correlation between transcription rates and transcript quantities could be found. For instance, *psbM* and *psbT* exhibit very low relative transcription rates in mutants (Figs. 14 B and 15 A), but the corresponding end-labeled RNAs were detectable in almost similar quantities in both wild-type and mutant (Figs. 14 C and 15 B). On the other hand, in mutants relative transcription rates of *psaI* only slightly differ in comparison to wild-type, but the resulting transcript quantities are significantly higher. Surprisingly, *petB* and *petD* transcription rates and transcript levels determined by the arrays differ, with intensities for *petD* being substantially higher. This is unexpected since *petB* and *petD* are cotranscribed as part of the *psbB* operon and not processed into individual RNAs in tobacco (Monde et al., 2000a). However, in addition to *petB*/*petD* containing processed RNA species including the 1.9 kb *psbH*/*petB*/*petD* and the 1.5 kb *petB*/*petD* transcripts a 1.2 kb *petH*/*petB* dicistronic RNA accumulates to substantial amounts. No monocistronic *petD* mRNA is detectable by Northern analysis (Fig. 16; Monde et al., 2000). Thus, the amount of *petB* containing transcripts is higher than

that of *petD*, as revealed by Northern analysis with strand-specific probes (Fig. 16). In the ”double-stranded” array approach, signal intensities for *petD* are influenced by antisense transcripts probably originating from run-through transcription of the *rpoA* or *aadA* gene flanking *petD* on the opposite DNA strand in wild-type and mutants, respectively (data not shown). More striking discrepancies appear when transcription rates are compared with data from Northern analyses. For example, the majority of genes coding for thylakoid proteins, such as *psbC*, *petB* and *petD*, accumulate only very low steady state transcript levels in mutants lacking PEP as judged by Northern analyses (Fig. 16 and data not shown), although they are transcribed with appreciable rates, obviously by NEP (Figs. 14 B and 15 A).

4.3.1.3. Qualitative differences between wild-type and ∆rpoA mutant

In addition to quantitative differences either in transcription rate or in RNA accumulation various genes for photosynthesis-related functions also exhibited conspicuous qualitative differences in transcript pattern between wild-type and PEP-deficient mutants. For some genes, such as *ndhJ*, the signal strength among individual bands of an RNA pattern differs between wild-type and mutant plastids (see Fig. 16). This may be caused by operon-internal promoters and/or differences in processing kinetics or RNA stability. In other cases, like *atpA*, *atpI*, *psaA*, *ycf3*, *rps14*, *rps18* or *ndhJ* (Fig. 16), high molecular weight precursors accumulate specifically in mutants lacking PEP (cf. Krause et al., 2000). For instance, whereas wild-type plastids accumulate *psaA* and *rps14* transcripts of approximately 5.2 kb in size, which correspond to the tricistronic *psaA*/*psaB*/*rps14* message, mutant leaf tissue lacks this band but exhibits large transcripts of approximately 7.5 and 9.5 kb. When probing for *ycf3*, signals of 7.5 and 9.5 kb are detectable again exclusively in mutants (Figs. 16 and 17). This suggests that in the absence of PEP *psaA* and *psaB* messages accumulate as parts of a large polycistronic *ycf3*/*psaA*/*psaB*/*rps14* RNA which is consistent with the localisation of a putative NEP promoter upstream of *ycf3* in plastids of mustard (Summer et al., 2000).

Aside of differences in transcript patterns between mutants and wild-type, there are also interesting similarities questioning the importance of PEP. For instance, the *petA* operon exhibits a transcript pattern virtually indistinguishable between ∆*rpoA* and wild-type plants, no matter, if a *petA*-, *ycf10*- or *ycf4*-probe was used (Fig. 18).

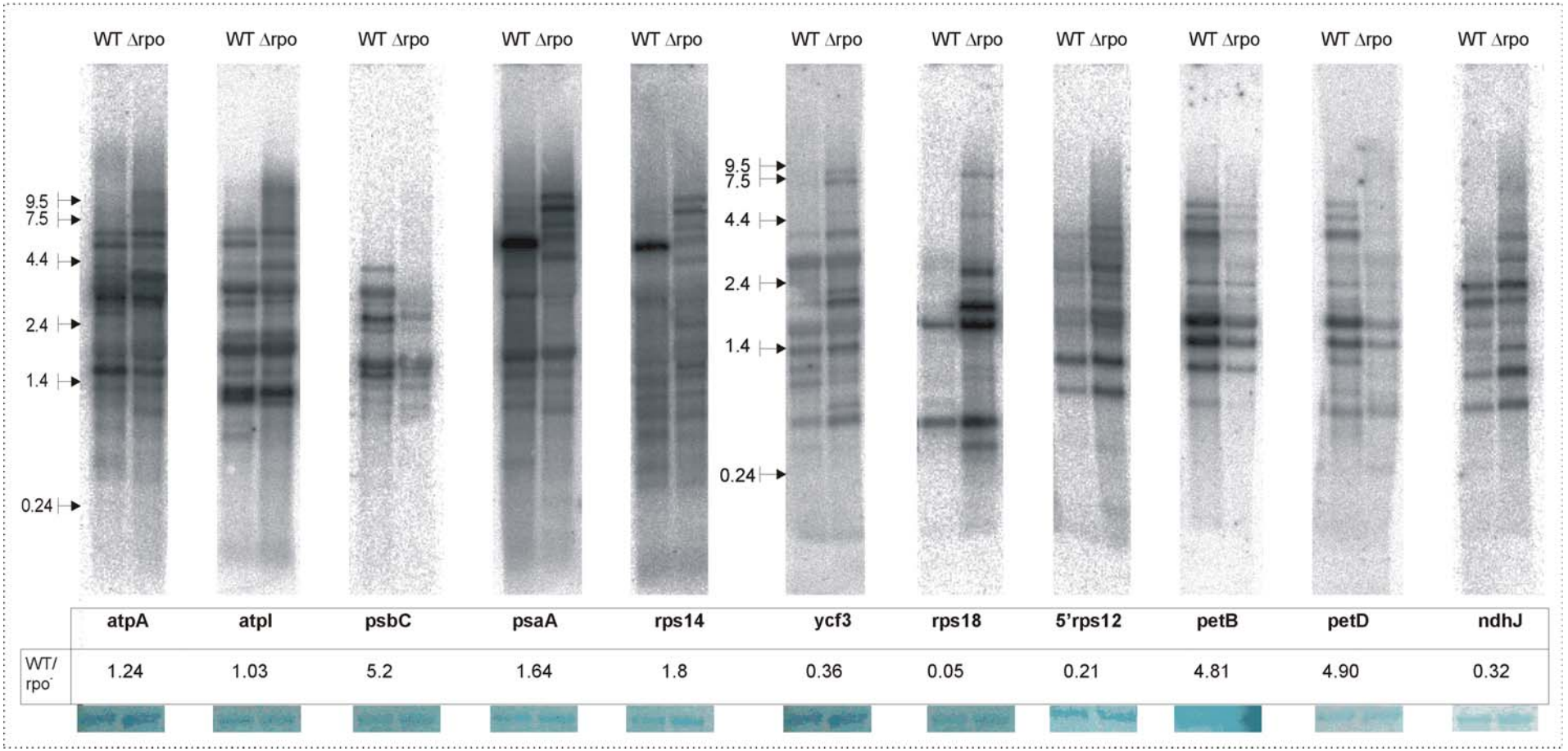


Figure 16. Selected Northern filters of RNA from tobacco wild-type (WT) and PEP-deficient mutants (Δ*rpoA*). Equal amounts of RNA were electrophoretically separated in agarose/formaldehyde gels and probed with *in vitro* transcripts specific to the mRNA of *atpA*, *atpI*, *psbC*, *psaA*, *rps14*, *ycf3*, *rps18*, 5'*rps12*, *petB*, *petD* and *ndhJ*, respectively. A 25S rRNA control for equal loading is shown below autoradiograms. Signals were quantified with a Fuji phosphorimager and the TINA20 (raytest) program.

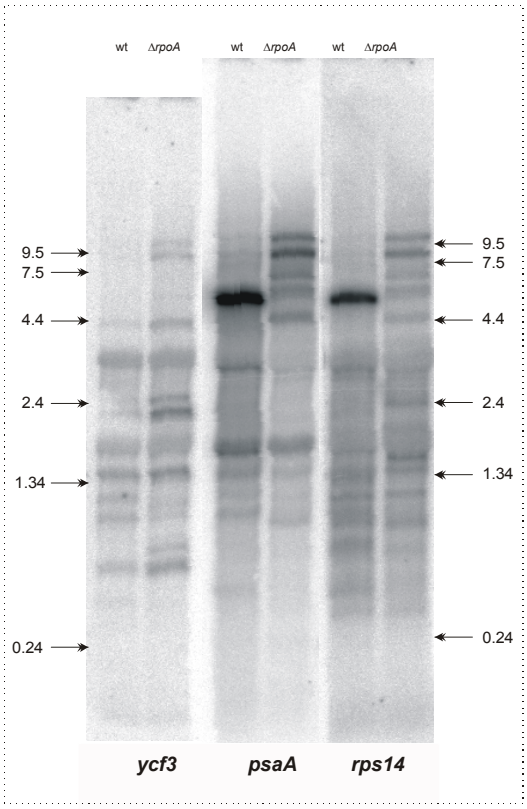


Figure 17. Transcript pattern of the *psaA/psaB/rps14* operon. Northern analysis of transcripts coding for the reaction center polypeptides of PSI (PsaA and PsaB), for Ycf3-protein and for: protein CS14 (5 µg RNA/lane separated on a 1% agarose gel and probed with gene-specific single-stranded RNA-probes). The major transcript covering the tricistronic message *psaA/psaB/rps14* of 5.2 kb is produced by PEP (Meng et al., 1988). Additional transcripts of approx. 9.5 and 7.5 kb accumulate only in *rpoA*-deficient plastids. This long transcript includes the upstream genes *ycf3* and *ORF74*.

Smaller differences concern only high molecular weight precursors at 4.5 and 6.3 kb that are typical for Δ*rpo* material and suggest a retarded processing kinetics. Western analyses revealed that cytochrome *f* is translated and accumulated in Δ*rpo* material although no thylakoid membrane system is developed (Fig. 18 B).

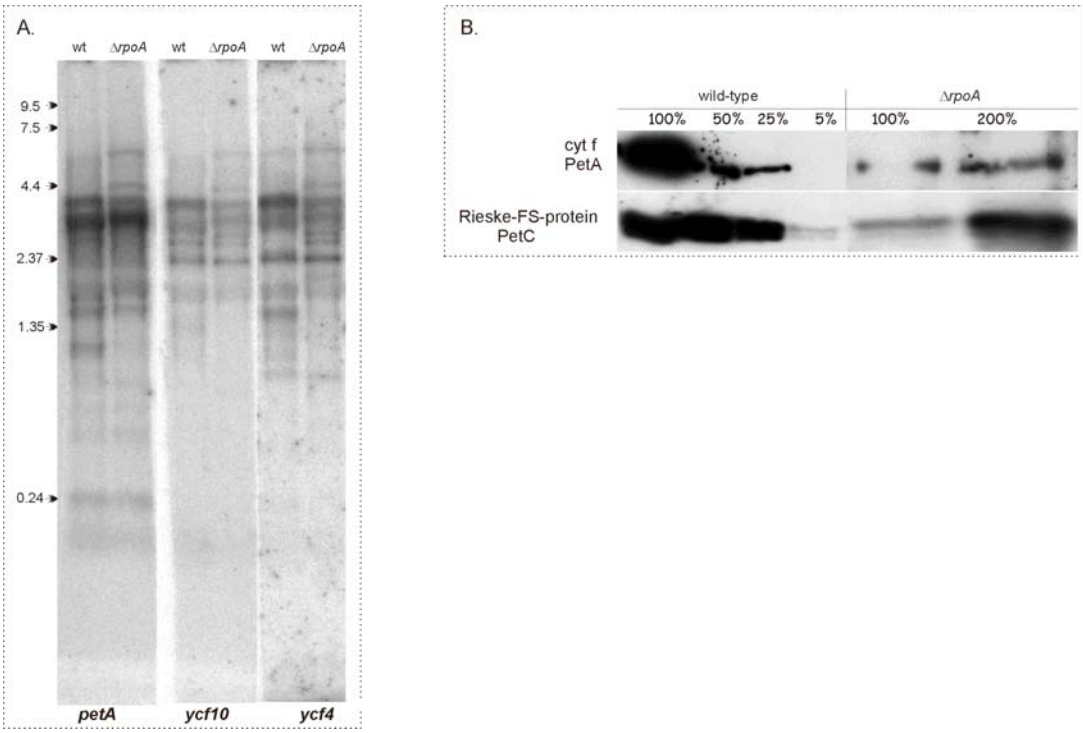


Figure 18. Northern analysis of the *petA* operon. 5 µg of total cellular RNA from wild-type tobacco (WT) and PEP-deficient tissue (Δ*rpoA*) were electrophoretically separated in a 1% agarose/formaldehyde gel, blotted and probed with *in vitro* transcripts specific to *petA*, *ycf10* or *ycf4*, respectively. B. Western blot analyses of PetA (plastid encoded) and Rieske protein (nuclear encoded), as cyt *b6/f* complex members. Total cellular protein of tobacco wild-type and PEP-lacking material were separated at 15% SDS-PAGE and probed with polyclonal, monospecific antibodies for cytochrome *f* and Rieske protein, respectively.

4.3.1.4. Protein accumulation in wild-type and Δrpo mutant

It is known that steady-state RNA concentrations and rates of mRNA translation are often discordant and substantial amounts of thylakoid proteins are frequently translated from minute RNA amounts (e.g. Herrmann et al., 1992). Immunoblot analysis, however, shows no detectable signals of full-size photosystem I core proteins (genes: *psaA* and *psaB*), CP43 of photosystem II (*psbC*) and cytochrome b_6 (*petB*) of the cytochrome b_6f complex in the PEP-deficient plastids (Fig. 19), although transcripts, even though in reduced amounts, are clearly present (Figs. 15 and 16).

In summary, immunoblot analyses (Fig. 19; *plus* complementing data of DeSantis-Maciossek et al., 1999) demonstrate that plastid-encoded thylakoid proteins are not or only barely detectable in plastids lacking an active PEP enzyme, in contrast to those of nuclear origin (e.g. PsbO) which accumulate to comparable levels in both, wild-type and PEP-deficient mutants. Combined with results from array-based expression profiling and from Northern analyses the data show that the lack of plastid-encoded polypeptides for photosynthesis-related functions in PEP-deficient

tobacco plants is not directly caused by a lack of transcription of their genes but depends on which one of the distinct RNA polymerases produced their mRNAs.

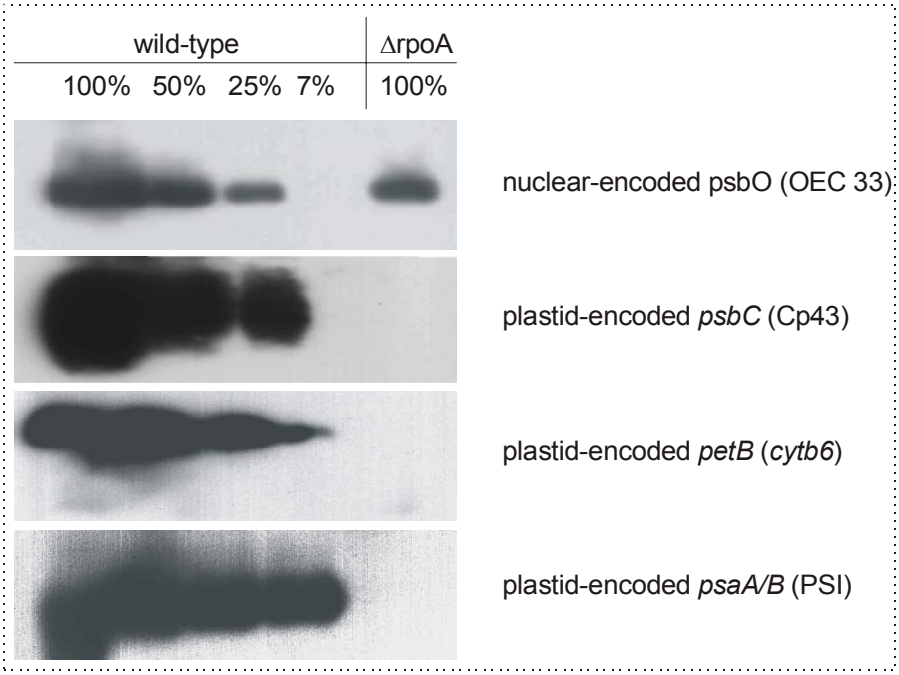


Figure 19. Western analysis of nuclear- and chloroplast-encoded polypeptides. Immunoblots of total proteins from tobacco wild-type and PEP-deficient mutants (Δ*rpoA*). Different amounts of total cellular proteins of wild-type and Δ*rpoA* mutants were loaded per lane and probed with polyclonal, monospecific antibodies against subunit CP43 of photosystem II (gene: *psbC*), cytochrome *b6* (*petB*), reaction centre proteins of photosystem I (*psaA* and *psaB*), and the nuclear-encoded 33 kDa polypeptide (*psbO*) of the oxygen-evolving system, respectively.

4.3.2. Genes encoding components of the genetic apparatus

Since translation takes place in PEP-deficient plastids (Allison et al., 1996; DeSantis-Maciossek et al., 1999), it is not surprising that run-on analyses reveal transcription of all genes coding for components of the organelle translation machinery. Consistent with previous observations (Hajdukiewicz et al., 1997; DeSantis-Macciosek et al., 1999), genes such as *rpoB, C1 and C2*, *trnL*(CAA), *trnR*(UCU), *trnS*(GCU) and *trnV*(GAC), *rps3*, *rps7*, *rps8*, *rps11*, *3'rps12*, *rps18* and *rps19*, *rpl16*, *rpl20*, *rpl22*, *rpl32*, *rpl36* and *infA* (which specifies a translational initiation factor) are transcribed at significantly higher rates in PEP-deficient mutants (Figs. 14 B and 15 A). On the other hand, comparable to genes for photosynthesis-related components shown above, relative transcription rates for several house-keeping genes in wild-type and mutants are quite similar (i.e., *rrn23*, *rrn16*, *rrn5* and *rrn4.5*, *trnG*(UCC), *trnI*(CAU), *trnL*(UAG), *trnR*(ACG), *trnS*(GGA), *trnT*(UGU), *trnP*(UGG) and *trnW*(CCA), *rps2*, *rps4*, *5'rps12*, *rps15*, and *rpl33*). Remarkably, various genes [i.e. *trnA*(UGC),

trnC(GCA), *trnD*(GUC), *trnE*(UUC), *trnF*(GAA), *trnG*(GCC), *trnH*(GUG), *trnK*(UUU), *trnL*(UAA), *trnM*(CAU), *trnfM*(CAU), *trnN*(GUU), *trnQ*(UUG), *trnS*(UGA), *trnT*(GGU), *trnV*(UAC), *trnY*(GUA), *rps14*, *rps16, rpl2*, *rpl23* and *ycf14*(*matK*), which is assumed to be involved in splicing of certain group II introns], display higher relative transcription rates in wild-type than in the mutant leaf tissue (see Figs. 14 B and 15 A).

As noted for genes for photosynthesis-related components, the transcript quantities again often do not correspond to transcription rates. For instance, *trnS*(GGA) shows similar relative transcription rates in mutant as in wild-type plastids (Figs. 14 B and 15 A), but the resulting transcript quantities hybridising to the *trnS*(GGA) gene fragment turned out to be higher in wild-type than in mutants (Figs. 14 C and 15 B). On the other hand, the *rpl2* gene is transcribed at nearly twofold higher rates in wild-type plastids (Figs. 14 B and 15 A), but in mutants the corresponding transcript quantity was found to be even more than twofold that of wild-type plastids (Figs. 14 C and 15 B). The amount of Rpl2 polypeptide, however, is significantly higher in wild-type plastids (DeSantis-Macciosek et al., 1999). This again clearly illustrates that, in analogy to the relationship between transcription rates and resulting RNA levels, only limited conclusions can be drawn from transcription rates or RNA concentrations on the accumulation of the corresponding protein, that correlates with recently published data (Eberhard et al., 2002).

4.3.3. Heterogenic operons encoding components of both, the photosynthesis and genetic apparatus

Different from plastid operons carrying genes either solely for components of the genetic machinery or for photosynthesis-related functions, *rps14,* which encodes the protein CS14 of the 30S ribosomal subunit and is crucial for the assembly of the ribosome (Maki et al., 2000), is transcribed as part of an operon also containing the two photosystem I genes *psaA* and *psaB* (see above; Meng et al., 1988). The operon has no obvious NEP promoter and seems to be transcribed preferentially by PEP (Hajdukiewicz et al., 1997) which in wild-type plastids mainly results in the tricistronic *psaA*/*psaB*/*rps14* message driven by a PEP promoter upstream of *psaA* (Meng et al., 1988). Data presented here demonstrate that *rps14* also is transcribed in PEP-deficient mutants but at a significantly lower rate compared to wild-type (Fig. 14 B

and 15 A). However, in contrast to the wild-type pattern no signal for the dominant 5.2 kb *psaA*/*psaB*/*rps14* cotranscript can be found in PEP-deficient leaf tissue as judged from Northern analyses, and even no monocistronic *rps14* mRNA (of approximately 0.5 kb) resulting from a potential NEP promoter is detectable (Fig. 17; Hajdukiewicz et al., 1997). In the absence of PEP, *rps14* mRNA seems to accumulate only as part of a large polycistronic *ycf3*/*psaA*/*psaB*/*rps14* RNA which is consistent with the presence of a putative NEP promoter upstream of *ycf3* (Summer et al., 2000). It should be noted that other transcripts with sizes of 2.2 and 4.2 kb not present in wild-type are detectable in mutant leaf material. Primer extension analysis of the region upstream to *rps14* generates a signal specific for mutant RNA at position –152 with respect to the translational start codon of *rps14* (Birgit Profanter, Diploma thesis, Munich, 2001). In addition, various other primer extension signals (i.e. at positions 173, 198, 211 and 250) which are more pronounced in RNA from PEP-deficient mutants than in wild-type RNA are also clearly detectable (Birgit Profanter, Diploma thesis, Munich, 2001). Since sequences surrounding these signals are identical in both, wild-type and mutant, premature termination of cDNA synthesis due to secondary structures in mRNAs can not account for those differences. Together with data from Northern analysis this strongly suggests that transcripts are subject to differential processing not only in a quantitative but also in a qualitative manner depending on the polymerase that produced them. Interestingly, the CTT motif, which has been found to constitute the core of the NEP promoter driving transcription of *rrn16* of tobacco (Nt Prrn16-64; Vera and Sugiura, 1995; Liere and Maliga, 1999) is found four nucleotides upstream of the primer extension signal at position –152, within the coding region of *psaB* (Birgit Profanter, Diploma thesis, Munich, 2001). Individual primer extension signals therefore alternatively may result from (specific or ”unspecific”) initiations of transcription by NEP.

Two editing sites have been identified in *rps14*-coding transcripts of tobacco (Hirose et al., 1998). Posttranscriptional C-to-U conversions at both positions lead to the restoration of codons for leucine residues conserved in *rps14* genes of other species. It is generally assumed, and has been experimentally shown for some cases (Bock et al., 1994; Zito et al., 1997; Sasaki et al., 2001) that such conserved amino acid residues restored by editing are of structural or functional importance for the affected polypeptide (Maier et al., 1996). Both sites of *rps14* transcripts are edited even in

PEP-deficient plastids. Also, 3‘ ends of the *rps14* message are found to be identical in both wild-type and mutant plastids (data not shown, Birgit Profanter, Diploma thesis, Munich, 2001). Thus, expression of a functional CS14 protein (Hirose et al., 1998), which is important in early stages of ribosomal 30S subunit assembly (Wimberly et al., 2000), appears to be guaranteed and efficient translation may compensate the only minute amounts of *rps14*-containing mRNA in the mutants. Alternatively, CS14 does, in discordance to published data (Wimberly et al., 2000), not play an important role in ribosome function. It is also possible that CS14 is translated from a polycistronic precursor or that a substitute is imported from the cytosol, encoded in the nucleus, as described for several genes of plastid origin (reviewed in Bruce, 2001).

4.3.4. Genes specifying other functions

As illustrated in Figure 14, *clpP*, encoding a catalytic subunit of the ATP-dependent Clp protease (Maurizi et al., 1990), belongs to those genes with quite different transcription in PEP-deficient leaf material. In wild-type, *clpP* is transcribed at significantly lower rates than in mutant plastids (Figs. 14 B and 15 A) and the resulting transcript levels determined by the array approach (Figs. 14 C and 15 B) are lower, too. Northern experiments revealed that the *clpP* mRNA is clearly detectable in wild-type (Fig. 20 A, Serino and Maliga, 1998) consistent with the finding that one of the four *clpP* promoters is recognized by PEP (Hajdukiewicz et al., 1997). Since three of the four promoters described upstream of *clpP* are NEP-dependent (PclpP-153; Nt PclpP-511 and Nt PclpP-173, Hajdukiewicz et al., 1997), and NEP seems upregulated in ∆*rpo*-material (Hajdukiewicz et al., 1997; DeSantis-Maciossek et al., 1999), an increased expression of this gene is not surprising.

However, does the increase in transcript amounts entail an increase in protein abundance? To test this, accumulation of both, the plastid-coded (pClpP) and the nuclear-coded subunits (nClpC, nClpP) were investigated immunologically (Fig. 19 B). All three subunits accumulate to much higher levels in PEP-deficient material than in wild-type, which is in contrast to all other plastid proteins tested (see above). The fact that the nuclear subunits also accumulate conspicuously might be explained as a compensatory expression, in order to maintain the known 1:1 stoichiometry of nClpX and pClpP. The overall boost of this protease could be a physiological answer

to the need to dispose of the many proteins produced that can not be assembled into a complex in the mutant.

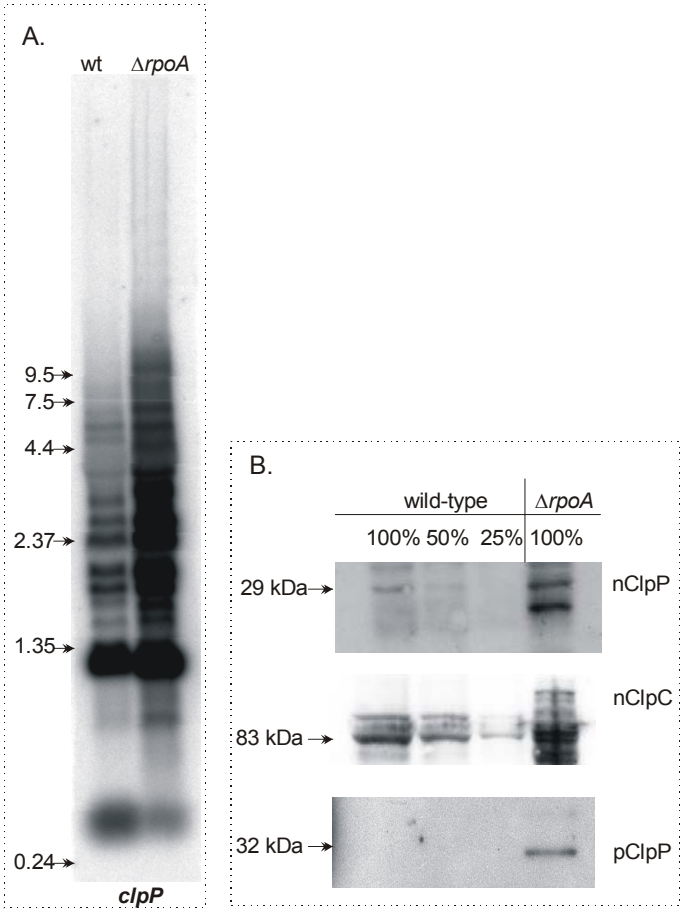


Figure 20. Northern and Western analysis of the *clp* transcripts and encoded proteins. (A) Equal amounts of total celullar RNA (5 µg) were electrophoretically separated in a 1% agarose/formaldehyde gel and probed with an *in vitro* generated transcript specific to the *clpP* mRNA. (B) Western analysis of total proteins from tobacco wild-type (WT) and PEP-deficient mutants (Δ*rpoA*). Total cellular proteins of wild-type and Δ*rpoA* mutants were loaded and probed with polyclonal monospecific antibodies against the nuclear-encoded subunit nClpP and nClpX that are posttranslationally imported into chloroplasts, and pClpP that represents the chloroplast encoded subunit. nClpX and pClpP are present within the organelle in a 1:1 ratio (Sokolenko et al., 1998).

On the other hand, *accD*, encoding one of the subunits of the prokaryotic-type acetyl-CoA carboxylase involved in lipid biosynthesis (Sasaki et al., 1993) shows high relative transcription rates in both, mutant and wild-type leaf tissue (Figs. 14 B and 15 A), but its mRNA accumulates in larger quantities only in PEP-deficient plastids as revealed by Northern analyses (data not shown, Hajdukiewizc et al., 1997; Krause et al., 2000).

4.3.5. Open reading frames

In addition to defined genes and well conserved *ycf*s, the tobacco plastid chromosome harbours 11 annotated ORFs (Shinozaki et al., 1986). Comparative analysis of chromosome regions encoding these ORFs in various plant species indicates that, although transcribed, most of them probably do not encode functional peptides (Schmitz-Linneweber et al., 2001). Whereas some of these ORFs exhibit relative transcription rates with no major difference between wild-type and PEP-deficient mutants (e.g. ORF70A and ORF115; Figs. 14 B and 15 A), some of them

exhibit higher expression rates in wild-type plastids (i.e. ORF105 and ORF75), others (i.e. ORF31 and 70B) show increased relative transcription rates in plastids lacking the PEP enzyme. As shown in Figure 14, transcription rates and RNA levels of these ORFs often correspond to those of flanking genes.

Of the more conserved hypothetical reading frames, *ycf2* has been analysed in greater detail. This gene has been found to be essential for plastid biogenesis, even though its function is unknown (Drescher et al., 2000).

Applying array techniques, the transcription rates of *ycf2* turned out to be higher in PEP-deficient mutants than in wild-type (Fig. 14 B). The most interesting difference was observed when RNA was hybridised against different parts of *ycf2*. In the PEP-deficient tissue, the transcription rate determined with a probe specific to the 5' coding region of *ycf2* was significantly lower than the transcription rate measured with a 3' specific probe (Fig. 14 B). This suggests that some internal promoter could be hidden within the coding region. Therefore a detaild transcript analysis within *ycf2*, as one of the largest open reading frames encoded in the plastid genome, was performed.

Transcript analysis further corroborated this assumption. Four different probes corresponding to different regions of the *ycf2* gene have been used in Northern experiments (Fig. 21). In general, transcript accumulation in wild-type is weaker than in the mutant (Fig. 21; Hajdukiewicz et al., 1997) and transcript patterns are not identical (see arrowheads in Fig. 21).

In a previous study, large transcripts initiating at a typical NEP-promoter upstream to *ycf2* have been described (Nt Pycf2-1577, Hajdukiewicz et al., 1997). These transcripts were also detected in the present work, being the only signals detected with a probe covering part of the 5' coding region (data not shown). When using 3' terminal probes (Fig. 21), additional, substantially more abundant signals were obtained. Most prominent are transcripts of 2.1 and 3.4 kb in the mutant. These signals are obtained with all four 3' terminal probes used, suggesting that this transcript covers the entire 3' region including the *ycf15* reading frame, the latter most likely being a pseudogene (Schmitz-Linneweber et al., 2001). The region covered by

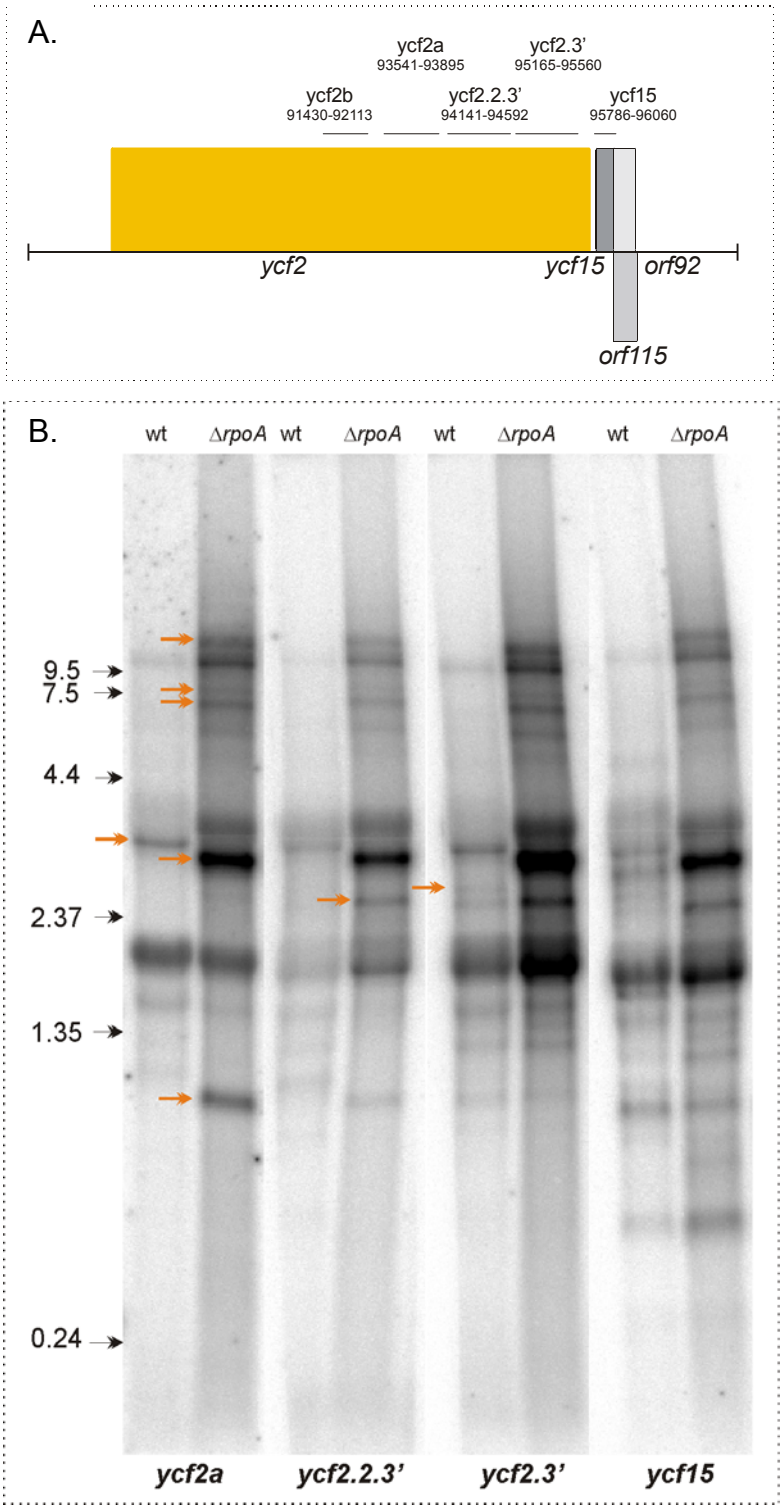


Figure 21. A. Schematical overview of the *ycf2* part of the plastid genome that has been covered by PCR products (shown as black lanes; amplification products are determined by bp position with respect to the tobacco plastid genome; Shinozaki et al., 1986) used for Northern analyses.
B. Northern analysis of the 3' region of *ycf2*. 5 µg of total cellular RNA from tobacco wild-type (wt) and PEP-deficient mutants (Δ*rpoA*) have been electrophoretically separated in agarose/ formaldehyde gels and probed with *in vitro* transcripts specific to the mRNA of *ycf2*. The four autoradiographs go back to the same gel and have been aligned here according to the marker (indicated on left). Orange arrows denote differences between wild-type and mutant transcript patterns.

the four probes is indeed 2.2 kb long and with probes corresponding to 3' terminal sequences (ORF92, see Fig. 14) no signals were detected.

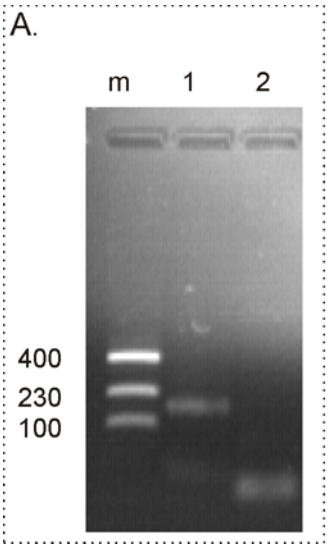
In order to exactly determine the 3' transcript end, a method based on ligation of an oligonucleotide to 5' ends of transcripts was used (Hayes et al., 1999, see Section Nr. 3.2.13.). Total plastid RNA was ligated to an RNA oligonucleotide of a non-plastid

sequence. This oligonucleotide can only be attached to the 3' end of an RNA. The RNA population was then reverse transcribed with a DNA oligonucleotide (PM13) complementary to the RNA-oligo and served as a template for amplification of the *ycf2* 3' terminal region. Combinations of gene-specific (ycf15-5'for and ycf15for2) and linker-specific primers (PM13) led to the amplification of several distinct bands (Fig. 22), which were cloned separately into pBSC-SK+ and sequenced. In this way, three recombinant DNAs were obtained that carried *ycf2*-specific sequences that terminated at positions 96038, 96168 and 96176 of the plastid chromosome (Acc. Z00444, Shinozaki et al., 1986). The low expression of *ycf2* in wild-type plastids probably was the reason for the low yield of clones. Position 96038 falls into the *ycf15* coding region. It has been speculated previously that *ycf15*, as it does not constitute a proper gene, is a conserved 3' regulatory element of *ycf2* (Schmitz-Linneweber et al., 2001). Often, regulatory sequences in the 3' UTR form extensive stem-loops, which serve to stabilize the RNA by impairing exonuclease activity (Drager et al., 1999). To test if any of the mapped 3' ends coincide with such a secondary structure, the region between nucleotide positions 96001 and 96051 (Shinozaki et al., 1986) was analysed by *mfold version 3.0* software (http://bioinfo.math.rpi.edu/~mfold/rna/form1.cgi).

According to the 3' ends detected and the lengths of the bands found in Northern experiments, the 5' end of the 2.1 kb transcript maps to a site inside of the *ycf2*, approximately around nucleotide position 93400 of the tobacco plastid chromosome. Possibly, this transcript is a processed form of the larger 3.4 kb transcript, for which the calculated start point is approximately position 94200. Capping experiments will aid to identify the transcription initiation site.

The data obtained suggest a functional divergence of the 5' and the 3' part of *ycf2*. To test this, separate BLAST searches (Altschul et al., 1990) were carried out for the part upstream and downstream of the potential transcript initiation site at nucleotide position 94200. As described previously the 3' part of *ycf2* exhibits similarity to the zinc-metalloprotease FtsH (Wolfe et al., 1994). Upstream parts however exhibit no significant similarities to other known proteins. Further experiments are needed to explain the curious bias for transcription of the *ycf2* gene.

A.



B.

```
95641 aacggaacgc tattggatca aatgacaaag acattgttga gaaaaagatg gcttttcccg
95701 gatgaaatga aaattggatt catgtaacag gagaaagatt tcccattcct tagccggaaa
95761 gatatgtggc catgaaagag ggattaagtg gaacagaatt gactgggtgg tagagtcgtg
95821 gaaacgcttg tttcttccat attttggacc ttagctccat ggaagaatat gttactgctg
95881 aaacacggaa gaattgaaat cttagatcaa aacactatgt atggatggta tgaactgcct
95941 aaacaagaat tcttgaacag caaacaacca gttcagatat tcacgaccaa gaagtactgg
96001 attctctttc ggataggccc tgaaaggaga aggaaggctg gaatgccaac aggcgtctat
96061 tatattgaat ttacccgata gtccccattt tgggaacgtc cagtgccaaa gtcactgaat
96121 gggtaagtcg ccaatccctg gactatgtaa tgtactttat ctgctgggtt acgggcgggc
96181 attttaccag aggtttctaa tctacccttg tgtgattcct gttgaagcat atactcgggg
96241 ggtgggtgca gggcggacga ttttaaagcg gactccccat tcattagata gagaagatca
96301 ccaagatttc gcgatccgct gccgaattta ttccaattcc aagagctcgg atcgaatcgg
96361 tatatcaata ccgattcgat ccgagctctc ttattgagaa tgctcattca atgagcattc
```

ycf15

© orf115

orf92

C.

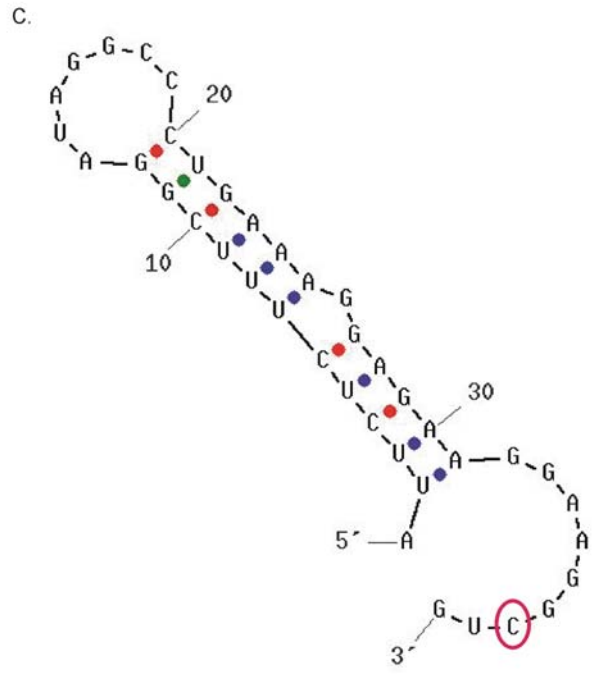


Figure 22. Molecular analyses of *ycf2* 3' UTR. (A) PCR products obtained when the RNA oligonucleotide-specific primer (PM13) together with a gene-specific primer were used for amplification (lane 1). The no-template negative control is shown in lane 2. Sizes of marker bands are shown at the left. (B) Graphical overview of the segment of the tobacco chloroplast genome (Shinozaki et al., 1986) with the respective positions of determined 3' transcripts ends (bases circled in red position: 96038, 96168 and 96176). (C) Predicted secondary structure of the determined 3' end sequence. The secondary structure shows a typical stem-loop structure with described processing end indicated with a red circle (Monde et al., 2000, Hicks et al., 2002).

4.3.6. Accumulation of aberrant transcripts for the *rpoB* operon in Δ*rpoA* mutants

The transcript pattern of the *rpoB* operon, which comprises the genes *rpoB*, *rpoC1*, *rpoC2*, coding for the β, β' and β'' subunits of the PEP-enzyme, has been analysed by Northern experiments. It has been assumed that this operon is transcribed solely by the NEP-polymerase (Hu et al., 1990, Hess et al., 1993, Hajdukiewicz et al., 1997). Thus, it is not surprising that transcript patterns in wild-type and Δ*rpoA* material are quite similar (Fig. 23), with only a few exceptions. First, longer transcripts at 6.9 and 6 kb are more abundant in the mutant, consistent with the general finding that processing seems to be slower in this material. In addition, a few bands between 0.5 and 2.3 kb are specific or far more abundant either in wild-type or in mutant material (arrows in Fig. 23). These RNA species are too small to encode any of the three genes and hence are probably degradation products. Differences in RNA degradation have been noted also for other genes (see above).

More interesting is a band of 2.3 kb which strongly accumulates in Δ*rpoA* tissue when probed with *rpoC1 in vitro* transcript. This specific transcript does not correlate to the spliced form of *rpoC1* mRNA (1713 bp, detected in both materials) nor to the unspliced form (2805 bp, detected in both materials, somewhat stronger in wild-type). Possibly, this transcript results from internal cleavage of *rpoB* mRNA or, alternatively, from a promoter inside *rpoB*. The nature of this band can not be completely resolved at the moment. Aside of differences between mutant and wild-type plastids, some general conclusions can be drawn from the pattern observed, in particular as no Northern experiments with comparable resolution regarding this operon have been presented so far in the literature. For instance, the bands at 6.0 and 6.9 kb are apparently covering *rpoB* and *rpoC1*, but not *rpoC2*, for which they were not detectable.

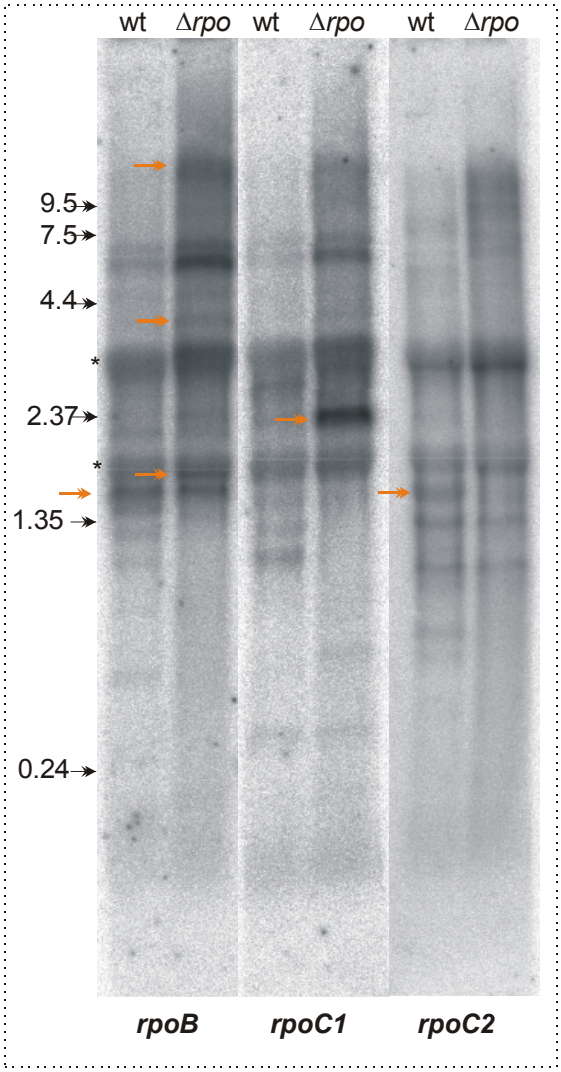


Figure 23. Northern analysis of the *rpoB* operon. 5 µg of total cellular RNA from tobacco wild-type (wt) and PEP-deficient mutants (Δ*rpoA*) have been electrophoretically separated in agarose/ formaldehyde gels and probed with *in vitro* transcripts specific for the mRNAs of *rpoB, rpoC1* and *rpoC2*. The three autoradiographs represent the same gel and have been aligned according to the markers (indicated on left). Arrows denote differences between wild-type and mutant transcript patterns.

A precursor comprising transcripts of all three *rpo* genes can be found at aproximately 13 kb, although the signal is very weak in the *rpoB* autoradiograph and barely detectable when probed for *rpoC1* and *rpoC2*. Nevertheless, it can be assumed that starting from a large precursor, first a dicistronic *rpoB/rpoC1* message is produced. In contrast to the identification of precursors, it is more difficult to find candidate transcripts representing the monocistonic mRNAs of the three genes. The *rpoB* reading frame is 3212 bp long. Curiously, bands detected with the *rpoB*-specific probe are too small (most prominent bands at 1.5 and 2.1 kb) and the reason for their accumulation is difficult to explain, unless these are again some degradation products. Only the signal corresponding to the ribosomal RNA (nuclear 25S rRNA, which is 3.3 kb long; in Fig. 23 marked by an asterisk) could potentially hide a band corresponding to the monicistronic message. For the reading frame *rpoC1* the monocistronic, spliced transcript (1731 bp) might also be hidden by ribosomal RNA (plastid 16S rRNA and nuclear 18S rRNA; in Fig. 23 marked by an asterisk). The stringency of washes after hybridisation was relatively high (65°C/0.2x SSC), which

usually prevents extensive cross-hybridisation. However, in cases where the target transcript is very rare, like for *rpo* genes, cross-hybridisation might still interfere with the detection of transcripts with the same size as ribosomal RNA. Specific blocking of ribosomal rRNA might help to overcome this problem. It still remains possible, that no monocistronic forms of *rpoB* and *rpoC1* exist and that translation is based on the dicistronic form. Finally, for *rpoC2* with a length of 4178 bp, ribosomal rRNA can no longer serve as a candidate preventing detection of monocistronic transcript. Unfortunately, no clear signal larger than 3.3 kb was observed. Only some diffuse bands in the high molecular weight range indicate *rpoC2*-specific transcripts, but surely not monocistronic forms. Thus, expression of *rpoC2* appears to be very weak, weaker than of other genes of the operon and no good candidate for a monocistronic nor for a polycistronic transcript could be identified. This discrepancy in transcript abundance for different *rpo* genes suggests that the stoichiometry of the polymerase subunits ($\alpha\beta$ $\beta'\beta''$) is maintained by translational and/or post-translational regulation.

Altogether, *rpo* gene expression is very weak and differences between *rpo* expression in wild-type and Δ*rpoA* material are small, in accordance with the expectation that this gene is transcribed from a NEP promoter. All three transcript patterns investigated exhibit an interesting (weird) accumulation of transcripts that are too short to cover the coding regions of the respective genes. An explanation for the abundance of these RNAs particularly occurring for this operon is speculative, but may lie in a slower degradation than usually occurring with RNA from other plastid operons.

4.4. PEP promoter studies

Plastid chromosomes are densely packed with genes, whereas the intergenic regions comprise only a relatively minor part of the whole circle, 26,9% in tobacco. *Cis* elements important for transcriptional initiation are predominantely located in these latter regions (Sugita and Sugiura, 1996). A particularly important promoter is located upstream of the *rpo* gene cluster, as transcription of these genes, which encode subunits of the PEP-polymerase, will eventually influence the overall transcriptional activity of the plastid. It was hypothesized that the *rpoB* operon is transcribed exclusively by NEP (Hajdukiewicz et al., 1997) and *in vitro* data indeed suggested that the *rpoB/C1/C2* message is generated from a canonical NEP promoter (Liere

und Maliga, 1999). This promoter was mapped at position -345 of the *rpoB* start codon and contains the CAT core motif typical for a NEP promoter (Weihe und Börner, 1999), but in addition other regulatory upstream elements are present in 5'-UTR of *rpoB* gene (Inada et al., 1997). In this study, an *in vivo* analysis of the *rpoB* promoter region has been performed.

4.4.1. Cloning strategy and plastid transformation studies

For the *in vivo* investigation of the *rpoB* promoter region, a plasmid carrying a point mutation within the defined core promoter CAT-sequence was prepared. The point mutation was obtained using a modified method provided by Clontech (St. Louis, USA). A 2079 bp fragment excised from clone BaT1 (Sugiura et al., 1986) with *HaeIII*/*SalI* endonucleases was inserted into the pBSC-SK II$^+$ vector to yield vectror pPrpoB-345. The fragment included the *rpoB* and the upstream *trnC(GCA)* gene together with the complete intergenic region. The mutation was introduced by annealing single stranded pPrpoB-345 (generated by denaturation for 15 min at 95^{o}C) with two oligonucleotides, one carrying the point mutation (the mutagenic primer "point mut wg": 5'-CGTCCAACCTTACACATAATAGT**C**TTATTACCATCTTT ACC-3') and a second one eliminating a restriction site (*PstI*) in the vector backbone (the so-called selection primer, see Fig. 24). Using T4 DNA polymerase and T4 DNA ligase, the complementary strand was filled in and circularised. Subsequently the completed plasmid was transformed into *E. coli* mutS cells that are incapable of repairing the mismatches introduced by primers. From this transformation, plasmids were isolated and digested with *PstI*, thereby linearizing the parental plasmids. After restricted digest with PstI of second transformation into *E. coli* DH5α were done, it allowed us to select for colonies carrying the mutation. 99% of all colonies tested carried the mutation.

In plastid transformation experiments, a resistance marker, the *aadA* gene coding for the adenylyltransferase, which detoxifies spectinomycin and related antiobiotics, was used to select for plastid transformants (Goldschidt-Clermont, 1991). This *aadA* gene equipped with 5' and 3' regulatory sequences (Koop et al., 1996) was introduced into the *SmaI* restriction site at the position 343 bp upstream of the mutated motif. The promoter of the resistance cassette (16S rrn promoter) faces in the direction of the *trnC(GCA)* gene, opposite of the transcription direction from PrpoB-345. The

sequence of the entire construct was checked by sequence analysis. No further mutation aside of the introduced one was detected.

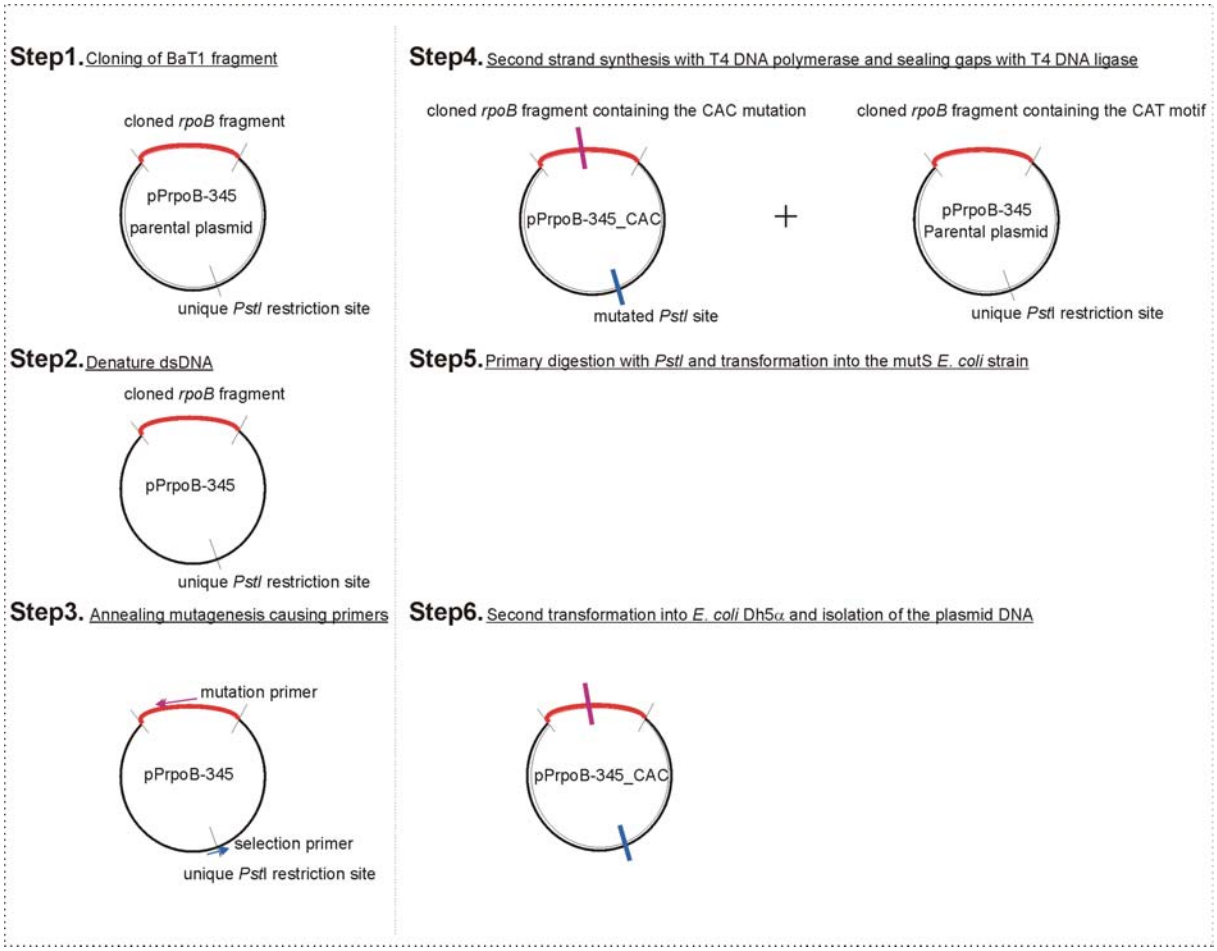


Figure 24. Cloning strategy to introduce a point mutation into the *rpoB* promoter -345. A plasmid containing the *rpoB* gene pPrpoB-345, was used for preparation of the pPrpoB-345_CAC plasmid via annealing of two oligonucleotides. The mutagenic primer exchanges bases in the core promoter region, the CAT motif into CAC. The selection primer was used to remove a *PstI* site in the backbone vector by mutating the *PstI* site. Mixture of plasmids (pPrpoB-345 as parental plasmid and pPrpoB-345_CAC as mutated plasmid) after first transformation into mutS *E. coli* cells were treated by *PstI* digestion. By this, only the mutated plasmid pPrpoB-345_CAC was amplified after retransformion into DH5α *E. coli* cells.

4.4.2. Analysis of *rpoB* promoter transformants

4.4.2.1. Homoplastomy check of transformed lines

Four spectinomycin-resistant transformants have been obtained after plastid transformation, designated mA, mC, mE and mG. After two cycles of regeneration, these plants were tested for correct integration of the transgene by PCR using *aadA*-specific primers in combination with primers outside the constructs (Fig. 25). PCR products spanning the flanking regions were obtained for three of the four lines (not for mA), which verifies that the transgene has been inserted correctly into the plastid

chromosome (Fig. 26).

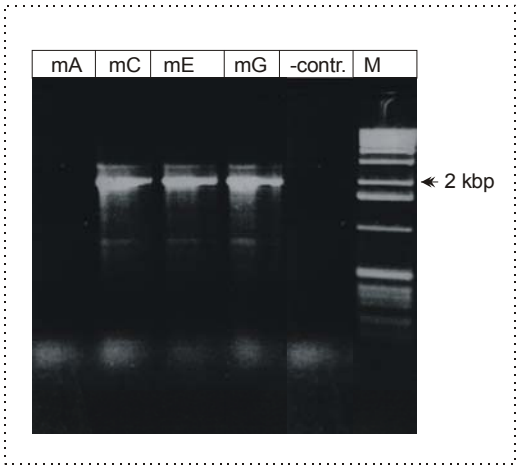


Figure 25. PCR analysis of transplastomic PrpoB-345 mutants. PCR products amplified with gene-specific and *aadA*-cassette primer (cpoBleft and aadAoutref) were separated in an 1% agarose gel. Amplified products of 2643 bp show the correct integration of the *aadA* cassette into the plastid chromosome. Plants mC, mG and mE were used for sequence analysis.

In addition to PCR-analysis, Southern analysis was performed to test integration of the transgene. After restriction of total plant DNA with *HindIII*, a band of 4.5 kbp, representing the restriction fragment containing the *aadA* gene, was found in the mutant lines mC, mE and mG, but not mA and wild-type (Fig. 26). Southern and PCR-analysis showed that the *aadA* cassette is present, but do not indicate, whether the point mutation has been successfully introduced into the plastid chromosome as well.

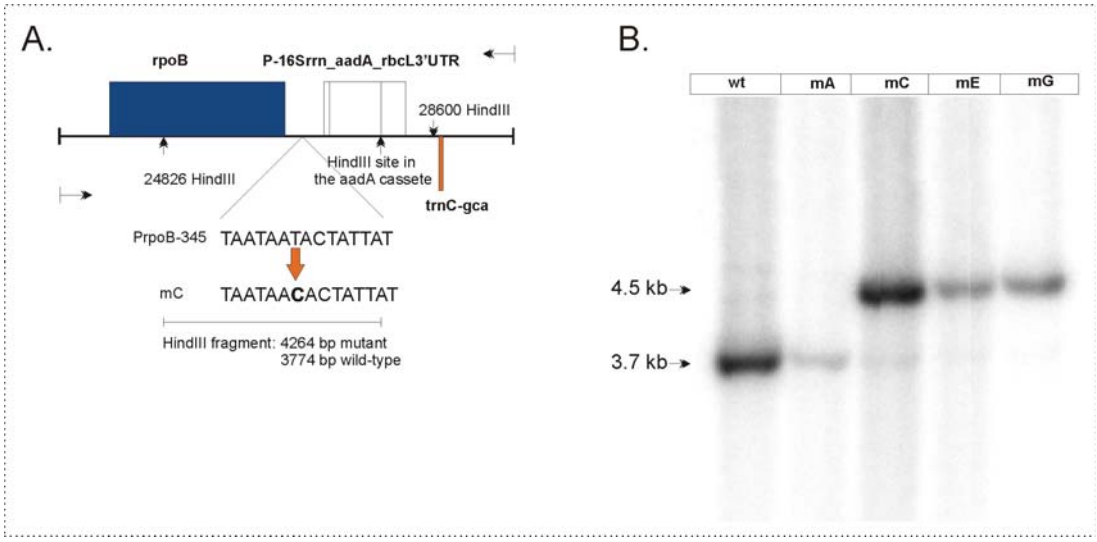


Figure 26. Southern analysis of transplastomic *rpoB* promoter mutants. (A) Schematic presentation of the *rpoB* promoter region including the introduced point mutation. *HindIII* restriction sites are indicated as well as the lengths of restriction fragments (arrows). (B) 5 µg of total genomic DNA from the transplastomic mutants mA, mC, mE and mG was digested with *HindIII* and separated in a 0.8% agarose gel. A gene-specific, radiolabeled probe including the promoter region PrpoB-345 was used for hybridisation. The sizes of bands are indicated at the left. This experiment was performed to check for homoplastomy of the transplastomic material. The digestion pattern indicates the homoplastomic status of the point mutation plant material (line mG).

To test this, primers aadAoutrev and the gene-specific primer HindIIIrpoB were used to amplify the *rpoB* promoter region of mutant mC. The PCR products were sequenced directly. The sequence confirmed the successful exchange of the wild-type T residue by the introduced C residue. No T signal was observed in the sequence of the mutant (Fig. 27), which again indicates that no wild-type chromosomes are present in this material any more. In contrast to mutant mC, mutants mE and mG did not exhibit a point mutation at the expected position (data not shown). Apparently, recombination between the *aadA* cassette and the point mutation occurred during or after transformation.

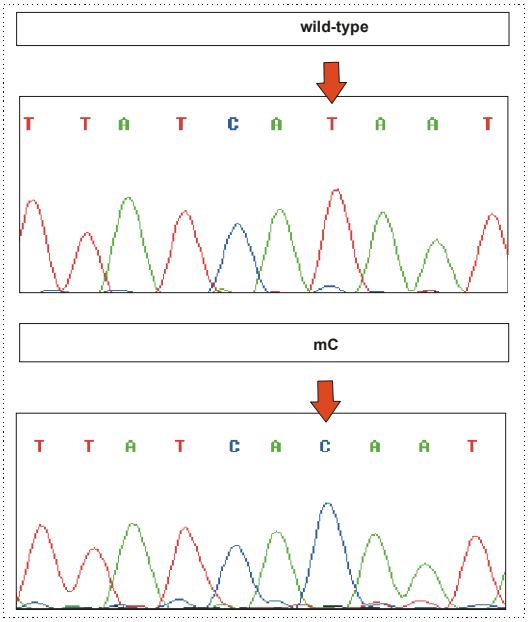


Figure 27. Sequence analysis of the *rpoB* promotor region of the mutant mC. Amplification products of the *rpoB* promotor region was obtained by PCR from total DNA of wild-type and the mutant mC using primers HindIIIrpoB and aadAoutrev and sequenced directly. Partial chromatograms show that the core promoter sequence "CAT" was sucsessfully replaced by "CAC" (arrow).

4.4.2.2. Phenotype characterisation of *rpoB* promoter mutants

Interestingly, the mutant plants did not show any phenotypic variation in comparison to wild-type. Overall morphology, leaf formation and colour as well as the ability to grow under autotrophic conditions were identical between wild-type and mutant mC (Fig. 28), different from known *rpo* knock-out plants (DeSantis Maciossek et al., 1999). This is a first indication that the mutation does not impair *rpoB* expression.

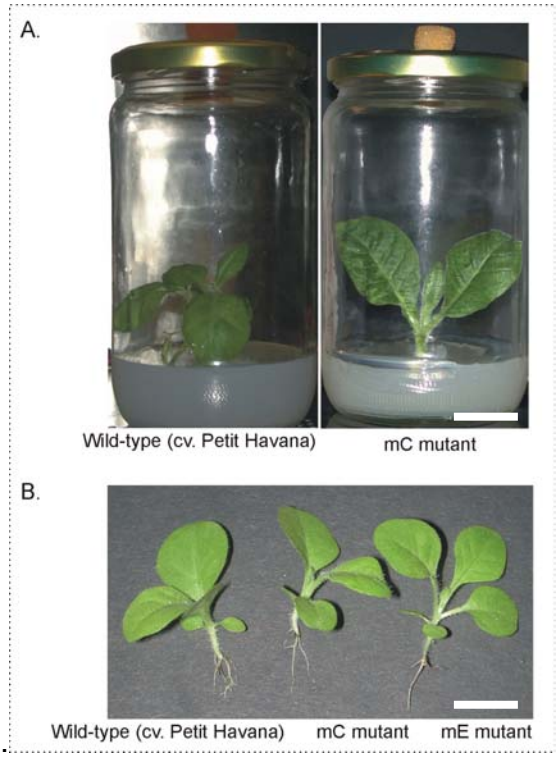


Figure 28. Comparison of wild-type and mutant phenotype. (A) 14 day-old seedlings of wild-type and the mC mutant have been cultivated in a growth chamber on MS-medium. The plants represent the outcome of two cycles of regeneration. (B) Three-week-old plants of wild-type, the mC and mE mutants (M2 generation) cultivated under autotrophic conditions. Size bar in A represents 1.8cm and in B 2cm.

4.4.2.3. Transcriptional characterisation of the *rpoB* promoter mutant

A defect in expression of *rpoB* should result in a pronounced alteration of gene expression of the entire plastid chromosome (DeSantis-Maciossek et al., 1999, Liere and Maliga, 1999). To investigate, whether the mutation of the core promoter motif has an impact on transcript accumulation of other genes, macroarray hybridisation was performed (Fig. 29). Even though the overall signal intensity in the mutant is lower, the expression pattern is very similar to wild-type and not at all resembling macroarrays hybridised with RNA from PEP-deficient material (cf. Fig. 14 C). This suggests that *rpoB* expression is not notably impaired by the point mutation.

To evaluate, whether the point mutation has an effect on expression of *rpoB* itself, transcript accumulation was investigated by Northern analysis. Identical amounts of RNA (6 µg) were hybridized with an *rpoB*-specific antisense RNA probe. A signal of approx. 6.9 kb represents a transcript encompassing the genes *rpoB* and *rpoC1*. This RNA-species was detected in wild-type as well as in mC and mE mutant plastids (Fig. 30 A). However, some smaller transcripts, at 0.7, 0.55, 0.45 and 0.35 kb are only detectable in both mutant plastids, but absent in wild-type (arrows at the right in Fig. 30 A).

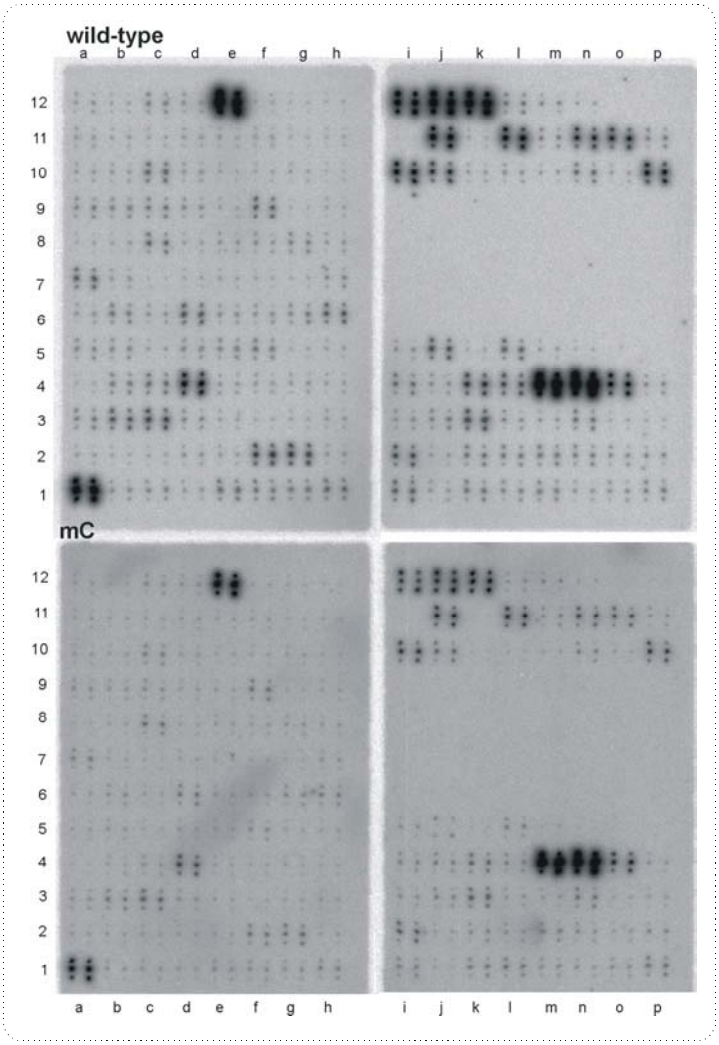


Figure 29. Expression profiles of wild-type and the *rpoB* core promoter mutant mC. Total plant RNA (5µg) was 5' labeled using T4 polynucleotide kinase and hybridized to macroarrays representing all plastid-encoded genes. For preparation of macroarrays and order of genes on the filter see Section 4.2.1.

In addition, overall transcript amounts are greatly increased in the two mutants. Hence, the two mutants are both qualitatively and quantitatively different from wild-type. Also, the transcript pattern found was not similar to that of PEP-deficient material (Δ*rpoA*), which indicates that they show no compensatory NEP overexpression as suggested for Δ*rpo* material (Allison et al., 1996, Hajdukiewicz et al., 1997). As both mutants (mC and mE) are affected similarly, the introduction of the *aadA* cassette seems to be responsible for this effect and not the point mutation. A second gene tested by Northern analysis, *trnC-gca*, seems less affected by the *aadA* insertion, although a moderate increase in signal intensity is observed as well (Fig. 30 B).

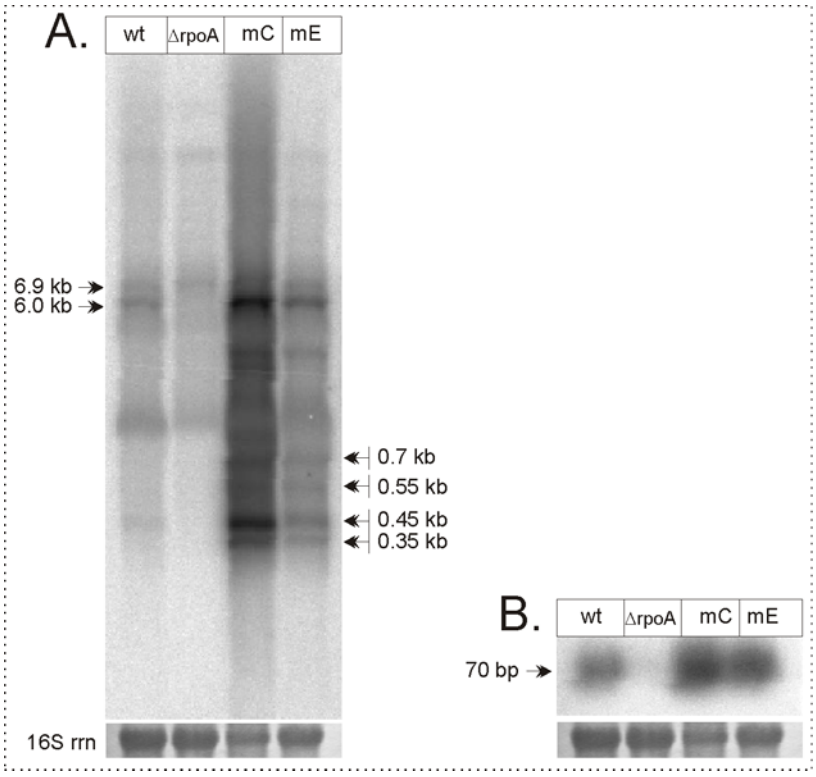


Figure 30. Transcript accumulation for the *rpoB* and *trnC-gca* gene of the mutant mC and wild-type. 6 µg of total cellular RNA of wild type and mutants mC, mE, and Δ*rpoA* were fractionated on a 1% agarose gel, transferred to nylon filter membranes and hybridized with single-stranded antisense RNA probes specific for (A) the *rpoB*, and (B) the *trnC-gca* gene. Wild-type transcripts detected at 6.9 and 6.0 kb show that transcription of *rpoB* gene is not impaired neither in the promoter mutant mC nor in mE (positive control mutant).

All in all, these data do not corroborate previous findings that the promoter PrpoB-345 is crucial for transcription of *rpoB*.

The accumulation of *rpoB* transcripts in the mC mutant suggests that the promoter at -345 is either not important for transcription of this gene or that additional, compensatory transcription initiation sites exist. To answer this question, primer extension experiments have been performed using a primer downstream of PrpoB-345 with RNA from wild-type and mutant plants.

Interestingly, in both RNA-samples no signal corresponding to the promoter detected *in vitro* by Liere and Maliga (1999) was observed. Instead, a signal at –294 bp relative to the start codon of *rpoB* was clearly visible (Fig. 31). The signal is not only present in wild-type, but also in mutants mC and mE as well as in PEP-lacking material. The fact that the signal is present without an active PEP enzyme suggests clearly that it results from a NEP driven transcript. It remains to be shown whether the signal is the result of a primary transcript or represents a processed form of a

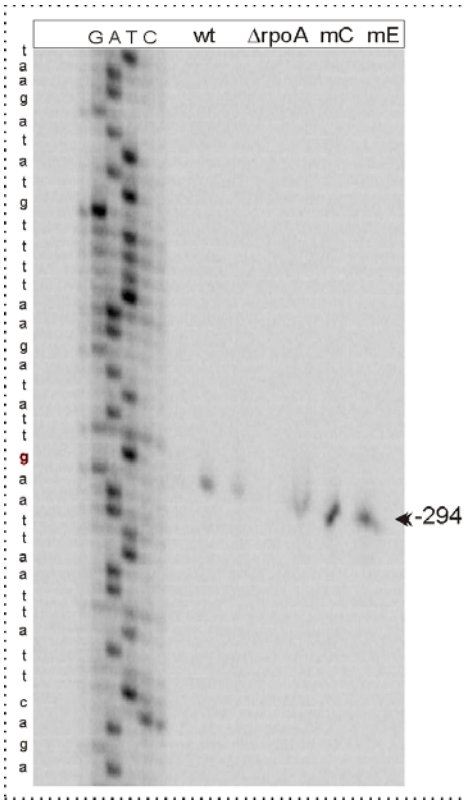


Figure 31. Primer extension analyses of the *rpoB* 5'-UTR. 50 µg of total cellular RNA from wild-type, mutants mC, mG, mE, and a PEP-deficient mutant (Δ*rpoA*) were reverse transcribed using the fluorescence-labeled primer rpoBpe. Extension products were analysed using a LICOR automatic DNA sequencer (model 4000 L^2). As a control, an amplification product spanning the *rpoB* 5'-UTR was sequenced with the same primer. Extension products terminating at -294 bp relative to the *rpoB* start codon were identified.

longer transcript initiating further upstream. Possibly, initiation starts at PrpoB-345, but the transcript is processed too quickly to be detected. However, when primers more closely situated to the *in vitro*-mapped PrpoB-345 were used (primer rpoBpe2 at –299), no signal was obtained as well. Capping experiments will aim to determine, whether the transcript end at -294 is of primary or secondary origin.

4.5. Expression of genes of the translation machinery in Δ*rpo* material

4.5.1. Expression of tRNA genes in tobacco plastids

In tobacco, altogether 30 tRNA-coding genes, 21 ribosomal protein-coding genes and 4 ribosomal rRNA coding genes have been annotated (Shinozaki et al., 1986). Applying 2D-gel electrophoresis and immunoprecipitation techniques, it has been shown that all ribosomal proteins encoded by the plastid chromosome are incorporated into 70S ribosomes (Yamaguchi et al, 2000). In contrast to this, the utilization of plastid-encoded tRNAs has not been well described. Theoretically, the tRNA set of the plastid chromosome is sufficient to decode all protein coding genes (Lohan and Wolfe, 1998). In plastids of parasitic beechdrops (e.g. *Epifagus virginiana*), the suit of tRNAs is greatly reduced, inadequate for decoding the remaining genes (Wolfe et al., 1991). In a situation like this, import of

extraplastidically encoded tRNAs might occur as has been shown for plant mitochondria (Small et al., 1998).

To investigate, whether the annotated tRNAs of the tobacco plastid chromosomes are expressed and what type of RNA polymerase is responsible for their transcription, Northern analysis of plastid RNA was performed. RNA was extracted from tobacco wild-type and PEP-deficient material (Δ*rpoA*). 4 µg of RNA were fractionated in a 1% agarose gel and hybridised with single-stranded antisense RNA probes. The results can be grouped into two classes: (i) transfer RNAs accumulating to similar levels in wild-type and PEP-deficient material (group I, Fig. 32). The genes *trnA(UGC)*, *trnF(GAA)*, *trnI(CAU)*, *trnI(GAU)*, *trnN(GUU)*, *trnP(UGG)*, *trnS(GCU)*, *trnS(GGA)* belong to this group. (ii) Transfer RNAs accumulating in wild-type, but are strongly decreased in *rpo*-deficient plants (Group II, Fig. 33). A fraction of species that belong to this group II (*trnC(GCA)*, *trnG(UCC)*, *trnG(GCC)* and *trnR(UCU)*) is reduced beyond the detection limit or possibly even not present in PEP-deficient tissue.

However, as several plastid-encoded reading frames are translated in Δ*rpo* material (*rpl2*, *clpP*, *petA*), and these reading frames at least theoretically need all tRNA species for translation, it has to be assumed that minor amounts of even these undetectable tRNAs are present in Δ*rpo* material. Alternatively, untypical pairing of other tRNAs or import of tRNAs from the cytosol may explain translation of these reading frames despite the absence of several tRNAs in Δ*rpoA* plastids. In addition to tRNAs that are not detectable, others, like *trnD(GUC)*, *trnE(UUC)*, *trnH(GUG)*, *trnI(gau), trnK(UUU)*, *trnL(CAA)*, *trnL(UAG), trnfM(CAU)*, *trnM(CAU)*, *trnQ(UUG)*, *trnR(ACG)*, *trnS(UGA)*, *trnT(GGU)*, *trnT(UGU)*, *trnW(CCA), trnV(UAC)*, and *trnY(GUA)* are reduced, but still present in Δ*rpo* material.

The reduction of tRNA amounts in the RNA polymerase knock-out suggests that transcript accumulation of tRNA genes belonging to group II depends on the PEP enzyme. Some are part of operons that are known to be transcribed by PEP, like *trnG(GCC)* as part of the *psbD*-operon, or *trnA(UGC)* and *trnI(GAU)* as part of the *rrn*-operon (see below). This explains why expression of group II tRNAs is under PEP control. For others, no PEP promoter has been described so far, like for example for

trnC(GCA) or *trnR(UCU)*. In general, the importance of the PEP enzyme for tRNA expression is surprising, in particular as it has been speculated that the NEP enzyme is responsible for all “housekeeping” genes including those of the genetic apparatus (Hajdukiewicz et al., 1997).

Quite instructive also is the comparison of tRNAs absent in ∆*rpo* tissue with the tRNA set of the parasitic plant *Epifagus virginiana* (Table 3). Of the 7 tRNAs that are not detectable or reduced to less than 20% of the wild-type amount in ∆*rpo* material, 6 are pseudogenes or completely missing in *Epifagus*. Of 7 further tRNAs absent from *Epifagus*, two [(trnT(GGU) and trnT(UGU)] are strongly reduced (5% and 12%, respectively). This implies that of 13 non-functional tRNAs in *Epifagus*, 8 are affected by the loss of PEP activity. This could indicate that PEP-dependent tRNAs are evolutionary less stable than tRNAs also or exclusively transcribed by NEP. The less expressed set of transfer RNAs, however seems not to influence protein accumulation, as illustrated by Western analyses between wild-type and PEP-lacking material. For example, pClpP protein accumulates in the PEP-lacking material tentimes more then in wild-type. Similarily, translation in *Epifagus* takes place, in the absence of certain tRNAs (see Figs. 32, 33 and Table 3).

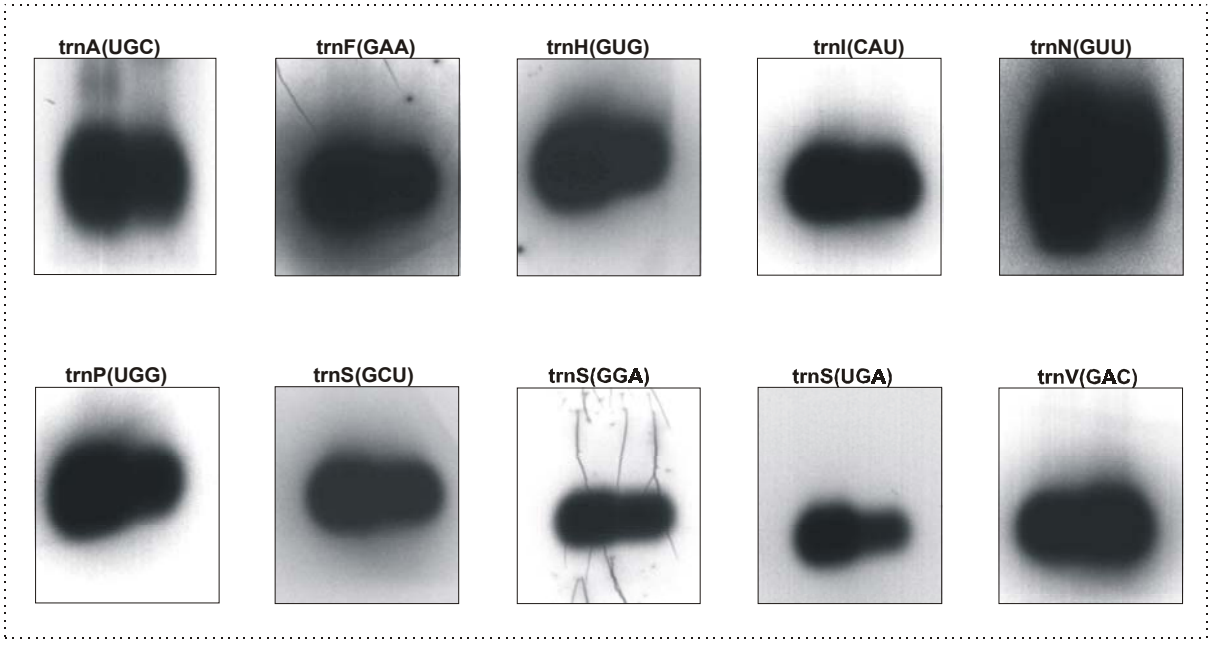


Figure 32. tRNAs accumulating to wild-type amounts in PEP-deficient material (group I tRNAs). 4 µg of total plant RNA from tobacco and ∆*rpoA* leaves were fractionated on a 1,2% agarose gel, blotted to nylon membranes and hybridised with strand-specific, single-stranded RNA probes. Due to high accumulation of respective tRNAs and their small size, the slots of blots are not indicated.

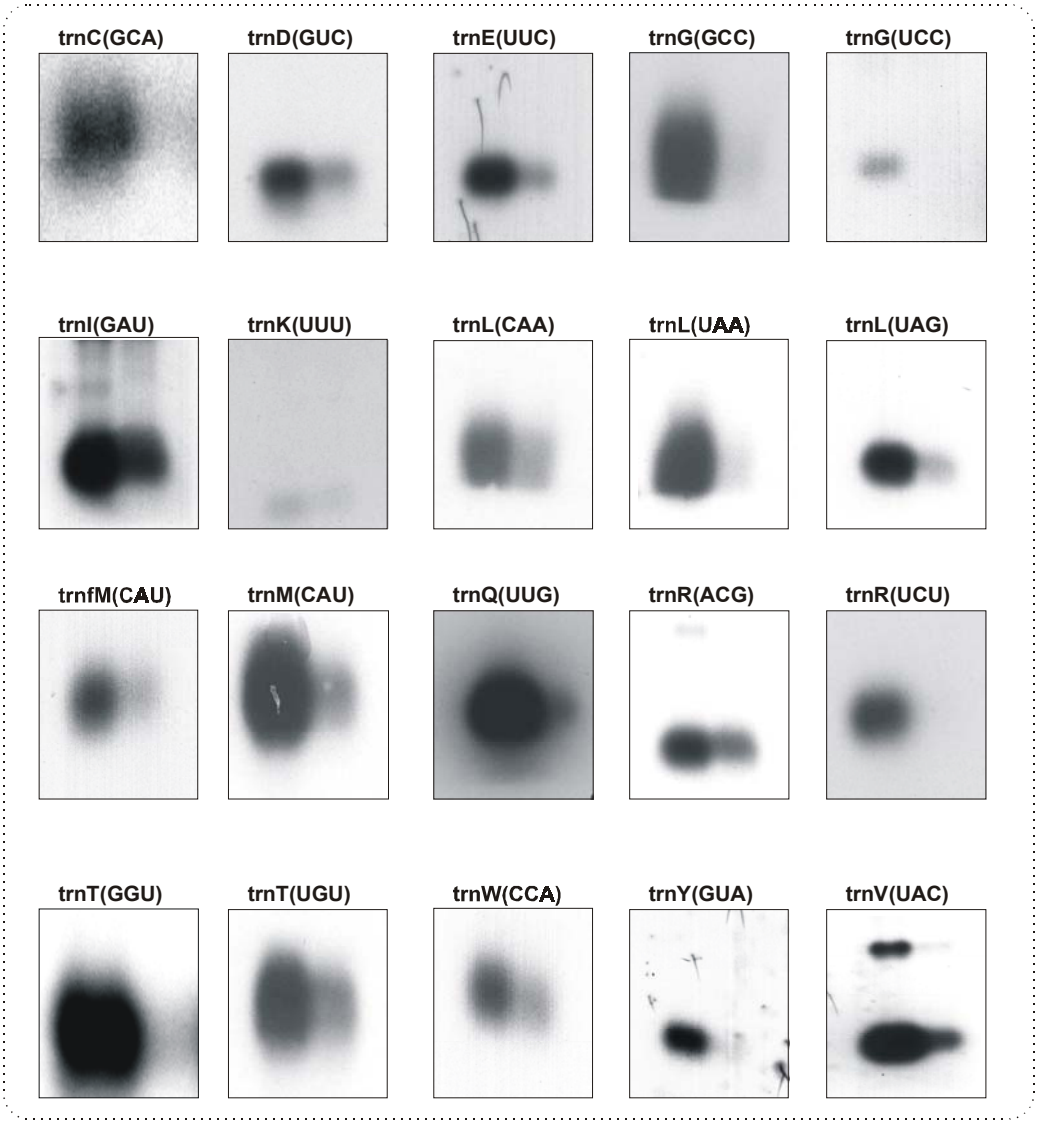


Figure 33. Reduced accumulation of certain tRNAs (group II tRNAs) in PEP-deficient material. Experimental procedures exactly as given in Figure 32.

Table 3. Comparison of tRNA accumulation in the ∆*rpoA* mutant with the tRNA content of *E. virginiana*.

Plastid tRNAs	Expression in ∆*rpo* relative to wild-type[1]	*Epifagus*
trnH(GUG) his	**++**	intact
trnK(UUU) lys	**o**	missing
trnQ(UUG) gln	**o**	intact
trnS(GCU) ser	**++**	intact
trnG(UCC) gly	**-**	missing
trnR(UCU) arg	**-**	pseudo
trnC(GCA) cys	**o**	pseudo
trnD(GUC) asp	**+**	intact
trnY(GUA) tyr	**-**	intact
trnE(UUC) glu	**o**	intact

trnT(GGU) thr	**o**	missing
trnS(UGA) ser	**++**	intact
trnG(GCC) gly	**-**	missing
trnfM(CAU) ile	**o**	intact
trnS(GGA) ser	**++**	pseudo
trnT(UGU) thr	**+**	missing
trnL(UAA) leu	**-**	missing
trnF(GAA) phe	**++**	intact
trnV(UAC) val	**+**	missing
trnM(CAU) met	**o**	intact
trnW(CCA) trp	**o**	intact
trnP(UGG) pro	**++**	intact
trnI(CAU) ile	**++**	intact
trnL(CAA) leu	**+**	intact
trnV(GAC) val	**++**	missing
trnI(GAU) ile	**+**	pseudo
trnA(UGC) ala	**++**	pseudo
trnR(ACG) arg	**+**	intact
trnN(GUU) asn	**++**	intact
trnL(UAG) leu	**o**	intact

[1] The relative abundance of tRNAs in Δ*rpoA* plants was measured by evaluating the signals in Fig. 32 and 33 with a phosphoimager. ++ = 80 - 100% of wild-type amounts; + = 20 - 80% of wild-type amounts; o = 0 - 20% of wild-type amounts; - = not detectable.

4.5.2. Expression of genes coding for ribosomal subunits in tobacco wild-type and PEP-deficient cells

As mentioned above, it has been speculated that genes of the genetic apparatus are transcribed predominantely by NEP, whereas photosynthesis-related genes are transcribed by PEP (Hajdukiewicz et al., 1997). As demonstrated for tRNA genes, this premise does not hold. To complete the survey of transcript accumulation of genes of the genetic apparatus, genes coding for ribosomal subunits and rRNA genes were analysed as well.

In general, it was assumed that mRNAs for ribosomal proteins accumulate to higher amounts in Δ*rpo*-tissue. This is in particular striking for *rpl14*, *rpl16*, *rpl32*, *rps7* and *rps18* (Fig. 34). Transcript abundance, like for *rpl22*, *rps8* and *rps16*, was low in both wild-type and Δ*rpo*-material. In contrast to these quantitative differences, qualitative differences are rare. For *rps7*, a large transcript of 5.4 kb seems to be absent in wild-type. This could be a precursor RNA which is less rapidly processed in the albino mutant than in wild-type. Also, the smallest transcript of *rps7* at 0.8 kb is barely detectable in the mutant. Possibly, some processing steps are impaired to generate

this transcript, or alternatively, this presumably monocistronic form is degraded in the mutant. Another qualitative difference was observed for *rps3* transcripts. Here, an RNA of 0.15 kb (arrow) is prominent in the mutant, but absent from the wild-type. Again, differences in processing can account for this. Finally, the *rps18* Northern exhibits several differences, both in the high- and low-molecular weight range, leading to a transcript pattern much more complex than in wild-type (arrows). In addition to the wild-type dicistronic transcript of 0.5 kb (*rpl33/rps18*) and the tricistronic *psaJ/rpl33/rps18*-transcript of 1.6 kb, five additional transcripts of 7.5, 4.6, 2.7, 2.0 and 0.4 kb are observed in PEP-deficient material (Fig. 34). Of these, the transcript of 2.7 kb most likely includes sequences of the *petL/petG* dicistron, and the 2.0 kb band would correspond to an intermediate between this pentacistronic and the typical wild-type tricistronic transcripts. The longer transcripts of 7.5 and 4.6 kb are not detected with probes corresponding to the upstream genes *petA* and *ORF99* (Fig. 18 and data not shown) so that these transcripts are apparently read-through products that do not terminate (or are not processed) at the regular *rps18* 3' regulatory sequence. It has to be assumed that these transcripts extend even beyond *rpl20* and *clpP*, which are on the opposite strand, and hence constitute antisense RNA to these genes. Interestingly, the 0.4 kb transcript found in ∆*rpo* material could correspond to a monocistronic form of *rps18* not observed in wild-type. Alternatively, this transcript is a non-functional degradation product of the dicistronic form that does accumulate in the mutant, but is rapidly degraded in wild-type. Both, kinetic reasons (faster degradation/slower processing) and differences in the processing machineries of mutant and wild-type plants could potentially account for these discrepancies.

An exception to the general rule that mRNAs for ribosomal subunits accumulate to higher amounts in ∆*rpo* material is the *rps14* gene, which codes for the CS14 protein as part of the small subunit of the ribosome. In mustard chloroplasts, this gene is transcribed by NEP from a typical NEP type II promoter (Ruf et al, 1993, Summer et al., 1998). As shown in Section 4.3.1.3. (Fig. 17) *rps14* is cotranscribed with *psaA* and *psaB* forming a tricistronic message in wild-type, which is not present in PEP-deficient material. In addition, in wild-type a monocistronic message of 0.5 kb is detectable (Fig. 17; Meng et al., 1988). Both, the dicistronic as well as the monocistronic form are not detectable in the mutant. Apparently, this ribosomal gene is transcribed by PEP alone.

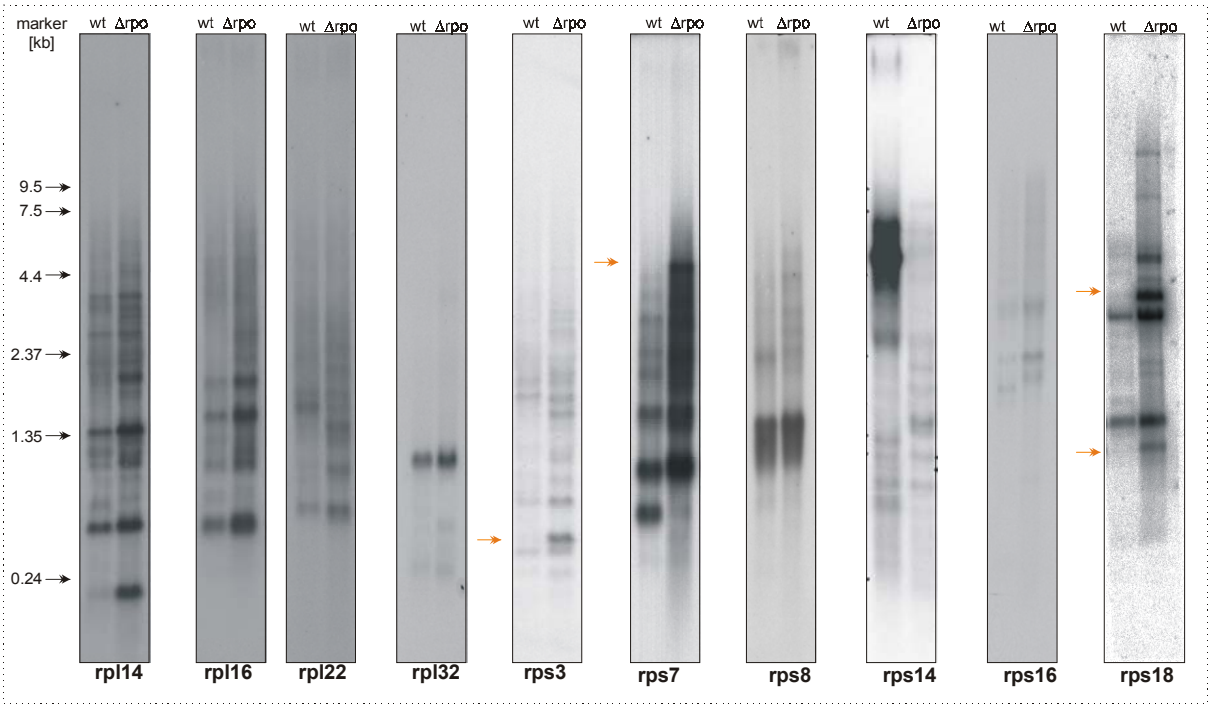


Figure 34. Northern analysis of transcripts coding for ribosomal subunits. 8 µg of total cellular RNA from wild-type and Δ*rpoA* leaves were fractionated on a 1% agarose gel, blotted onto nylon membranes and hybridised with single-stranded antisense RNA probes. Arrows mark transcripts present in Δ*rpo* tissue, but absent in wild-type.

4.5.3. Expression of the rRNA operon

The four plastid ribosomal genes *rrn16*, *rrn23*, *rrn4.5* and *rrn5* together with two tRNA genes are organized in an operon that is repeated in the IR_A and IR_B region of the plastid chromosome. The operon is highly expressed in wild-type and its promoter structure has been studied in detail, by capping experiments it has been shown that in tobacco two major promoters are active, the consensus type PEP-dependent promoter P1 (Nt Prrn-116; Vera and Sugiura, 1995), and a NEP-type promoter P2 (Nt Prrn16-64; Shylavi et al., 1998). Northern analysis has been performed for all 6 genes and led to some unexpected observations (Fig. 35). First, steady transcript amounts of *rrn16* and *rrn23* are clearly reduced in Δ*rpo* material. Transcripts of *rrn4.5* and *rrn5* are reduced as well, though to a lesser extent. At first glance, this would suggest that PEP activity contributes substantial amounts of transcripts to rRNAs. However, transcription rates of the four *rrn* genes are identical to wild-type (Figs. 14 B and 15 A). This means that not the transcriptional activity, but the stability of the produced rRNAs is reduced in the mutants, leading to lower steady-state levels, even though transcription rates are normal. This is in accordance with the reduced number

of ribosomes, stabilising rRNAs, in PEP lacking mutants (DeSantis-Maciosseck et al., 1999).

Other differences may lie in the processing of the precursors. Indeed, some smaller transcripts are not only reduced in the mutant, but almost not detectable (arrows in the *rrn16* and *rrn23* Northern), which indicates that processing is impaired in ∆*rpo* material, too. This is further corroborated by the fact that the precursor of *rrn4.5* accumulates normally in the mutant, whereas the mature form at 0.1 kb is approximately tenfold reduced. At least for *rrn*23 transcripts qualitative and quantitative differences in wild-type and mutants can be explained by the fact that the fraction of *rrn* molecules integrated into ribosomal structure breaks at the definied positions (“hidden breaks”) resulting in a population of low molecular fragments. The reduced number of ribosomes in plastids lacking PEP, therefore, also has an influence on transcript patterns as observed in Northern experiments (De Santis-Maciosseck et al., 1999). A reduction of the mature form in ∆*rpo* material is also observed for *trnI(GAU)*, together with the absence of intermediate transcripts (arrows). Comparably, the *trnA(UGC)* precursor is absent in the mutant as well, although the mature form is present in wild-type amounts. As these two tRNA genes harbour introns, it can be speculated that splicing of *trnI(GAU)* is less efficient in the mutant in comparison to wild-type and also in comparison to that of *trnA(UGC)*. In any case, the amounts of these two tRNAs in the mutants relative to wild-type exceed those of the surrounding rRNAs, which indicates that either the tRNAs have independent promoters or that these tRNAs are more stable/more efficiently processed from the precursors in the mutant than the rRNAs. All in all, transcript accumulation of the *rrn*-operon is reduced in PEP-deficient material, most likely because transcripts are less stable and not because transcription rates are reduced.

To sum up, results obtained for transcript accumulation of genes of the genetic apparatus: There are tRNAs, mRNAs for ribosomal proteins as well as *rRNAs* that show a reduction in steady-state transcript amounts in ∆*rpo* mutants. This is not in agreement with previous models, claiming that these genes are all NEP-dependent and therefore should show a compensatory higher accumulation in ∆*rpo* material than in wild-type.

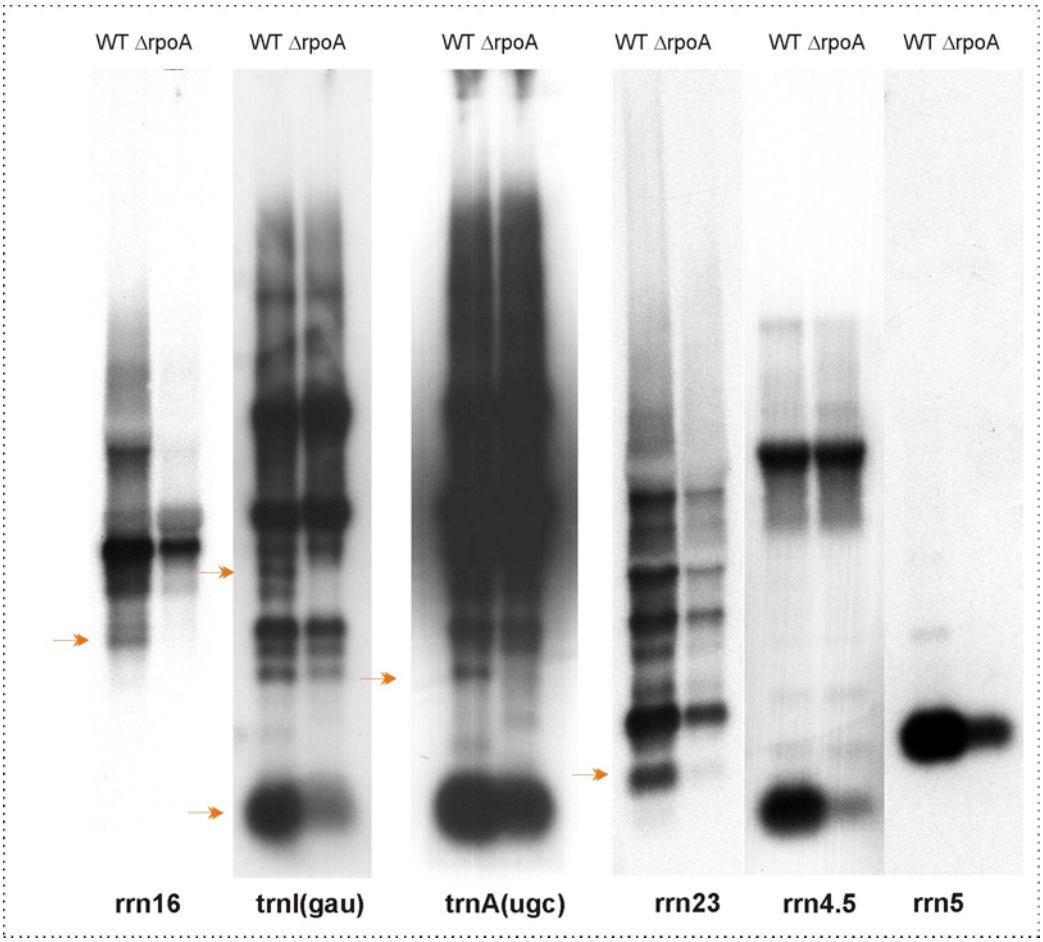


Figure 35. Northern analysis of genes coding for ribosomal RNAs. 3 µg of total cellular RNA from wild-type and Δ*rpoA* leaves were fractionated on a 1% agarose gel, blotted to nylon membranes and hybridised with single-stranded antisense RNA probes. Arrows denote differences between Δ*rpo* and wild-type tissue.

4.6. Analysis of the biogenesis of the cytochrome b_6/f complex

Oxygenic photosynthesis is an energy-transducing process which converts light energy into biochemical energy and is mediated by five major multi-protein complexes in the thylakoid membrane, the photosystems I and II, the ATPase and NADH-complexes, and the cytochrome b_6/f complex (cyt b_6/f). This last mentioned complex transfers electrons from reduced plastoquinone to plastocyanin, a soluble electron carrier located in the thylakoid lumen. Both, the nuclear and plastid genome contribute subunits to this complex, as is typical for complexes of the thylakoid membrane.

Mutants of genes coding for subunits of the cyt b_6/f complex are characterized by their light sensitivity and their *hcf* (high fluorescence) phenotype (Barkan et al., 1998, Meurer et al., 1999). At present, studies on *petA*, *petB* and *petD* of *Chlamydomonas reinhardtii* and tobacco have shown that cytochrome *f* has a prominent role in

complex assembly, as it recruits the Rieske protein (*petC*) and forms a precomplex which acts as an anchor for further assembly (Kuras et al., 1994). However, a comprehensive approach to unravel the temporal and spatial assembly of this complex is still missing.

4.6.1. Cloning strategy and homoplastomicity check of transformed pet-deficient lines

To elucidate the assembly of the cytochrome b_6/f complex, a transgenic approach was chosen. Applying biolistic transformation, knock-out mutants of plastid-encoded *petA* (coding for cytochrome *f*), *petB* (cytochrome b_6; cyt b_6) , *petD* (subunit IV), *petG* (subunit V), *petL* (3.5 kDa subunit) and *petN* (subunit VIII) were prepared. This was accomplished by introducing pBlueScript KS II$^+$-based plasmids that carried the coding region of the *pet* genes either disrupted or in some cases completely replaced by a selectable marker cassette containing the bacterial *aad*A gene, which confers resistance to spectinomycin and streptomycin into tobacco leaf chloroplasts.

As depicted in Fig. 36, the *petA* mutant was made by inserting the *aad*A cassette into a *NspV* site, 251 bp downstream of the *petA* start codon. In case of *petB*, the *aadA* cassette was inserted between a *ClaI* and a *SnaBI*-site (positions 77.860 and 78.823 of the tobacco plastid chromosome, Acc. Z00444), thereby deleting a 963 bp fragment representing almost the entire exon 2 of *petB* and 360 bp of the intron (Fig. 37). The *petD* knock-out was achieved by inserting the cassette into the reading frame using a singular *Bpu1102I* site, which is located 900 bp downstream of the start codon inside the intron (Fig. 38). The *petG* mutant was obtained by replacing a *Bsu15I*/*PacI* fragment with the *aadA* cassette (Fig. 39). This removes the entire *petG* gene. For *petL*, both an insertional approach as well as a complete removal of the reading frame was undertaken (Fig. 40). In construct 1, the *aadA* cassette was inserted into a *Bpa10I* site, 51 bp downstream of the *petL* start codon. As it appears possible that a truncated version of PetL is translated from this construct that may still maintain, at least in part, the function of the full-length protein, a deletion construct was prepared as well. In construct 2, a *NdeI* site in conjunction with a *SphI* site was used to remove the complete *petL* coding region (Fig. 40). Finally, *petN* was also removed completely using two *MunI* sites flanking the gene, 277 bp upstream of the start codon and 43 bp downstream of the stop codon (Fig. 41).

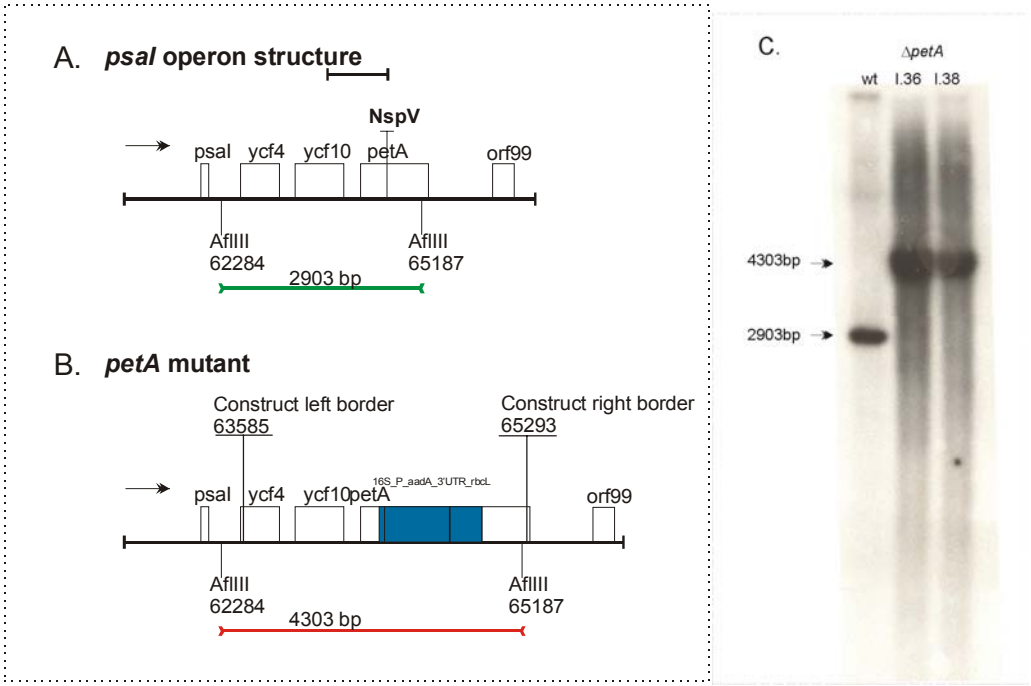


Figure 36. Schematic presentation of the *psaI* operon without (A) and with *aadA* cassette integrated in *petA*. (B). Boxes denote genes. The arrow indicates the direction of transcription. The bar marks the probe (oligonucleotides used for PCR: cemAfor2 and petArev2) used for Southern analysis. Restriction sites used for inserting the *aadA* cassette are boldfaced (NspV). Numbers refer to positions on the tobacco plastid chromosome, deposited as entry Z00044 in genbank. (C) Southern analysis of wild-type and Δ*petA* plants. 200 ng of total leaf DNA of the two mutant lines I.36 and I.38 and of wild-type were digested with *AflIII*, fractionated in a 0.8% agarose gel and blotted onto nylon membranes. Hybridisation was carried out with a radiolabeled, single-stranded RNA probe generated by *in vitro* transcription using a PCR product as template. The position of the probe is marked by a black bar in (A). The size of detected signals are marked in green for wild-type and in red for mutant.

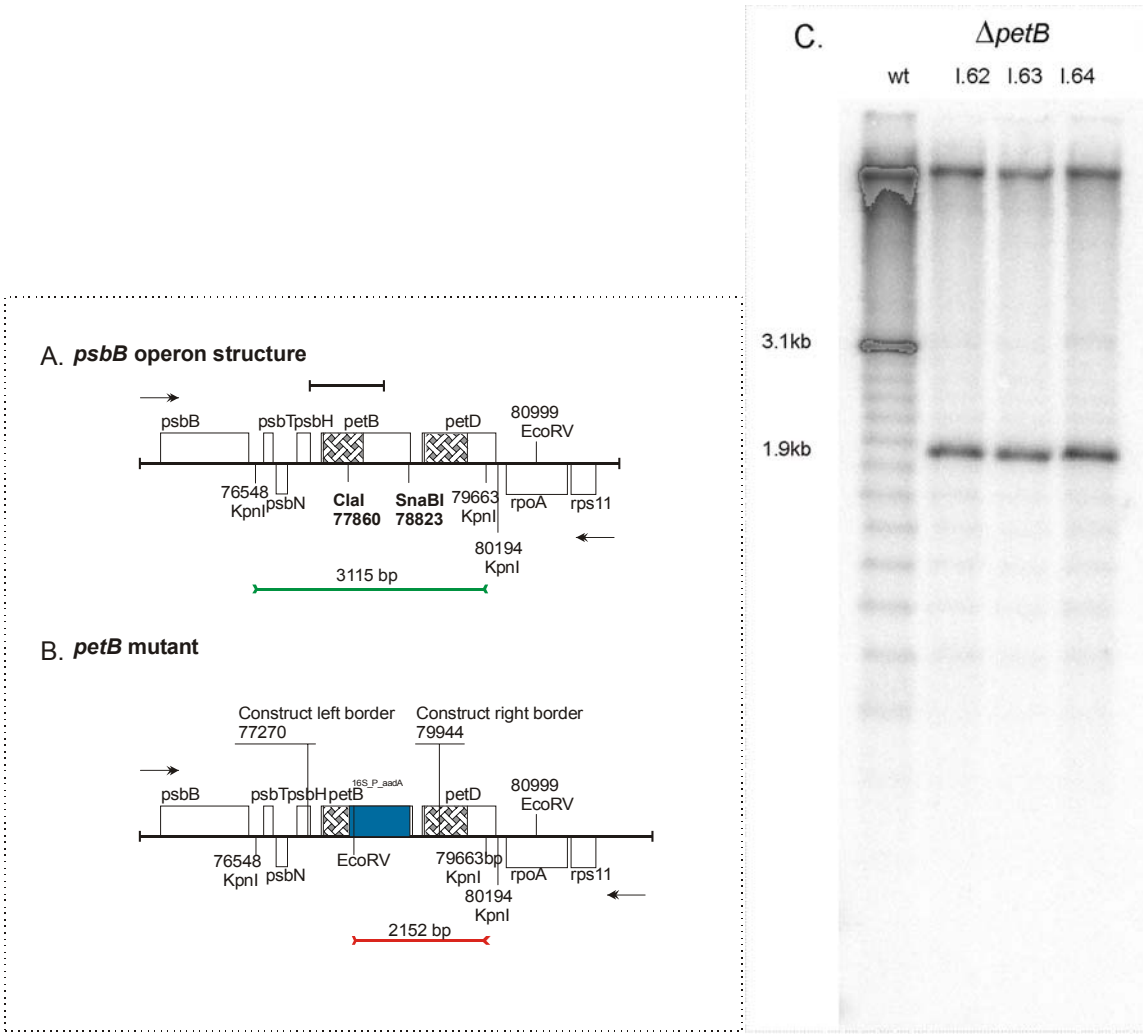


Figure 37. Schematic presentation of the *psbB* operon without (A) and with *aadA* cassette integrated in *petB*. (B). Legend as in Fig. 36. Chequered boxes = introns. (C) Southern-analysis of wild-type and Δ*petB* plants. Legend as in Fig. 35, but the DNA was digested with *KpnI* and *EcoRV*. Hybridisation was carried out with a radiolabeled probe generated by random labelling using a PCR product as template (oligonucleotide used for probe: psbH3'for+petDrev). The position of the probe is marked by a black bar in (A). The sizes of detected signals are marked in green for wild-type and in red for mutant.

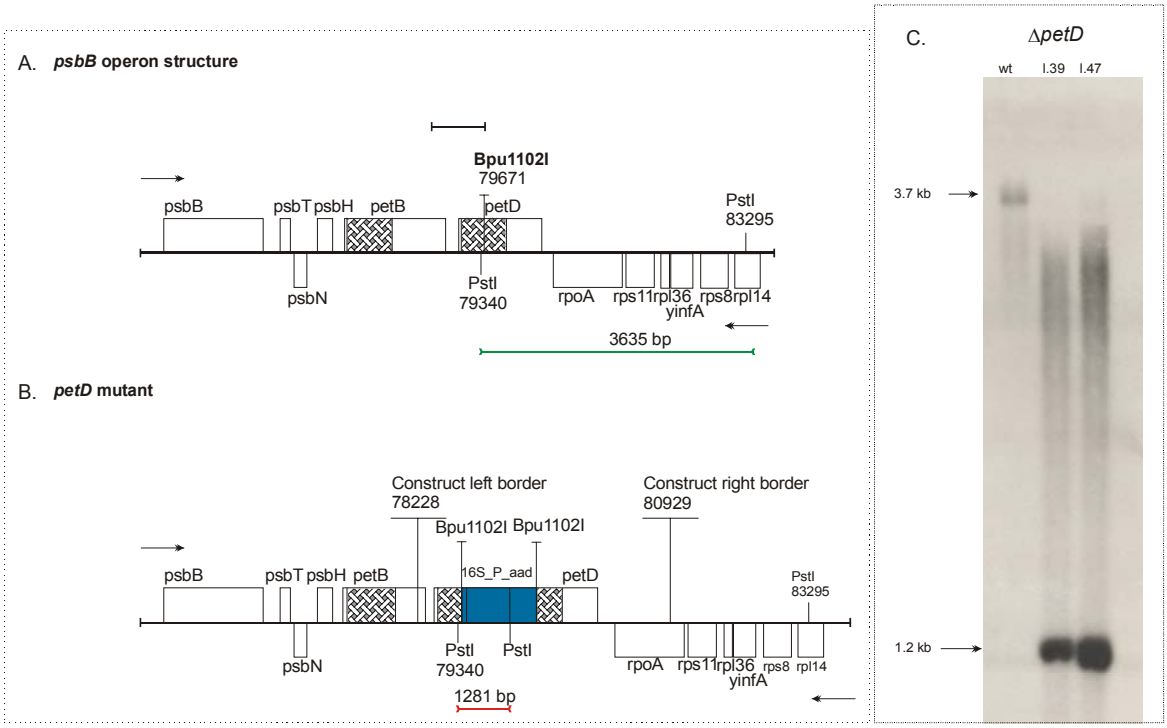


Figure 38. Schematic presentation of the *psbB* operon without (A) and with *aadA* cassette integrated in *petD*. (B). Legend as in Fig. 36. (C) Southern analysis as in Fig. 35, but the DNA was digested with *PstI*. Hybridisation was carried out as for *petB* (Fig. 36), but with probe amplified with oligonucleotides petB3'for and petDrev. The position of the probe is marked as a black bar in (A). The sizes of detected signals are marked in green for wild-type and in red for mutant.

For every construct, more than 5 lines have been obtained that carried the insertion as verified by PCR analysis (data not shown). Of these, 1 - 3 lines have been maintained in tissue culture for several cycles of regeneration in order to allow segregation of the mutation under selective pressure (500 µg/ml spectinomycin). After every regeneration cycle, the relative amounts of wild-type to mutant plastome was tested by Southern analysis. Results are shown adjacent to the schematic were presentation of each construct in Figs. 35 - 40. In all cases, mutants after the third regeneration cycle have been analysed. As can be seen for individual mutant lines of *petA*, *petB*, *petD, petL* constructs 1 and 2; *petN* and *petG*, no signal could be detected that would correspond to the wild-type restriction fragment. This indicates that the plants are homoplastomic with regard to the mutation. Only one line of Δ*petN*, l.63 exhibited an aberrant signal pattern neither corresponding to the wild-type nor to the mutant pattern. Thus, this mutant did not represent the desired knock-out and was discarded.

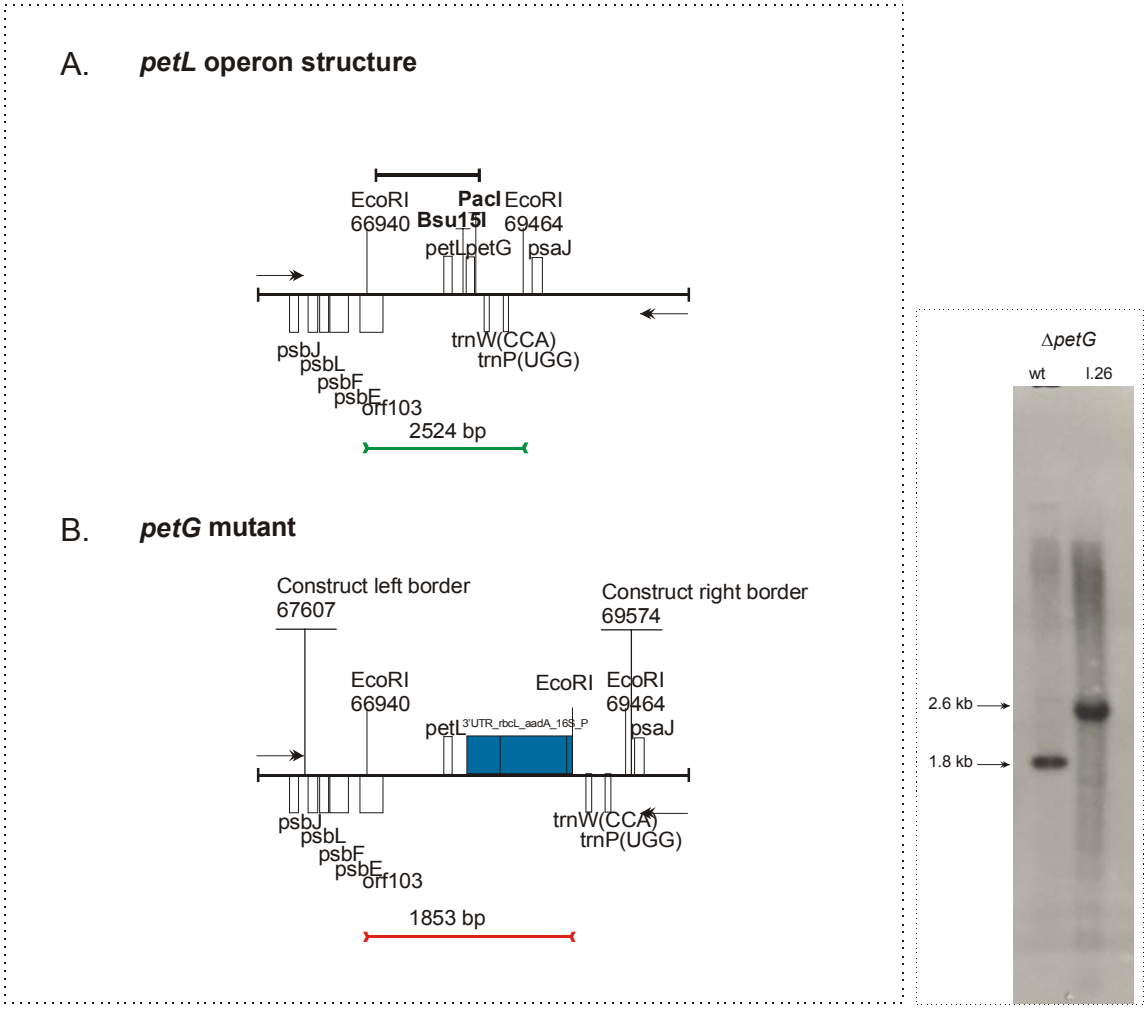


Figure 39: Schematic presentation of the *petL/petG* operon without (A) and with *aadA* cassette the integrated in *petG*. (B). Legend as in Fig. 36. (C) Southern analysis as in Fig. 35, but the DNA was digested with *EcoRI*. Hybridisation was carried out as described in Fig. 36, but the probe was amplified using oligonucleotides orf103for2 and AT7petG. The sizes of detected signals are marked in green for wild-type and in red for mutant.

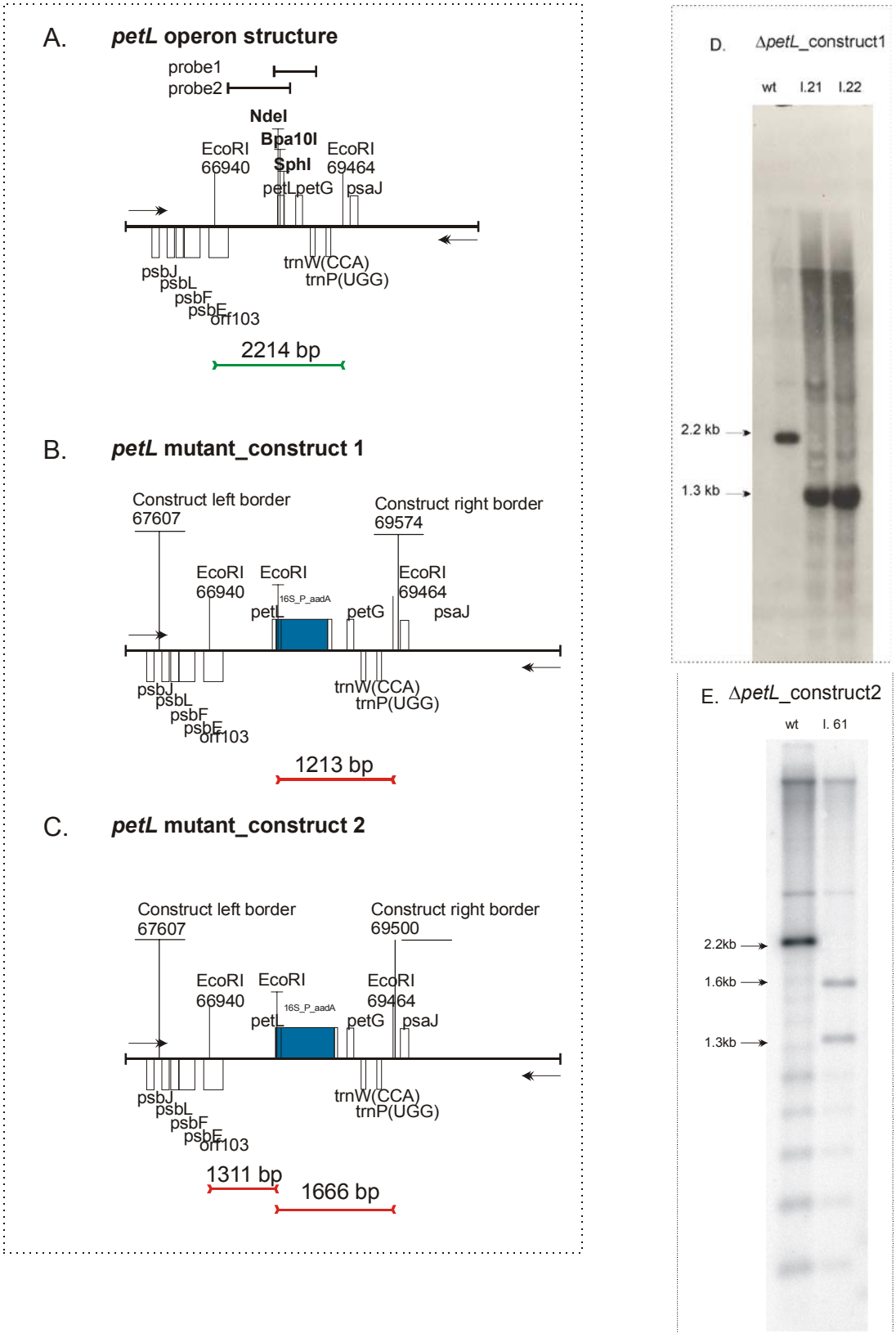


Figure 40. Schematic presentation of the *petL/petG* operon without (A) and with *aadA* cassette integrated at two different positions (B and C). In (B), the cassette has been inserted in the *petL* reading frame (*Bpa10I*), whereas in (C), the entire gene was deleted. Legend as in Fig. 36. (D and E) Southern analysis as described in Fig. 36, but the DNA was digested with *EcoRI*. Hybridisation was carried out as described in Fig. 37, but probe 1 was amplified using oligonucleotides ApetL and AT7petG, and in the case for probe 2 orf103for2 and AT7petL. The sizes of detected signals are marked in green for wild-type and in red for mutant.

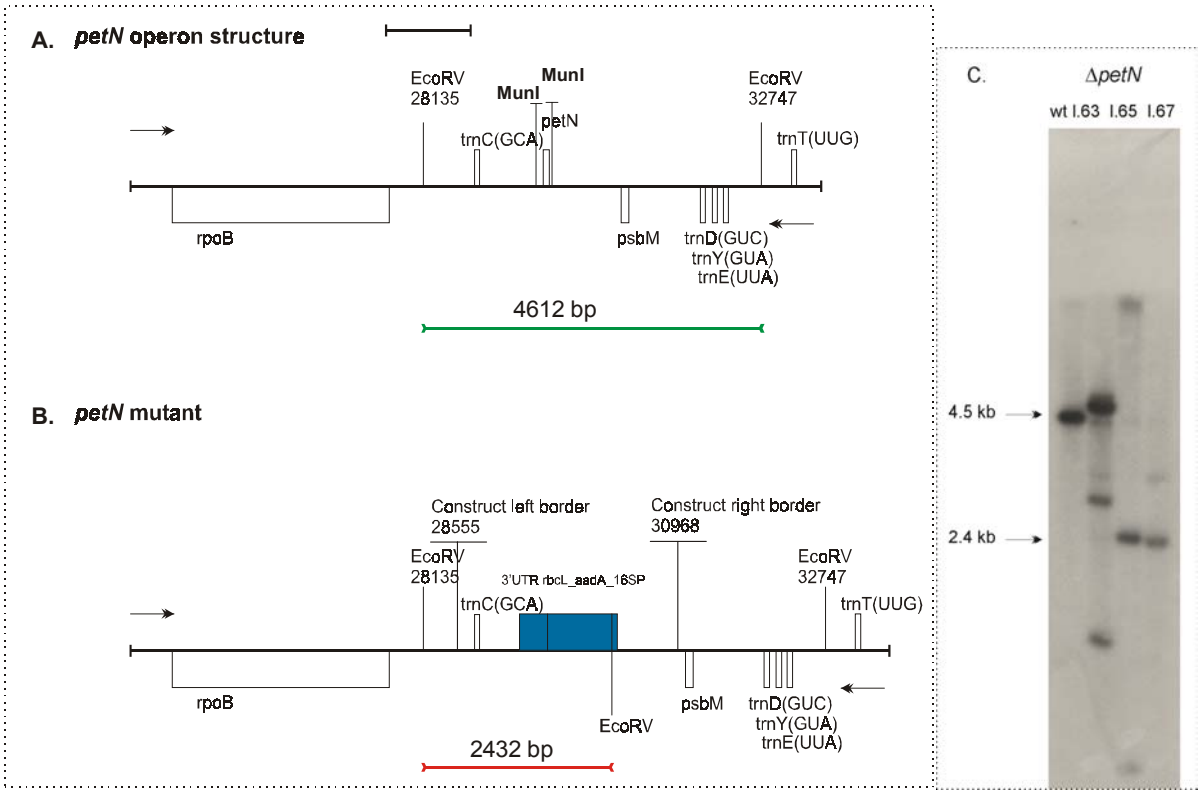


Figure 41. Schematic presentation of the *petN* coding region without (A) and with integrated *aadA* cassette (B). Legend as in Fig. 36. (C) Southern analysis as described in Fig. 36, but the DNA was digested with *EcoRV*. Hybridisation was carried out as described in Fig. 36, but the probe was amplified using oligonucleotides cpoBleft and AtrnCgca.

4.6.2. Phenotype characterisation of the transplastomic mutant lines

With the exception of Δ*petL* plants, all homoplastomic mutants show some photobleaching, that is, under normal tissue culture conditions in medium light (ML; 100 µE/m^2s^1) they bleach and some become eventually white like Δ*petA*, Δ*petD* and Δ*petN*, others retain a pale green colour like Δ*petB* and Δ*petG*. Mutants remained green, when kept under very low fluence rates, low light conditions (LL; 4 µE/m^2s^1; Fig. 42 A and B). All mutants, except of Δ*petL*, show growth retardation and slightly aberrant leaf formation in comparison to wild-type (Figs. 42 a and 42 b, and data not shown). When grown on soil, only Δ*petL* plants were viable, whereas the other mutants died within a few days after they were planted out onto the earth. This shows that, aside of Δ*petL*, all mutants have lost the ability for autotrophic growth. Interestingly, the *petL* knock-out described in *Chlamydomonas rheinhardtii* is not capable of growing without external carbon source, which suggests that the two homologous genes serve different functions in the two organisms.

There are additional clear differences in phenotypes between mutants. For instance, Δ*petN* plants are only slightly green even under low light, whereas others are almost coloured like wild-type (Δ*petB*). This observation indicates that each mutant has a different impact on chloroplast biogenesis.

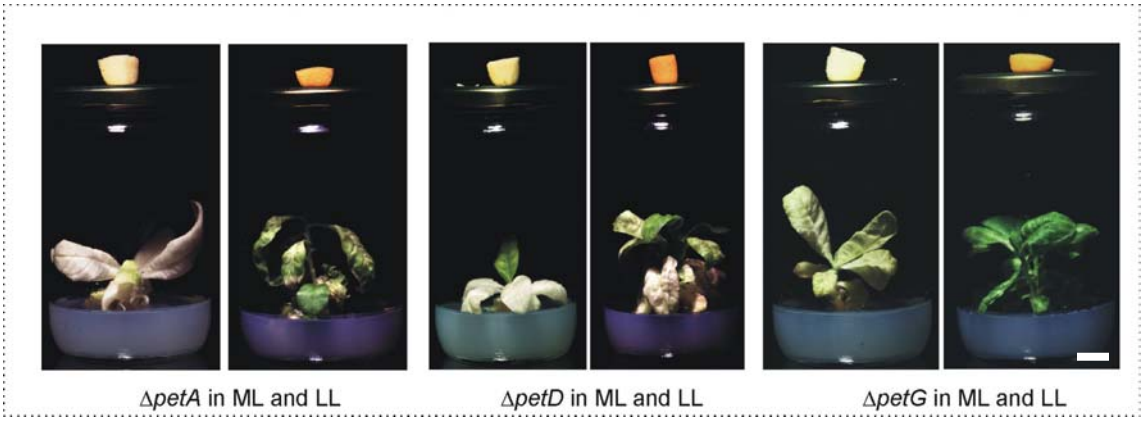


Figure 42a. The Δ*petA*, Δ*petD* and Δ*petG* mutant phenotype differences when growth of 2 weeks-old plant material was under two different light intensities: low light fluence LL (4 – 10 μE/m^2s^1) and normal light fluence ML (100 μE/m^2s^1). Size bar in white represents 2cm.

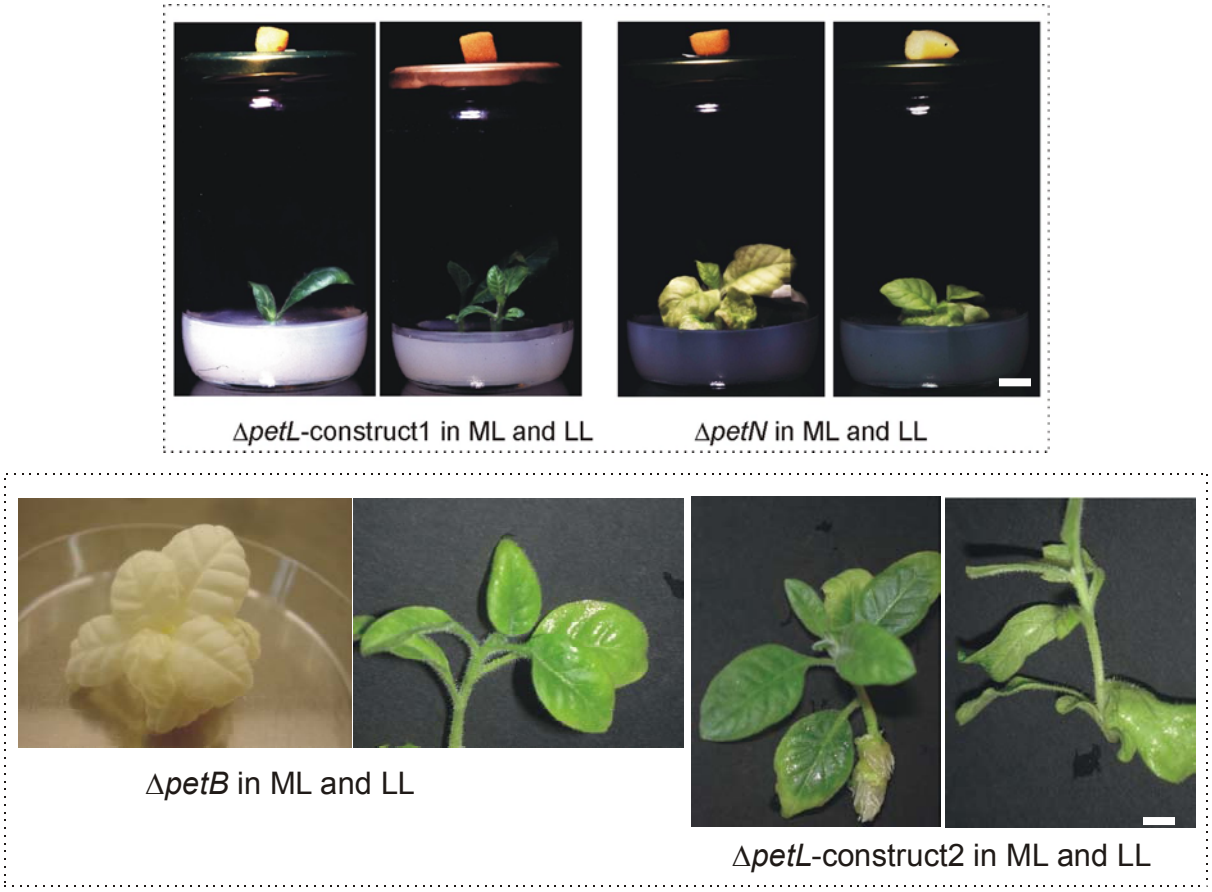


Figure 42b. The Δ*petL* construct1, and Δ*petL* construct2, Δ*petB* and Δ*petN* mutant phenotype differences when growth of 2 weeks-old plant material was under two different light intensities: low light fluence LL (4 – 10 μE/m^2s^1) and normal light fluence ML (100 μE/m^2s^1). Size bar in white represents 2cm.

4.7. Protein analysis of the transplastomic lines

4.7.1. Complex formation

To investigate, whether other complexes of thylakoid membranes of mutants are altered in comparison to wild-type, two-dimensional gel electrophoresis was applied. To this end, thylakoid proteins were prepared from 2 week-old wild-type, Δ*petA* and Δ*petL* plants that have been grown under low-light fluences (4 – 10 $\mu E/m^2s^1$). These protein preparations were fractionated according to their isolectrical point (in a pH gradient from 4.0 to 7.0) in the first dimension and in an SDS-PAGE in the second dimension (see also Materials and Methods). The protein patterns of wild-type and Δ*petL* were almost identical which is in accordance with the fluorescence measurements and the phenotype (Fig. 42 and data not shown). Only a slight reduction in subunits of the cytochrome b_6/f complex is observed (circled in Fig. 43; as tested immunologically by Alexander Hepp, Diploma thesis, Munich, 2001). Apparently, the loss of *petL* does not impose severe problems on thylakoid biogenesis. In contrast to this, the protein profile of the Δ*petA* mutant deviates strongly from the wild-type situation. Most conspicuously, the spots corresponding to the plastid-encoded core components of the cytochrome b_6/f complex are clearly reduced (circled in Fig. 43). In addition, other proteins seem to be reduced as well. However, no qualitative differences were noted, that is, all proteins found in wild-type are also present in the mutant, although at lower abundance. This suggests that the loss of individual subunits of the cytochrome b_6/f complex does have a negative impact on the accumulation of other proteins of the complex, although it is not possible with the current approach to distinguish between primary and secondary effects.

Similar results were obtained when thylakoid membrane complexes were separated by sucrose density gradient centrifugation. Plastids from low-light-grown (4 - 10 $\mu E/m^2s^1$) mutants Δ*petL*, Δ*petG* and wild-type were isolated and solubilized with 0.06% dodecyl-maltoside. The resulting crude extracts were loaded on a 0.1 to 1 M sucrose gradient and centrifuged for 16.5 h at 187.000x g in an SW40 rotor in an ultracentrifuge (see also Materials and Methods). This allows separation of intact complexes of the thylakoid membrane that are visible due to their pigment content (Swiatek et al., 2001). As shown in Fig. 44, the major complexes observed in wild-

type, like the LHCII antenna monomers and trimers and the different forms of PSI and PSII complexes, are also seen in the mutants Δ*petL* and Δ*petG*. The presence of the major complexes and their higher molecular weight aggregates, like antenna-PS-associations in the mutants shows that the disruption of a *pet* gene does not affect the biogenesis of the other thylakoid supramolecular complexes, which is in accordance with the data from 2-dimensional gel electrophoresis.

4.7.2. Immunological characterisation of transplastomic lines

Although two-dimensional gel electrophoresis gives a good overview of protein accumulation in the mutant material, this does not apply to the cytochrome b_6/f complex. Thus, this complex and the direct influence of the loss of a subunit on other proteins of the same or different multiprotein complexes can be assessed only and more precisely immunologically by Western analysis. This was done for all mutants and wild-type with antibodies against five subunits of the cytochrome b_6/f complex, four of them plastid- and one nuclear-encoded (Rieske protein). In addition, PsaC, a subunit of PS I, the D1 protein of PSII, PsbE of cytochrome b_{559} and the α-subunit of the ATP synthase were analysed as well. The results are presented in Fig. 45.

As expected, cytochrome *f*, cytochrome b_6 and subunit IV are missing in the corresponding knock-out mutants, which corroborates that these mutants are homoplastomic. Of the seven mutants investigated, three, Δ*petB*, Δ*petD* and Δ*petG,* show the most pronounced reduction of other subunits of the cytochrome b_6/f complex. In Δ*petG*, except PetL no other subunit could be detected, whereas in Δ*petB* and Δ*petD*, small amounts of cytochrome *f* and also of the Rieske protein are present. Apparently, without either of these proteins, the stability of other subunits of the complex is impaired, perhaps because assembly does not proceed. This indicates that PetB, PetD and PetG participate in early steps of complex formation, consistent with earlier data in *Chlamydomonas* (Monde et al, 2000, Kuras et al., 1994). In addition, PetG also seems to participate in assembling this premature complex. Two further mutants, Δ*petA* and Δ*petN*, also show a pronounced reduction in the investigated subunits. In Δ*petA* plants, PetD and PetC are present at approximately 5% of wild-type levels. In contrast to this, PetB appears to be absent.

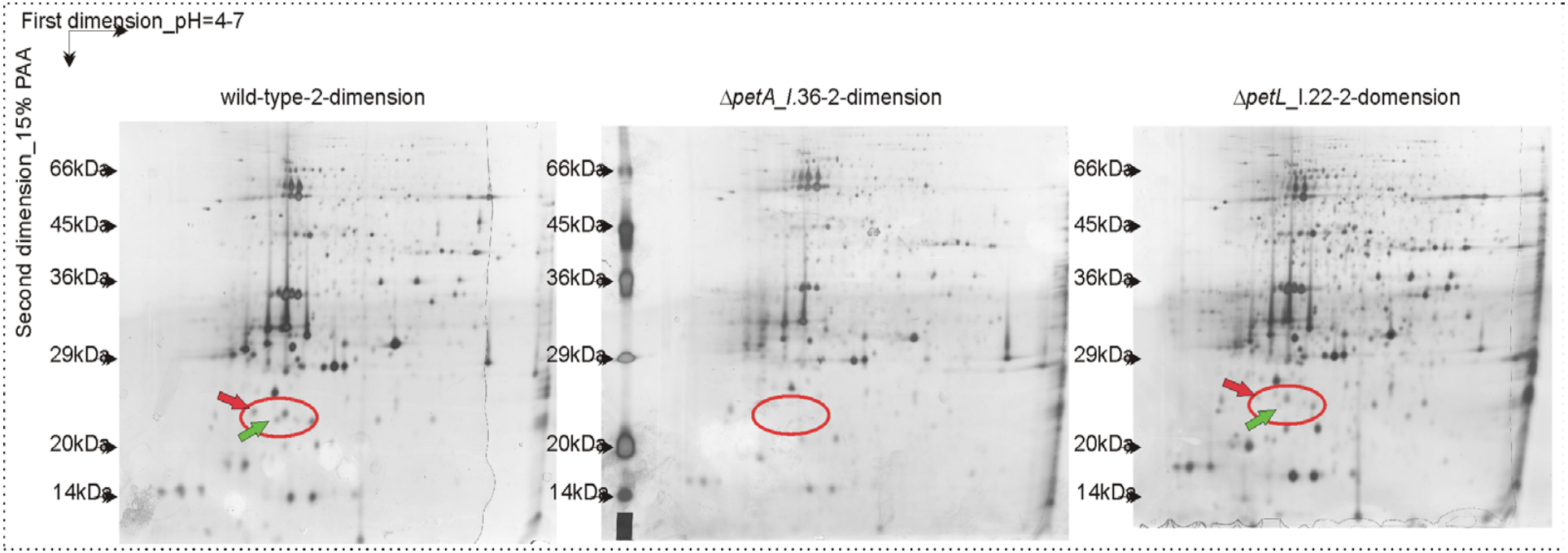


Figure 43. Analysis of Δ*pet* mutants in comparison to wild-type by two-dimensional gel-electrophoresis. Thylakoid membranes were prepared from wild-type, Δ*petA* and Δ*petL* mutant leaves (grown at low light intensities, 4 - 10 µE/m^2s^1). Isolated thylakoid membranes (according to Müller and Eichacker, 1999) were separated in the first dimension according to their isoelectric point (pH gradient from 4.0 to 7.0). The resulting slab gels were applied to an SDS PAGE (second dimension). Proteins were visualized by silver stain. The circle marks core subunits of the *cytb6/f* complex, red arrows show PetC (Rieske protein) and green arrows show accumulated *petD* gene product (subunit IV, according to immunological data by Alexander Hep, Diploma thesis, Munich, 2001). PetC and PetD were not detectable in the mutant Δ*petA*, whereas in Δ*petL* mutant they were found only slightly reduced compared to wild-type amounts.

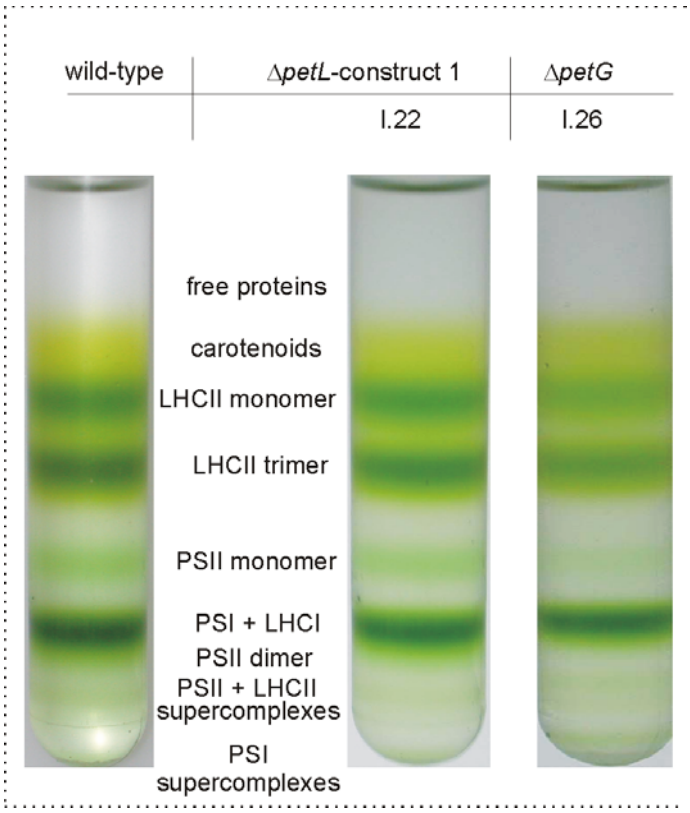


Figure 44. Sucrose density gradient analysis of thylakoid membrane complexes in Δ*petG* and Δ*petL* in comparison to wild-type. An equal number of isolated plastids (3 x 10^8) from wild-type and mutants (Δ*petL* construct 1 l.22 and Δ*petG* l.26) was solubilized with TMK buffer containing 0.06% dodecyl-maltoside. These crude chloroplast fractions were loaded onto a 0.1 to 1.0 M sucrose gradient and centrifuged at 187.000x g for 16.5 h in an rotor SW40. Multiprotein complexes are identified according to their colour, density and composition as described in Swiatek et al. (2001).

However, the PetB antibody is less sensitive than the PetD and PetC antibodies, as no signal in wild-type at a dilution of 5% is visible. Hence, it can not be ruled out that PetB accumulates in low amounts as PetD and PetC in this mutant. The accumulation of PetD, PetC and possibly also PetB in *petA* knock-out plants indicates that PetA is not necessary for the initial integration and accumulation of the core components and may therefore enter the complex later during assembly. In Δ*petN* plants, no PetB and no PetC is detectable. However, substantial amounts of PetA (more than 25% of the wild-type quantity) and low amounts of PetD are present. The signal corresponding to PetD is not an artifact from the adjacent lane, as it has been repeatedly found also in other experiments (data not shown).

Again, the relatively low sensitivity of the PetB antibody does not allow to rule out the presence of small amounts of PetB in this mutant. Evidently, the loss of *petN* does not impair strongly accumulation of *petA* and is also not abolishing, though drastically

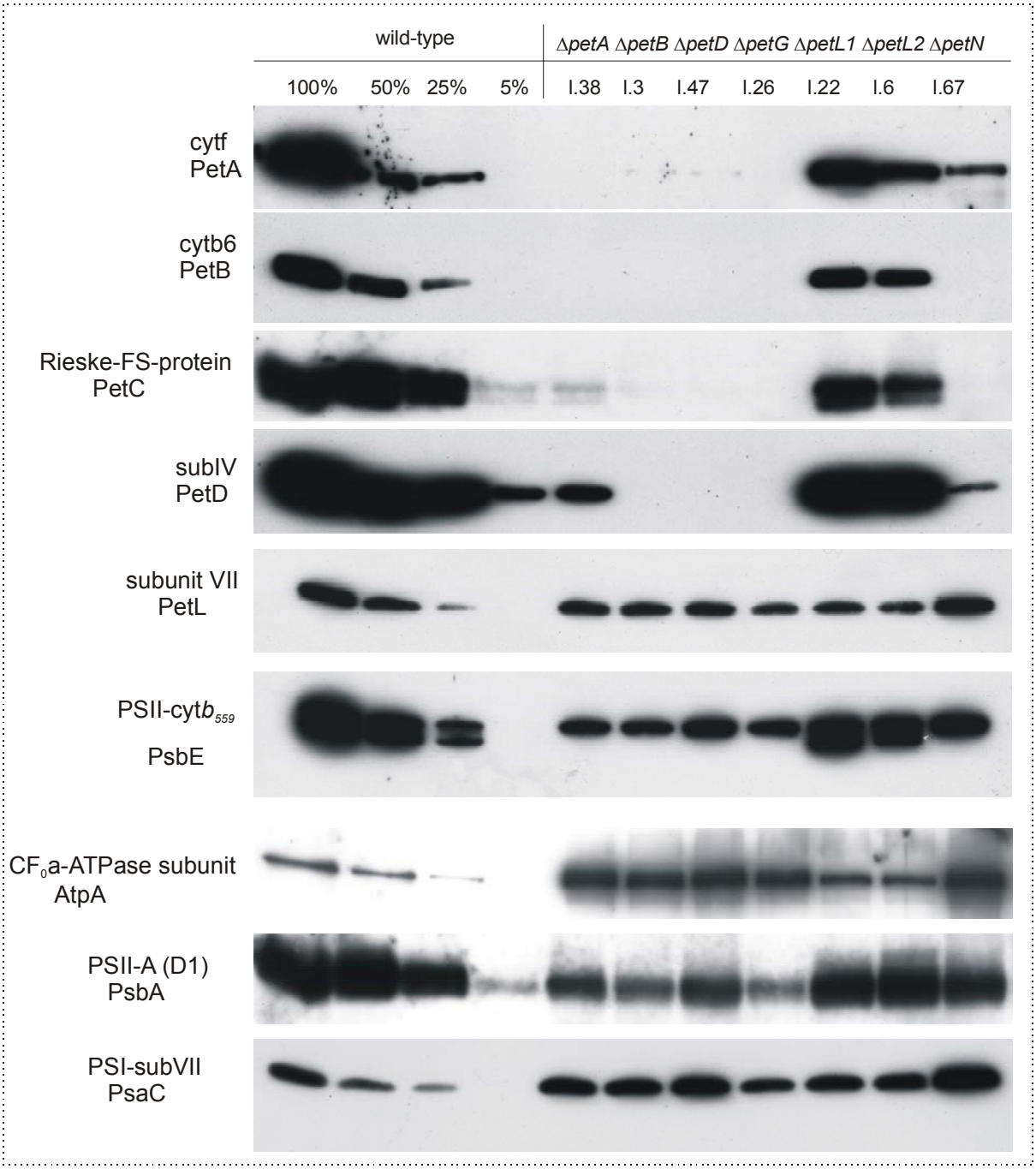


Figure 45. Immunoblot analysis of of thylakoid proteins of *cytb6/f* complex mutants. Equal amounts of proteins from isolated thylakoid membranes from mutant plants grown for two weeks at low light intensities (4 - 10 µE/m^2s^1) were fractionated by a 15% PAGE. In addition, several dilutions of wild-type thylakoid proteins were loaded in order to estimate the relative amounts of proteins accumulating in the mutants. These gels were blotted to PVDF membranes and incubated with monospecific, polyclonal antisera directed against subunits of thylakoid membrane complexes as indicated at the left.

reducing, PetD accumulation. Δ*petL* plants finally accumulate PetA, PetB, PetD and PetC in almost normal amounts (50 - 75%), no matter whether the knock-out is due to simple insertion of the selective marker (Δ*petL*-construct1) or due to deletion of the entire reading frame (Δ*petL*-contruct2). Surprisingly, when αPetL antibodies (kindly

provided by Prof. Rochaix, Geneva) were used for detection of PetL polypeptides, it was possible to detect the normal amounts of subunit VII in all cytochrome *b*$_6$*f* complex mutants, even in Δ*petL* knock-out mutants (see Fig. 52). The level of accumulated protein in Δ*petL* knock-out mutant, either in the insertional (Δ*petL*_construct1) or in the deletion mutant (Δ*petL*_construct2) was not impaired. This can be explained as (i) an effect of still not homoplastomic material. This would mean that Southern analysis was not sensitive enough to verify homoplastomy of this transplastomic material or (ii) the nucleus encodes a copy of this gene which, after import into plastids, is integrated into thylakoid membranes. The detected accumulation of all subunits of the cytochrome *b*$_6$*f* complex in Δ*petL* knock-outs reveal conclusion that at least the plastid-encoded PetL is not essential for stabilizing the other subunits of cytochrome *b*$_6$*f* complex.

Taken together these data suggest that the loss of PetB, PetD and PetG has a large negative impact on the accumulation of the other subunits, consistent with the role of these subunits in early complex formation in *Chlamydomonas* (Monde, et al., 2000, Kuras et al., 1994). The effect of the *petA* and *petN* knock-outs is less pronounced, suggesting that these subunits enter the complex later (see Figs. 43, 44 and 45). The plastid-encoded PetL, finally, does not seem to play a prominent role in complex formation, as all subunits investigated in this mutant accumulate almost to wild-type amounts.

Analysing the accumulation of the D1 protein of PSII and the α-subunit of the ATP synthase revealed some interesting cross-relations between the biogenesis of different thylakoid complexes. First, the α-subunit accumulates to higher levels in all knock-out mutants with the exception of Δ*petL*. This could be due to an overrepresentation of ATP synthase subunits as a consequence of the reduction of cytochrome *b*$_6$*f* subunits in the protein extract. If this were the case, subunits of PS II should be overrepresented as well. However, D1 does not exceed wild-type levels in the mutants. Rather a marked reduction of D1 can be observed, particularly striking for Δ*petG*, but also for Δ*petB* and still weaker for Δ*petD* and Δ*petA*. Only Δ*petL* plants possess approximately normal levels of D1, and the PetN knock-out has substantial amounts of this protein as well. Apparently, the reduction of D1 is a general feature of the knock-outs (with the exception of Δ*petL*). Western analysis with PsbE antisera

show that cytochrome b_{559} which is a constituent of photosystem II as well exhibits a similar pattern. The data do not reveal whether this is an indirect effect or whether there is a direct regulatory relation between the two complexes, mediated for example by PetG, the subunit that has the greatest effect on D1 when lost (see Fig. 45).

5. Discussion

5.1. Conserved *RpoT* genes of both parental lines of tobacco

Six genes encoding bacteriophage-type RNA polymerases were identified in the *N. tabacum* nuclear genome. Based on their homology to *Arabidopsis* RpoT sequences, the genes were classified into three groups designated *NictaRpoT;1*, *NictaRpoT;2* and *NictaRpoT;3*, respectively. Each group consists of two members displaying more than 96% amino acid identity within the coding regions. Comparisons with genomic fragments amplified from the diploid parent species of allotetraploid *Nicotiana tabacum* (Okamura and Goldberg, 1985) showed that the two members of each group are orthologuous genes inherited from *N. tomentosiformis* and *N. sylvestris*, respectively (Figs. 3 and 4). According to their ancestry, the *Nicotiana tabacum* genes were therefore designated as either “tom”- or “syl”-forms (see Figs. 3, 4 and 5).

Interestingly, all six *RpoT* genes of tobacco are expressed: amplification of coding regions from oligo dT-primed cDNA resulted in all three cases in a mixture of tom- and syl-specific transcripts. A functional conservation of genes of both parents following an allopolyploidization event has been described, for example, for nitrate reductase, subunits of photosystem I as well as for auxin-binding proteins (Vaucheret et al., 1989; Obokata et al., 1993; Leblanc et al., 1997) and is thought to be the rule rather than an exception (reviewed in Comai, 2000). In a number of reported cases, however, parental gene copies were reported to have recombined blockwise (Sperisen et al., 1991) or acquired frame-shift mutations (Lee et al., 1998). For *N. tabacum RpoT* genes, neither evidence for recombination among parental forms nor for silencing could be detected. Experiments with natural and synthetic polyploids of *Brassica* (Song et *al*., 1993, 1995) indicated that tetraploid nuclear genomes are more related to the maternal parent having contributed the cytoplasm. This could illustrate constraints of a functional nuclear-organellar communication on a successful alloploidization event.

In *N. tabacum*, the cytoplasm is a derivative of the *N. sylvestris* parent. Nevertheless, nuclear genes encoding phage-type organellar RNAP are conserved from both the *sylvestris* and the *tomentosiformis* parental genomes. The few differences observed

between the “tom” and “syl” *RpoT* forms do probably not interfere with the enzymatic activity of the gene products (see below). Eukaryotic *RpoT* polymerases are known to depend on (nuclear-encoded) specificity factors to recognise the organellar promoters (Hess and Börner, 1999). It will be interesting to see, once such factors have been isolated, whether both parental gene sets of these interacting polypeptides are conserved as well.

In *N. tabacum,* full length transcripts were characterised for the three *tomentosiformis*-derived genes (*NictaRpoT-tom*). For comparison, *RpoT* coding regions were amplified from *N. sylvestris* (*NicsyRpoT*, see Fig. 3). Alignments of the latter with partial sequences available for the three *sylvestris*-derived genes of *N. tabacum* (*NictaRpoT-syl*) revealed only very few nucleotide exchanges acquired during the estimated six million years since allotetraploidization (99.6 -100% identity in coding regions). This suggests selection with regard to functional sustainment of both parental polymerases.

5.1.1. Phylogenetic origin of the polymerase genes

The striking variety of RNA polymerases found in the plastid compartment can at present not be explained easily in terms of function. However, the analysis of tobacco NEPs, identified during this work, together with data on organellar RNA polymerases in other systems allows several conclusions of its evolution.

The phage-type RNA polymerases obviously did not primarily originate in the genome of the host that acquired the ancestors of mitochondria and plastids. Based on the three-dimensional structures and sequence comparisons, Cermakian et al., (1997) have shown that phage-like RNA polymerases stem from an ancestral DNA polymerase I or a primitive reverse transcriptase. How NEP invaded the eukaryotic genome is however unresolved, although an acquisition by lateral gene tranfer is favoured (Hess and Börner, 1999). It is assumed that initially a signle ancient NEP enzyme was acquired and later duplicated, which is corroborated by similar intron-exon organisation of all NEPs studied so far. This situation of two NEPs is found in the moss *Physcomitrella patens* and also in the monocot maize. This appears to be ancestral to the situation in dicots, where an additional duplication yielded in a third gene. Interestingly, in *Physcomitrella*, both enzymes are dually targeted, whereas in

maize, each enzyme is exclusively directed to only one compartment. Apparently, the targeting signal of duplicated gene became accustomed to the plastid import apparatus in higher plants and specialisation of NEPs might be a derived trait. In this light, the additional duplication in dicots, which yielded an enzyme that became targeted to both compartments, seems an evolutionary regression. Further functional analysis has to verify, whether and what the differences are between the dually targeted *Physcomitrella* and dicot enzymes. The last step known in the evolution of NEPs is the increase in gene number by polyploidisation as determined in this study. This greatly enlarges the field for evolution to produce functional variants and it will be exciting to study paralogue NEPs in alloploids.

In contrast to the NEP family, only one PEP enzyme per organelle has been detected so far. With its prokaryotic organisation, PEP obviously is a direct descendant of the RNA polymerase present in the initial endosymbiont (Igloi and Kössel, 1991). Recently, *rpo* genes coding for PEP subunits have been described from mitochondria of the primitive flagellate *Reclinomonas americana* (Burger et al., 1996, Lang et al., 1997), which suggests that the purely NEP-driven transcription system of higher eukaryote mitochondria is derived from a dual situation that we find presently in the plastid. One may speculate that plastids will ultimately loose PEP as well and hence will completely depend on nuclear regulatory dominance in transcription. This dominance is already apparent when considering that the regulatory subunits of PEP, the sigma-like factors, are encoded by nuclear genes. Interestingly, it has recently been shown that in maize, one of these factors is targeted not only to the plastid but also to the mitochondrium (Beardslee et al., 2002). Whether this means that it serves as a specificity factor for NEP or for some other task has not been shown so far.

5.1.2. Targeting of RpoT;1 and RpoT;2 genes into organelles

To prove the targeting properties of *Nicotiana RpoT* amino termini the subcellular targeting properties of GFP fusion proteins were examined. The import sequence of 153 amino acids of the RpoT;1 gene product translocated the GFP protein into the mitochondrial compartment. *RpoT;2*, in contrast to the *RpoT;1* gene, encodes an N-terminal sequence which directs GFP fusion proteins to two different types of organelles: green fluorescence was observed in mitochondria as well as in plastids (Fig. 6). Thus, the *Nicotiana RpoT;2* displays properties identical to its *Arabidopsis*

homologue. Most interestingly, not only the dual targeting characteristics, but also a second methionine residue (aa 35 in *NicsyRpoT;2*) is conserved between the two species. In *Arabidopsis*, forced initiation at this second methionine codon has been demonstrated to give rise to an exclusively mitochondrial protein, whereas dual targeting has been shown to be a feature of a translational start at the first ATG (Hedtke et al., 2000).

An initiation at the second methionine could be achieved either by shortened mRNAs lacking the upstream AUG, or by “leaky scanning” ribosomes (reviewed in Small et al., 1998). Various 5´ RACE reactions performed on *RpoT;2* did not indicate a presence of such 5´ shortened transcripts. Alternative translation intiation by “leaky scanning” ribosomes, giving rise to bifunctional genes, has been repeatedly described in plants (Mireau et al., 1996; Small et al., 1998; Souciet et al., 1999; Watanabe et al., 2001). The initiation of translation has been shown to depend on the the sequence context around the AUG codon (Kozak, 1991). Most probably, the surroundng sequences of the *RpoT;2* gene are of functional importance. The different plant *RpoT* sequences are obviously duplication products of a mitochondrial-targeted ancestor (Börner et al., 1999; Hess and Börner, 1999).

5.1.3. Non-AUG initiation of translation in RpoT;3

The existence and functional relevance of nuclear-encoded plastid RNA polymerases (NEP) has been unequivocally demonstrated in a number of higher plant systems (Hess et al., 1993, Morden et al., 1991, Allison et al., 1996). Tobacco, especially due to the feasibility of plastid transformation, has played a major role in elucidating NEP functions and promoter preferences.

For *RpoT;3* plastid import was demonstrated (Fig. 6). Characterization of RpoT;3 cDNA 5´ ends failed to reveal an in-frame ATG giving rise to a polypeptide containing a plastid transit peptide. The segment from the *RpoT;3* translation product starting at position 59 neither did in predictions nor in GFP-fusion experiments display plastid targeting properties. The possibility of an incomplete coding sequence was ruled out by analysis of *N. sylvestris RpoT;3* cDNA (Fig. 5). Careful examination of 5´ sequences revealed a potential non-AUG initiation codon conserved in all three

Nicotiana RpoT;3 transcripts. Therefore, three different GFP fusions, each including 12 nucleotides upstream of the CTG, were designed to test the translation initiation capacity of the CUG codon as well as the targeting properties of the derived polypeptide (Hedtke Boris, Berlin). The results proved a plastid import preference of the derived amino terminus as well as an efficient translation initiation at the CUG position that could be eliminated by modifying the codon to CAC (Hedtke et al., 2002).

5.1.4. *RpoT* transcript accumulation

Quantitative RT-PCR showed different transcript amounts of the three *RpoT* genes in wild-type leaves (Hedtke et al., 2002). *NictaRpoT;2* mRNA accumulates to higher levels than the transcript of the exclusively mitochondrial targeted *RpoT;1*, underlining the functional relevance of the dual targeted protein. *RpoT;3* mRNA is the most abundant RpoT transcript in leave tissue. This is in contrast to the assumed role of NEP in early developmental stages of plastids (Hajdukiewicz et al., 1997) and suggests a role of the NEP enzyme in the transcription within mature chloroplasts.

To get insights into the “division of labor” among the transcriptional machineries relying on RpoT core enzymes remains a challenging task. By the characterisation of the *RpoT* gene family in tobacco, with its prospects of well-established plastid reverse genetics as well as detailed knowledge of plastid promoter architecture, new possibilities to dissect *RpoT* functions will emerge.

5.2. Transcription studies of plastids based on array data

Transcription in plastids is unique since it is the only “prokaryotic system” known to operate with different kinds of RNA polymerases. In the nucleus the existence of different RNA polymerases has been established for long and the roles of the individual enzymes was shown to operate in the transcription of different gene classes. Whereas the eukaryotic RNA polymerase I transcribes rRNA genes, type II transcribes mRNAs, and RNA polymerase III is involved in the synthesis of generally small RNA molecules, including *5S* rRNA and tRNAs (Chambon, 1975). Similarly, the different RNA polymerases acting in plastids were proposed to be involved in

differential transcription of photosynthesis-related and house-keeping genes (Hajdukiewicz et al., 1997).

In this study, we have presented an array-based comparative analysis of gene expression profiles of entire plastid chromosomes in leaves of tobacco wild-type and mutants lacking the plastid-encoded RNA polymerase. Whereas transcripts in wild-type may derive from both, PEP and NEP promoters, in mutant plastids transcription initiation is limited to promoters used by NEP and perhaps the recently detected nuclear-encoded third type of plastid RNA polymerase that does not correspond to the phage-type NEP enzymes of 110 kDa (Bligny et al., 2000). Different probe types and expression levels have been compared in order to estimate the reliability and limitations of the approach. Data were evaluated with regard to (i) differential transcription rates and resulting transcript quantities of individual genes/operons by the wild-type or PEP-deficient transcription apparatus (quantitative point of view; see also Fig. 14), (ii) differential initiation/termination and processing of transcripts either driven by NEP or PEP (qualitative point of view; see also Figs. 14, 15 and 16 in Chapter 4.2.3.), (iii) differential degradation/stabilisation, and (iv) differential translation of NEP- and PEP-driven transcripts. It appears that no universal relationship between transcription rates, transcripts levels and the amount of translation products exists, and also no differential, gene class-specific transcription by NEP or PEP, different from the nuclear system.

The evaluation of array-based expression profiles in plastids is particularly challenging in comparison to nuclear expression profiles (e.g. DeRisi et al., 1997; Lashkari et al., 1997; Wodicka et al., 1998) because plastid genes are organised in polycistronic transcription units that are subsequently processed into complex sets of overlapping transcripts which may differ in stability and hence in relative stationary concentrations. Furthermore, transcription of individual genes of a given operon can be driven by several, even intracistronic, promoters (Orozco et al., 1990; Kapoor et al., 1994; Hajdukiewicz et al., 1997) which adds to the complexity of transcript patterns.

5.2.1. The choice of the probe type; technical aspects

Whereas the transcription rates of single plastid genes can be estimated by hybridising labeled run-on transcripts to the arrays, transcript steady-state levels in principle can be assessed by either using labeled RNA or cDNA derived by reverse transcription of plastid RNA as probes. Results show that cDNA probes are not recommendable for the array-based transcription profiling of plastid genomes, since these are transcribed in polycistronic units and reverse transcription can lead to a heterogenous cDNA population covering the upstream cistrons more abundantly than the downstream ones. This can lead to incorrect expression level estimations. Concerning the use of RNA probes, application of transcript probes 5' end-labeled by PNK turned out as most suitable, since 3' end labeling using T4 RNA ligase can lead to co-ligation of independent transcripts (Fig. 12) and therefore may produce artifacts in the determination of actual transcript levels. PNK on the other hand selectively labels 5' hydroxyl termini and therefore excludes primary transcripts from probing (Fig. 13). The fraction of primary transcripts in plastids, however, is generally low, in *Chlamydomonas* not even detectable by capping experiments (Monde et al., 2000b). The finding that transcript levels of the first cistron of a polycistronic operon are not generally under-represented (Figs. 13 and 14) clearly shows that only a minor fraction of transcripts, of both wild-type and mutant plastids, contain triphosphate groups at their 5' ends. Even those transcripts which are known to undergo only limited processing in tobacco plastids (e.g. Eibl et al., 1999) are significantly labeled by PNK, which suggests that their 5' ends are present in the dephosphorylated form. Thus, the error introduced by using PNK-labeled probes is negligible. On the other hand, although PNK labels only a subpopulation of RNAs the transcript quantities estimated by the array approach, especially in PEP-deficient mutants, often exceed those determined by Northern experiments (e.g. *petB*, *petD*, *psbC* or *rps14*, Fig. 15).

These discrepancies can be explained in several ways. (i) Since the arrayed double-stranded DNA fragments hybridise to both, sense and antisense transcripts, a fraction of the signals measured may trace back to antisense RNA. The influence of antisense transcription on array-based data has been checked by the reciprocal experiment in which filters with dotted RNA were individually probed with gene-specific *in vitro* transcripts (Fig. 10). It turned out that wild-type/mutant signal ratios

for transcripts less abundant in mutant leaf tissue were higher using strand-specific probes (Fig. 10). The opposite was observed for transcripts more abundant in mutants. These generally exhibited lower wild-type/mutant transcript ratios when using strand-specific probes. This can be caused by fractions of antisense RNA, with regional similarity in both mutant and wild-type, which hybridise to double-stranded amplicons but are not detected with strand-specific probes. (ii) It is conceivable that a significant fraction of end-labeled RNA molecules which hybridise to a distinct gene probe constitute run-through transcripts initiating at upstream genes (or downstream genes on the opposite DNA strand, as in the case of *petD*) as well as to transcripts initiating within the coding region (sense and/or antisense strand) of the tested gene. For instance, a promoter specifically used by NEP has been identified upstream of *clpP* on the antisense strand within the *psbB* coding region (Hajdukiewicz et al., 1997). (iii) Also, “unspecifically” initiated transcripts may account for the observed discrepancies. Promoter core sequences recognised by NEP exhibit only a weak consensus motif (Weihe and Börner, 1999). Using an *in vitro* transcription assay, directed mutagenesis of a 15 nucleotide segment, containing the P*rpoB* -354 NEP promoter, revealed a major decrease in transcription activity (<30%) only if the core CRT sequence had been changed whereas other positions seem to be less important (Liere and Maliga, 1999). Therefore, NEP transcription may initiate at various sites spread over the plastid chromosome leading to a heterogenous population of perhaps mostly abortive transcripts. This may find support in numerous primer extension signals not only found in intergenic but also in coding regions in PEP-deficient leaf tissue (Birgit Profanter, Diploma thesis, Munich, 2001). However, at least in cases where signals appeared from both, wild-type and mutant RNA, such bands may also be caused by premature termination of cDNA synthesis due to RNA secondary structures. Moreover, both sense and antisense RNA fragments resulting from “unspecific” transcription initiation are able to hybridise to the respective gene-specific PCR products on the array. (iv) Finally, small transcript fragments (i.e. degradation products of unspecifically synthesised or not well stabilised RNAs), which during RNA gel electrophoresis are separated from distinct bands prior to blotting and appear as background in gels, may escape detection in Northern filters but could contribute significantly to the increased transcript quantities measured by array hybridisation.

5.2.2. Transcription and transcript analysis

By hybridising run-on transcripts from wild-type and PEP-deficient plastids to filter arrays carrying all plastid-encoded genes (Fig. 14 B and C) it could be shown that NEP, as well as probably PEP, alone is able to transcribe the entire genetic information of the tobacco plastid chromosome. The transcription rates of individual genes or operons however often do not correlate with steady-state transcript levels, as determined by hybridising end-labeled RNA to arrays (Figs. 14 C and E) and by Northern analysis (Figs. 15, 16, 17 and 19), and only a distinct subset of the corresponding proteins accumulates in PEP-deficient plastids (Figs. 17 B, 18 and 19; DeSantis-Maciossek et al., 1999). In fact, independent of the function of the encoded gene product or of the multisubunit assembly (i.e. photosynthesis-related structure or ribosome), all three patterns, increase, decrease and virtually no change of RNA levels, have been observed comparing transcript/transcription profiles of wild-type and PEP-deficient plastids. Taken together, these findings exclude that selective accumulation of non-photosynthesis related gene products in plants lacking PEP is based primarily on the selective transcription of the corresponding genes (cf. also Krause et al., 2000) as in the nucleus (Green and Chambon, 1988). This raises the question upon the cause and physiological relevance for the observed differences in the expression pattern.

The steady state level of an mRNA is dependent on both its kinetics of synthesis as well as of degradation. The interplay between transcriptional regulation and control of RNA stability is known to be quite complex in plastids (reviewed in Monde et al., 2000b). RNA stabilities can vary during plastid development and their regulation during chloroplast biogenesis in response to photosynthetic activity has been suggested (Monde et al., 2000b). Interestingly, the redox state may influence RNA degradation rates, as in chloroplasts of *Chlamydomdonas reinhardtii* (Salvador and Klein, 1999). Because of the different redox conditions in wild-type and PEP-deficient leaf tissues (DeSantis-Maciossek et al., 1999), RNA stability may considerably differ between transcripts of one and the same gene in the two materials. Stability of single transcripts possibly could also differ in dependence of its synthesis by either NEP or PEP. It is known that stabilisation/degradation of plastid transcripts is often modulated by *cis* elements located predominantly in the 5' and 3' UTRs (Hayes et al.,

1999; Schuster et al., 1999). In addition, transcript degradation can also be initiated by endonucleolytic cleavages within the coding region of an mRNA (Klaff, 1995). Since the PEP and NEP enzymes operate with different promoters which may be separated by hundreds of base pairs, at least 5' ends of the primary transcripts produced by the two polymerases do not coincide. This could influence transcript stability. It is conceivable that (hitherto unknown) polymerase-specific transcription termination signals account for differences in transcript accumulation as well, since the 3' terminal sequence of a plastid transcript, often folding into a secondary structure, is critical for its stability (Hayes et al., 1999, Monde et al., 2000). Run-on transcription in plastids treated with tagetitoxin, a potent inhibitor of PEP, was shown to be linear for over one hour in the presence of heparin (Kim et al., 1993). This suggests that at least *in vitro* NEP does not leave the template which may be the cause for the overall transcription of the plastid chromosome observed in PEP-deficient plants and, implicitly, 3' transcript termini different from those of PEP-driven RNAs.

One of the more interesting findings of the study is that transcripts differ not only quantitatively, but in some cases (e.g. *psaA*, *rps14*, *ycf3* see Figs. 14 B, C, D and E) also qualitatively in wild-type and PEP-deficient plastids (for example Fig. 15). This may be caused by differential processing of transcripts driven by either PEP or NEP. In mammalian cells it has been shown that splicing, processing of 3' ends and termination of transcription all depend on the presence of a carboxy-terminal domain in RNA polymerase II (McCracken et al., 1997). Specific interaction of processing factors with one of the different plastid RNA polymerases could be the cause for the observed differences in transcript patterns as well. Such factors acting at the interface between transcription and RNA processing may be represented by PEP associated nuclear-encoded polypeptides recently identified in mustard (Pfannschmidt et al., 2000). The fate of a transcript may therefore already be determined by its synthesis by one of the different RNA polymerases.

It is known that substantial amounts of thylakoid proteins are often translated from minute amounts of RNA (e.g. Herrmann et al., 1992). However, although corresponding transcripts accumulate to detectable amounts even in PEP-deficient plastids, a number of proteins with photosynthesis-related functions are not

detectable in Δ*rpo* mutant leaf tissue (e.g. PsaA and PsaB photosystem I core proteins; Fig. 18), but some of them are (e.g. pClpP, see Figs. 17 and 19). This may be ascribed to the fact that the message can not be translated in a certain context (e.g. from the polycistronic RNA in the case of *psaA* and *psaB*) as reported for maize *petD* (Barkan et al., 1994) and tobacco *ndhD* (Hirose and Sugiura, 1997) mRNAs. Translation of individual plastid transcripts has also been described to be under control by epistasy in *Chlamydomonas* (Wollman et al., 1995). It is therefore conceivable that the absence of thylakoid membrane complexes in PEP-deficient tobacco plants prevents translation of certain mRNAs by epistatic effects, too. On the other hand, peptides (and other components of the photosynthetic apparatus) may escape detection if they are made but rapidly degraded, because the thylakoid components are not assembled in mutants. Then, however, it remains puzzling why nuclear-encoded thylakoid components clearly accumulate close to wild-type levels in the mutant plastids in the absence of their plastid-encoded partners (Fig. 18).

Nearly all data on NEP-specific transcription patterns were derived from transplastomic tobacco (e.g. Hajdukiewicz et al., 1997; Krause et al., 2000) and naturally arisen mutants (e.g. Hübschmann and Börner, 1998; Silhavy and Maliga, 1998a) lacking the PEP enzyme. The PEP-deficient tobacco plants, as in this work, were generally cultivated in tissue culture and varying quantities of supplements like amino acids, hormones and vitamins may influence gene expression. These ”artifical” systems may be useful to dissect the activity of different RNA polymerases, but it is difficult to envisage presently (with the data available) that at any developmental stage NEP acts without a contemporary activity of the PEP enzyme. Notably, since transcription of PEP-coding genes appears to be driven by NEP, which results in relatively high transcription rates and transcript accumulation of the plastid-encoded *rpo* genes in PEP-deficient mutants, one would expect an increased expression of the PEP genes as soon as NEP is active. A major task of the NEP polymerase could be to ensure an increased expression of the PEP enzyme at certain stages during plastid ontogeny and to allow pronounced expression of non-photosynthesis-related genes (e.g. *accD*, *clpP* and *ycf2*) in certain cell/tissue types. In fact, expression of PEP is significantly increased at an early stage of chloroplast development (Baumgartner et al., 1993; Inada et al., 1996), the period where NEP activity was

proposed to be highest (Hajdukiewicz et al., 1997) which suggests that NEP regulates PEP.

5.3. A new NEP-promoter upstream of the *rpoB* operon?

The plastid-encoded polymerase utilizes promoters of the prokaryotic type with canonical –10 and –35 boxes that are recognized by the sigma subunit of the holoenzyme. In contrast to this, NEP has entirely different *cis* element requirements. Promoter studies in PEP-deficient material allowed the identification of a common CRT motif necessary for initiation of transcription (Weihe et al., 1999). A highly similar promoter structure has been found in mitochondria, which points at an overall conservation of the transcriptional machinery in both organelles. In particular the T of the CRT motif, which seems to be used as the transcriptional initiation site (Liere und Maliga 1999), is strongly conserved.

The major catalytic subunits of the PEP enzyme are encoded by the plastid *rpo* genes. The promoter of the *rpoB/C1/C2* gene cluster has been analysed *in vitro* and appeared to be a typical NEP promoter (Liere und Maliga, 1999). It has been suggested, that NEP transcribes the *rpo* genes during early stages of plastid development thereby boosting plastid transcription (Hajdukiewicz et al., 1997). Hence NEP would regulate PEP.

In this work, *in vivo* promoter studies have been carried out in order to evaluate the importance of the *in vitro* characterized “CRT” promoter of the *rpoB* gene. Therefore, transplastomic tobacco lines carrying a CAC instead of the wild-type CAT core promoter motif have been prepared. Surprisingly, the transformed plants look like wild-type and not at all like the knock-out phenotypes of *rpo* genes, which are albino (Allison et al., 1996; DeSantis-Maciossek et al., 1999). Plants that carry the mutation linked to the *aadA*-cassette and plants without mutation but with the accompanying *aadA*-cassette exhibit identical transcript patterns (Fig. 29 A). Hence, the mutation of the core sequence motif does not influence the expression of the *rpo* gene cluster. However, introduction of the cassette itself entails a pronounced increase in transcript accumulation in comparison to wild-type and PEP-deficient material as both transplastomic lines differ qualitatively and quantitatively from these plants. An

explanation could be that a transcriptional repressor element of *rpoB* is disrupted or shifted by the presence of the *aadA* cassette and hence the element would loose its capacity for repression. In fact, such a negative regulator of *rpoB* transcription has been described at position –1158 to –1184 relative to the *rpoB* start codon (Ishida et al., 1996). In mutant lines, this element is separated from the *rpoB* gene by an additional 1.4 kb of the *aadA* cassette. Likely, this impairs its function and explains the strong expression of *rpoB* in the mutants.

As mentioned above, the *rpoB* transcript pattern is not affected by the point mutation. This is further corroborated by mapping of transcript 5' ends. The single transcript 5' end detected at –294 is found in the mC, mE mutants, wild-type, and Δ*rpoA* tissue. The latter finding implies that the transcriptional activity detected definitely originates in a nucleus-encoded enzyme. No further transcript ends were found. How can these *in vivo* data be explained?

(i) The T to C exchange does not notably affect promoter activity *in vivo*, even though it does *in vitro*. The message initiating at the T in wild-type or C in the mutant is produced and rapidly processed to yield the transcript end found 51 bp downstream (at –294 relative to the *rpoB* start codon).

(ii) The promoter found *in vitro* is not used *in vivo*. Instead, a different, hitherto unknown promoter is utilized by NEP. In the sequence vicinity of the –294 transcript end, no NEP promoter was found, while potential elements for a PEP promoter are present (Fig. 30). If the –294 site is a primary transcript end, a new type of NEP-promoter would have to be postulated here. Possibly, a NEP enzyme in particular responsible for *rpo* transcription exists in plastids, just like - or even identical to - the nucleus encoded polymerase known to act specifically on ribosomal RNA gene transcription (Bligny et al., 1999). Otherwise, the –294 would be a secondarily generated transcript end, whereas the real initiation site and promoter is further upstream. This open question can be settled by capping-experiments.

It is inherently difficult to clarify the role of components of a syntonically regulated system by removing one of the "team mates". Recently published data on pale green *Arabidopsis* mutants defective in the expression of one of the six nuclear-encoded

sigma factors, namely SIG2 (Shirano et al., 2000; Kanamaru et al., 2001), are in line with this assumption. Sigma factors interact with the PEP core for regulation (Allison, 2000). In the absence of SIG2, transcripts for several tRNAs are drastically reduced (Kanamaru et al., 2001), whereas transcripts for photosynthesis-related functions accumulate to wild-type levels. Thus, implying a SIG2 activation of PEP, also expression of various genes for the genetic machinery is strongly dependent on PEP. The finding that all tested “NEP transcripts” (e.g. *clpP* and *accD*) were increased in the absence of SIG2 complicates the situation further. It suggests a direct or indirect influence of SIG2 on NEP activity, too. However, the amount of ClpP is increased, but the amounts of AccD polypeptide were not increased in spite of a several-fold accumulation of their mRNAs. This may be caused by a general decrease in translational activity in the SIG2 mutants which also seems to be reflected in the reduced amounts of photosynthesis-related polypeptides produced from transcripts present in wild-type levels (Kanamaru et al., 2001). These data can not be reconciled with a simple interactive regulation of NEP and PEP.

Analysis of tRNA expression in wild-type and PEP-deficient plastids showed that 20 of 30 tRNAs accumulate to only barely detectable amounts in mutant material. Western results obtained for Rpl2 (DeSantiss-Macioseck et al., 1999), together with pClpP and cytochrome *f* show that even though less amounts of individual tRNAs are present in PEP-lacking material, translation does occur. It remains to be seen if the low expression of certain plastid tRNAs in PEP-deficient mutants is compensated by the import of cytosolic tRNAs as also speculated for *Epifagus virginiana* which does not encode a complete set of tRNAs in its plastid chromosome. Alternatively, translation products detected in PEP-deficient plastids may contain misincorporated amino acids. The accumulation of “mistranslated gene products” could be the reason for the high level of ClpP enzyme found in the mutant plastids, which may be needed to remove these non-natural polypeptides.

In summary, accumulating data on the various RNA polymerases involved in the transcription of plastid chromosomes clearly point out the complexity of mechanisms controlling and implementing plastid gene expression. The plastid is an essential genetic element of the plant cell that traces back to an endosymbiotic event. Data on the expression of plastid chromosomes and its evolution add new insight to the entire

genetic set-up of the plant cell. Our current knowledge might well be only scratching the surface of what awaits. To assess the role of the different plastid RNA polymerases it is mandatory to investigate their activities in biologically relevant situations, i.e. distinct tissues, cells and developmental stages. Without doubt, this is an exceedingly challenging venture and will continue to be a lively field for the future.

5.4. Assembly of cytochrome *b*$_6$*/f* complex

The cytochrome *b*$_6$*/f* complex, the least complex of the major assemblies of the thylakoid membrane is functionally situated between two larger, chlorophyll-containing complexes, PSII and PSI, in the electron transport chain of the thylakoid membrane. It catalyses the reduction of plastocyanin at the expense of reduced plastoquinon, contributing to the formation of a proton gradient across the membrane used for ATP synthesis (Choquet et al., 2000). The cytochrome complex is homologous to the mitochondrial and bacterial cytochrome *bc*$_1$ complexes that have similar cytochrome *b*- and *c*-type subunits. Isolation of nuclear mutants of subunits of the complex together with reverse genetics of plastid-encoded subunits by plastid transformation have led to the identification of subunit interactions. A comprehensive overview of the assembly is however still lacking. One goal of this study therefore was to generate knock-out mutants for each of the plastid-encoded subunits of the complex, which was achieved by biolistic plastid transformation.

The mutant material obtained was proven to be homoplastomic by Southern and Northern analysis and had the typical pale *hcf*-phenotype (except the *petL* mutants) described earlier for some of the *pet* mutants of *Chlamydomonas* and tobacco (Barkan et al., 1994, Hager et al., 1999, Meurer et all., 1998, Monde et al., 2000; Takahashi et al., 1999).

The method of choice for pinpointing assembly defects in the knock-out material is the immunoblot analysis of cytochrome *b*$_6$*/f* complex subunits. Specific antisera were used to detect cytochrome *f* (PetA), cytochrome *b*$_6$ (PetB), subunit IV (PetD), subunit VII (PetL) and the nuclear-encoded Rieske Fe-S protein (PetC) in all knock-out mutants. In addition, to check the accumulation of other thylakoid membrane complexes, antisera for the α-subunit of the ATPase (AtpA), 9 kD protein of PSI

(PsaC), 9 kDa α subunit of cytochrome b_{559} (PsbE) and D2 of PS II (PsbD) were tested. The presence or absence of subunits in each mutant led to the following conclusions regarding the assembly scheme of the cytochrome b_6/f complex in higher plants.

1. Δ*petG*

None of the investigated cytochrome b_6/f subunits was found to accumulate in the Δ*petG* mutants, which suggests that most probably subunit V is a core component and appears first during assembly.

2. Δ*petB*/Δ*petD*

In contrast to Δ*petG*, Δ*petB* and Δ*petD* mutants partially accumulate to some extent other cytochrome b_6/f subunits, for example cytochrome *f* (more then 5%) or Rieske protein (~ 3%). This finding is in accordance with data from *Chlamydomonas* that point out a role of cytochrome b_6 and subunit IV in early complex assembly (Kuras and Wollman, 1994, Wollman, 1999). One migth speculate that subunit IV, cytochrome b_6 and subunit V form a pre-complex as they have a similar effect on the accumulation of the other subunits.

Cytochrome *f*, which is believed to be the membrane-anchor for the pre-complex, accumulates in these two mutants only in minor amounts, in accordance with the idea that cytochrome *f* down-regulates its own accumulation when not incorporated into the complex (so-called CES regulation; control of epistatic synthesis; Baymann et al., 1999, Choquet et al., 1998). The Rieske protein also fails to accumulate in Δ*petB* and Δ*petD* mutants similar to the situation in the corresponding knock-out mutants in *Chlamydomonas* (Kuras et al., 1995).

3. Δ*petA*

It has so far been assumed that the Rieske protein is only stable when associated with membranes, like for instance when it is anchored to the membrane *via* cytochrome *f* (Kuras et al., 1994). In the Δ*petA* mutant however, the Rieske protein is quite stable, probably because it is also able to associate with a pre-complex in the stroma consisting of the cytochrome b_6, subunit IV and potentially subunit V (Gupta et al., 2002). In fact, subunit IV accumulates to substantial amounts, whereas for

cytochrome b_6, no entirely conclusive data are available, since the wild-type signal could not be detected at a 5% dilution. The reduction of subunit IV could mean that the accumulation of the pre-complex could be directly impaired due to the failure to interact with the membrane anchor of cytochrome *f*. Alternatively, the loss of cytochrome *f* as an enhancer of *petB/D* mRNA translation could lead to a decreased accumulation of cytochrome b_6 and subunit IV.

4. ∆*petN*

The loss of the *petN* gene product entails a severe reduction of the pre-complex as indicated by the loss of cytochrome b_6 and subunit IV. This suggests a function of PetN in pre-complex assembly, probably as a constituting subunit, which has been postulated (Hager et al., 1999).

Astonishingly, cytochrome *f* accumulates to substantial levels (approximately 25%) in this mutant. Possibly, the low amounts of a residual pre-complex detected are sufficient to stabilize some cytochrome *f*. However, if the pre-complex stabilizes cytochrome *f* in a 1:1 stoichiometry, then the 2% accumulation of subunit IV relative to wild-type can not account for a 25% accumulation of cytochrome *f*. Therefore, PetN must have an additional effect on cytochrome *f*. One possibility is that PetN could reduce cytochrome *f* accumulation by repressing its translation and would therefore be a new CES-component of the complex (Chouqet et al., 1998). It should be noted that Hager et al. (1999) did not detect any cytochrome *f* in their ∆petN knock-out mutants. As their knock-out strategy was comparable to ours – the reading frame in both mutants is completely deleted – external factors, as for example tissue culture conditions, may account for the observed differences.

4. ∆*petL*

In *Chlamydomonas* cells deficient for *petL* (subunit VII), other subunits of the cytochrome b_6/f complex accumulate substantially (50% of wild-type levels, Takahashi et al., 1996). The tobacco ortholog of *Chlamydomas petL* was knocked-out by insertional (∆*petL*1) as well as by deletional (∆*petL*2) mutagenesis in the course of this work. The two mutants do not show any substantial differences in the accumulation of other subunits of the complex consistent with the *Chlamydomonas* work. Cytochrome *f*, cytochrome b_6, subunit IV and Rieske accumulate to about 50%

of the wild-type levels. Both mutants are also not impaired in chlorophyll accumulation. Surprisingly, Western analyses when total cellular extract was used show accumulation of the PetL subunit in both types of mutants, Δ*petL*1 and Δ*petL*2, as well as in all other deletion mutants of the cytochrome complex. As the Southern data clearly show that both mutants are homoplastomic (Fig. 40), the protein is most probably not expressed from the plastid gene. It is tempting to speculate that an active copy of the gene is present in the nucleus and its gene product imported into plastids. Such a rather recent transfer of an organellar gene has occurred in other plant lineages as well (Hansmann et al., 2000, Kadowaki et al., 1996, Martin et al., 1990, Sanchez et al., 1996). However, the simultaneous presence of an intact plastid and a functional nuclear copy would be unprecedented in plants. In any case, the presence of the PetL protein in both mutants indicates that these plants are not functional knock-outs of this gene and hence, it is impossible to draw any conclusions from this material regarding the assembly of the holocomplex. The proposed involvement of petL in dimerisation of the cytochrome b_6/f complex (Breyton et al., 1997) has to be studied by different means. Loss of cytochrome b_6/f subunits could potentially have an impact on the accumulation of subunits of other thylakoid membrane complexes. Therefore, selected subunits of the ATP synthetase, PSI and PSII have been analysed serologically in the seven tobacco *pet*-mutants. Interestingly, the subunits of the ATPase and PSI are more abundant in mutants than in wild-type, with the exception of Δ*petL*, which shows signals similar to wild-type. In contrast, PSII-subunits are reduced in all Δ*pet* mutants again with the exception of Δ*petL*. Western analyses using an antibody against cytochrome b_{559} (PsbE) exhibit two bands in wild-type and in the Δ*petL*-mutants. So far, no processing of cyt b_{559} has been described in the literature so that the nature of this band remains enigmatic at the moment. In any case, loss of the cytochrome b_6/f complex does seem to have an impact on the turnover of PSII proteins.

All in all, this work provides a basis for understanding the biogenesis of the cytochrome b_6/f complex in higher plants. In the future, isolation of the native complex from transplastomic mutants and *in vivo* labeling experiments should give a more detailed insight into what happens when single subunits are missing. In particular cytochrome *f* (Δ*petA*) and subunit VIII (Δ*petN*) are interesting lines for studies as they seem to act as control factors of complex assembly. This could also

give a more comprehensive view of how CES regulation works in higher plants in comparison to algae, like *Chlamydomonas*.

Aside of interactions between subunits of the complex, additional (presumably nuclear) factors may be involved in expression and assembly of the complex (Barkan et al., 1994; Wollman, 1999 and articles therein). Furthermore, cross-interactions with other multiprotein complexes, in particular the balancing of the machinery involved in electron transport, has to be taken into account, as well.

6. Summary

The plant cell contains an integrated, compartmentalised genetic system that comprises three subgenomes: nucleus, chondriome and plastome. In this study, the examination of the plastid compartment was studied focusing on two aspects, plastid transcription and assembly processes of the cytochrome b_6f complex.

1. Tobacco nuclear homologues of *Arabidopsis* RNA polymerases have been isolated. According to the allotetraploid nature of tobacco as many as six different cDNAs have been found. These cDNAs could be assigned to the orthologs present in the parental lines of tobacco, namely *N. tomentosiformis* and *N. sylvestris*. All six genes are expressed. Targeting studies using GFP fusion constructs showed that one pair of RNA polymerases is targeted to the mitochondrion, the second pair to the plastid, and the third pair to the both compartments. Its transient peptide is modularly organised. Translation of the plastid enzyme initiates a GUG codon, a second example in plants. This confirms and extends the results from *Arabidopsis* and shows that the organization of the NEP enzyme in dicots is conserved.

2. A macroarray was developed for studying the expression of entire plastid chromosomes. Application of this new approach on tissue deficient in PEP activity allowed to determine qualitatively and quantitatively the impact of the two polymerases on the expression of every single gene on the plastid chromosome. The use of 5' labeling of RNA ends proved to be most reliable for probe labeling.The approach proved for the first time that almost the entire plastid chromosome is a template for both, the NEP as well as the PEP enzymes. This cannot be reconciled with data from the literature. In fact, it was found that NEP is responsible for the expression of many genes coding for the components of the photosynthetic apparatus, a task formerly attributed exclusively to PEP. Northern analysis substatiated the array-based data, but also provided evidence that several tRNAs are transcribed by PEP only, a finding that indicates import of nuclear-encoded tRNAs into plastids. A second aspect of this study concerned the comparison of different expression levels, beside arrays on stationary

RNA concentrations, arrays with run-on transcripts, Northern and Western analyses of related regions of the plastid chromosome. Surprisingly, no obvious regulatory relations between these levels were found, indicating their highly complex interplay during biogenesis.

3. Mutational analysis of the *rpoB* promoter revealed that the NEP promoter deduced from *in vitro* mapping does not have any relevance for the expression of the operon *in vivo*. This questions previous assumptions that PEP is under simple regulatory control of NEP. Interestingly, the transcription initiation site determined in this study is also not consistant with previous data. Rather, it is situated closely to a putative PEP promoter, which may imply that autoregulatory loops play a role in PEP expression and therefore sheds new light on the regulatory interaction of the two enzymatic systems.

4. The cytochrome b_6f complex is not only involved in electron transport and proton pumping, but plays a crucial role in regulating thylakoid membrane dynamics (redox sensor) as well. It consists of eight subunit species. Each by each, all six plastid-encoded subunits of the cytochrome b_6f complex were inactivated by targeted gene disruption. Immunological studies of both nuclear- and plastid-encoded subunits in these mutants allowed to study the hierarchy of assembly, if any, of this complex. The data point out unexpected differences to the assembly of the complex in *Chlamydomonas*, for instance, in the formation of intermediate complexes. In addition, evidence is provided that the plastid-encoded *petL* gene is not essential for complex formation, possibly because a nuclear encoded homologue can serve as a substitute. This work forms a basis for studying the biogenesis of the thylakoid membrane complex in higher plants, which appears to be different from that of alga. It may also aid for approaches to verify or disprove constraints for phylogenetic gene translocation from plastids to the nucleus.

7. Abbreviations

Amp	ampicillin
AMP	adenosinemonophopsphate
APS	amoniumpersulphat
ATP	adenosinetriphosphate
BAC	bacterial artificial chromosome
BAP	6-benzylamino-purine
Bp	basepair
BSA	albumin, calf bovine serum, fraction V
Ci	Curie
cp	chloroplast
cpm	counts per minute
CTAB	hexadecyltrimethyl-ammonium bromide
dATP	desoxy-adenosintriphosphate
dCTP	desoxy-cytosintriphosphate
DMSO	dimethyl-sulfoxide
dGTP	desoxy-guanosintriphosphate
dNTP	desoxy-nucleotidetriphosphate
dTTP	desoxy-thymidintriphosphate
DEAE	diethyl-aminoethyl
DNA	desoxyribonucleic acid
DNAse	desoxyribonuclease
dsDNA	double-stranded DNA
DTT	1,4-dithiotreithol
EDTA	ethylendiamine-tetra-acetate
EGTA	ethyleneglycol-bis-(aminoethylether)-N,N,N',N'-tetraacetic acid
h	hour
HEPES	N-2-hydroxyethylpiperazin-N'-2-ethansulfonic acid
IAA	α-indolyl-acetic acid
IPTG	isopropyl-b-D-thiogalactopyranoside
kbp	kilobase pairs
M	molar
mCi	milliCurie
MES	morpholinoethane sulfonic acid
min	minute

mM	millimolar
MOPS	2-morpholinoethansulfonic acid
MS	Murashige and Skoog-medium
mt	mitochondria
NAA	β-naphtalene acetic acid
nt	nucleotides
OD	optical density
ON	over night
PEG	polyethylenglycol
PMSF	phenylmethylsulphonylfluoride
PVP	polyvinylpyrrolidone
RNA	ribonucleic acid
RNAse	ribonuclease
rpm	rounds per minute
RT	room temperature
sec	seconds
SDS	sodium dodecylsulfate
TEMED	N,N,N',N'-tetraethylmethylethylendienamine
Tris	trishydroxyaminomethane
U	units
Vol	volume
v/v	volume per volume
w/v	weight per volume
w/w	weight per weight
X-Gal	5-bromo-4-chloro-3-indolyl-β-D-galactopiranoside

8. Literature

Abdallah F, Salamini F, Leister D. (2000) A prediction of the size and evolutionary origin of the proteome of chloroplasts of Arabidopsis. Trends Plant Sci., **5** (4):141-2.

Allen JF, Forsberg J. (2001) Molecular recognition in thylakoid structure and function. Trends Plant Sci., **6** (7):317-26. Review.

Allison LA, Simon LD, Maliga P. (1996) Deletion of rpoB reveals a second distinct transcription system in plastids of higher plants. EMBO J., **15** (11):2802-9.

Allison LA. (2000) The role of sigma factors in plastid transcription. Biochimie. 2000, **82** (6-7):537-48. Review.

Barakat A, Szick-Miranda K, Chang IF, Guyot R, Blanc G, Cooke R, Delseny M, Bailey-Serres J. (2001) The organization of cytoplasmic ribosomal protein genes in the Arabidopsis genome. Plant Physiol., **127** (2):398-415.

Barkan A. (1989) Tissue-dependent plastid RNA splicing in maize: transcripts from four plastid genes are predominantly unspliced in leaf meristems and roots. Plant Cell., **1** (4):437-45.

Barkan A, Walker M, Nolasco M, Johnson D. (1994) A nuclear mutation in maize blocks the processing and translation of several chloroplast mRNAs and provides evidence for the differential translation of alternative mRNA forms. EMBO J., **13** (13):3170-81.

Barkan A and Goldschmidt-Clermont M. (2000) Participation of nuclear genes in chloroplast gene expression. Biochimie., **82** (6-7):559-72. Review.

Bayley CC, Morgan M, Dale EC, Ow DW. (1992) Exchange of gene activity in transgenic plants catalyzed by the Cre-lox site-specific recombination system. Plant Mol Biol., **18** (2):353-61.

Besendahl A, Qiu YL, Lee J, Palmer JD, Bhattacharya D. (2000) The cyanobacterial origin and vertical transmission of the plastid tRNA(Leu) group-I intron. Curr Genet. ;**37**(1):12-23.

Binder S. and Brennicke A. (1993) A tRNA gene transcription initiation site is similar to mRNA and rRNA promoters in plant mitochondria. Nucleic Acids Res., **21**(22):5012-9.

Binder S, Hatzack F, Brennicke A. (1995) A novel pea mitochondrial in vitro transcription system recognizes homologous and heterologous mRNA and tRNA promoters., J Biol Chem., **270** (38):22182-9.

Binder S, Marchfelder A, Brennicke A. (1996) Regulation of gene expression in plant mitochondria. Plant Mol Biol., **32** (1-2):303-14.

Bligny M, Courtois F, Thaminy S, Chang CC, Lagrange T, Baruah-Wolff J, Stern D, Lerbs-Mache S. (2000) Regulation of plastid rDNA transcription by interaction of CDF2 with two different RNA polymerases. EMBO J., **19** (8):1851-60.

Bock R, Kossel H, Maliga P. (1994) Introduction of a heterologous editing site into the tobacco plastid genome: the lack of RNA editing leads to a mutant phenotype. EMBO J., **13** (19):4623-8.

Bogorad L. (1981) Chloroplasts. J Cell Biol., **91**(3 Pt 2):256s-270s. Review.

Boudreau E, Nickelsen J, Lemaire SD, Ossenbuhl F, Rochaix JD. (2000) The Nac2 gene of Chlamydomonas encodes a chloroplast TPR-like protein involved in psbD mRNA stability. EMBO J., **19** (13):3366-76.

Bruce BD and Malkin R. (1991) Biosynthesis of the chloroplast cytochrome b6f complex: studies in a photosynthetic mutant of Lemna. Plant Cell., **3** (2):203-12.

Casano LM, Zapata JM, Martin M, Sabater B. (2000) Chlororespiration and poising of cyclic electron transport. Plastoquinone as electron transporter between thylakoid NADH dehydrogenase and peroxidase. J Biol Chem. **275** (2):942-8.

Cermakian N, Ikeda TM, Cedergren R, Gray MW. (1996) Sequences homologous to yeast mitochondrial and bacteriophage T3 and T7 RNA polymerases are widespread throughout the eukaryotic lineage. Nucleic Acids Res., **24** (4):648-54.

Choquet Y, Stern DB, Wostrikoff K, Kuras R, Girard-Bascou J, Wollman FA. Translation of cytochrome f is autoregulated through the 5' untranslated region of petA mRNA in Chlamydomonas chloroplasts. Proc Natl Acad Sci U S A., **95** (8):4380-5.

Comai L, Tyagi AP, Winter K, Holmes-Davis R, Reynolds SH, Stevens Y, Byers B. (2000) Phenotypic instability and rapid gene silencing in newly formed arabidopsis allotetraploids. Plant Cell., **12** (9):1551-68.

Eberhard, S, Drapeir, D. and Wollman, F.A. (2002) Searching limiting steps in the expression of chloroplast-encoded proteins: relations between gene copy number, transcription, transcript abundance and translation rate in the chloroplast of Chlamydomonas reinhardtii. Plant J., **31**, 149-161.

Eibl C, Zou Z, Beck a, Kim M, Mullet J, Koop HU. (1999) In vivo analysis of plastid psbA, rbcL and rpl32 UTR elements by chloroplast transformation: tobacco plastid gene expression is controlled by modulation of transcript levels and translation efficiency Plant J., **19** (3):333-45.

Eisermann A, Tiller K, Link G. (1990) In vitro transcription and DNA binding characteristics of chloroplast and etioplast extracts from mustard (Sinapis alba) indicate differential usage of the psbA promoter. EMBO J., **9** (12):3981-7.

Esposito D, Hicks AJ, Stern DB., (2001) A role for initiation codon context in chloroplast translation. Plant Cell., **13** (10):2373-84.

Erie DA, Yager TD, von Hippel PH. (1992) The single-nucleotide addition cycle in transcription: a biophysical and biochemical perspective. Annu Rev Biophys Biomol Struct., **21**:379-415. Review.

Felder S, Meierhoff K, Sane AP, Meurer J, Driemel C, Plucken H, Klaff P, Stein B, Bechtold N, Westhoff P. (2001) The nucleus-encoded HCF107 gene of Arabidopsis provides a link between intercistronic RNA processing and the accumulation of translation-competent psbH transcripts in chloroplasts. Plant Cell., **13** (9):2127-41.

Ferat JL, Michel F. Group II self-splicing introns in bacteria. Nature. 1993 Jul 22; **364**(6435):358-61.

Fisk DG, Walker MB, Barkan A. (1999) Molecular cloning of the maize gene crp1 reveals similarity between regulators of mitochondrial and chloroplast gene expression. EMBO J., **18** (9):2621-30.

Freyer R, Kiefer-Meyer MC, Kossel H. (1997) Occurrence of plastid RNA editing in all major lineages of land plants.Proc Natl Acad Sci U S A. **94** (12):6285-90.

Gockel G. and Hachtel W. (2000) Complete gene map of the plastid genome of the nonphotosynthetic euglenoid flagellate Astasia longa. Protist., **151** (4):347-51.

Green S, Chambon P. (1988) Nuclear receptors enhance our understanding of transcription regulation. Trends Genet. **4** (11):309-14. Review.

Hager M, Biehler K, Illerhaus J, Ruf S, Bock R. (1999) Targeted inactivation of the smallest plastid genome-encoded open reading frame reveals a novel and essential subunit of the cytochrome b(6)f complex. EMBO J., **18** (21):5834-42.

Hajdukiewicz PT, Allison LA, 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** (13):4041-8.

Hanic-Joyce PJ, Gray MW. (1991) Accurate transcription of a plant mitochondrial gene in vitro. Mol Cell Biol., **11** (4):2035-9.

Hallick RB, Hong L, Drager RG, Favreau MR, Monfort A, Orsat B, Spielmann A, Stutz E. (1993) Complete sequence of Euglena gracilis chloroplast DNA. Nucleic Acids Res., **21** (15):3537-44.

Harris N, Foster JM, Kumar A, Davies HV, Gebhardt C, Wray JL. (2000) Two cDNAs representing alleles of the nitrate reductase gene of potato (Solanum tuberosum L. cv. Desiree): sequence analysis, genomic organization and expression. J Exp Bot., **51** (347):1017-26.

Hausmann, S. and Martin W. (2000) Phylogeny pf 33 ribosomal and six other proteins encoded in an ancient gene cluster that is conserved across prokaryotic genomes: influence of excluding poorly alignable sites from analysis. Int. Syst. Evol. Microbiol., **50**Pt4, 1655-1663.

Hawley DK, McClure WR. (1983) Compilation and analysis of Escherichia coli promoter DNA sequences. Nucleic Acids Res., **11** (8):2237-55. Review.

Hayes R, Kudla J, Gruissem W. (1999) Degrading chloroplast mRNA: the role of polyadenylation. Trends Biochem Sci., **24** (5):199-202. Review.

Hedtke B, Borner T, Weihe A. (1997) Mitochondrial and chloroplast phage-type RNA polymerases in Arabidopsis. Science, **277** (5327):809-11.

Hess WR, Prombona A, Fieder B, Subramanian AR, Borner T. (1993) Chloroplast rps15 and the rpoB/C1/C2 gene cluster are strongly transcribed in ribosome-deficient plastids: evidence for a functioning non-chloroplast-encoded RNA polymerase. EMBO J., **12** (2):563-71.

Hess WR, Muller A, Nagy F, Borner T. (1994) Ribosome-deficient plastids affect transcription of light-induced nuclear genes: genetic evidence for a plastid-derived signal. Mol Gen Genet., 242(3):305-12.

Hess W.R. and Börner T. (1999a) Organellar RNA polymerases of higher plants. Int Rev Cytol., **190**:1-59.

Hedtke B, Wagner I, Borner T, Hess WR. (1999b) Inter-organellar crosstalk in higher plants: impaired chloroplast development affects mitochondrial gene and transcript levels. Plant J., **19** (6):635-43.

Helmann J.D. and Chamberlin MJ. (1988) Structure and function of bacterial sigma factors.
Annu Rev Biochem., **57**:839-72. Review.

Helmann JD. (1999) Anti-sigma factors. Curr Opin Microbiol., **2** (2):135-41. Review.

Hentze MW. (2001) Protein synthesis. Believe it or not-translation in the nucleus. Science., **293** (5532):1058-9.

Hepp, A. (2002) Phopsphoproteine und Phospholylierungen in der Thylakoidmembran. Diploma Thesis. Ludwig-Maximilian-Universität.

Herrmann RG, Westehoff P, Link G. (1992) Biogenesis of plastids in higher plants. In: Herrmann RG. (Ed.) Plant Gene Research, 6, Cell Organells, Springer Verlag, 275-349.

Herrmann RG. (1996) Photosynthesis research: aspects and perspectives. In: Frontiers in Molecular Biology, Genetics in Photosynthesis. Andersson B., Salter AH, and Barter J (Ed.). Oxford University Press, 1-44.

Herrmann RG. (1997) Eukaryotism, toward a new interpretation. In: Eukaryotism and symbiosis: intertaxonic combination versus symbiotic adaption. Schenk HEA., Herrmann RG, Jeon KW. Und Schwemmler W. (Ed.) Springer-Verlag Berlin Heidelberg,73-118.

Higgs DC, Shapiro RS, Kindle KL, Stern DB. (1999) Small cis-acting sequences that specify secondary structures in a chloroplast mRNA are essential for RNA stability and translation. Mol Cell Biol., **19** (12):8479-91.

von Hippel PH, Fairfield FR, Dolejsi MK. (1994) On the processivity of polymerases. Ann N Y Acad Sci., **726**:118-31.

Hiratsuka J, Shimada H, Whittier R, Ishibashi T, Sakamoto M, Mori M, Kondo C, Honji Y, Sun CR and Meng BY. (1989) The complete sequence of the rice (Oryza sativa) chloroplast genome: intermolecular recombination between distinct tRNA genes accounts for a major plastid DNA inversion during the evolution of the cereals., Mol Gen Genet., **217** (2-3):185-94.

Hirose T, Kusumegi T, Sugiura M. (1998) Translation of tobacco chloroplast rps14 mRNA depends on a Shine-Dalgarno-like sequence in the 5'-untranslated region but not on internal RNA editing in the coding region. FEBS Lett., **430** (3):257-60.

Hoch B, Maier RM, Appel K, Igloi GL, Kossel H. (1991) Editing of a chloroplast mRNA by creation of an initiation codon. Nature., **353** (6340):178-80.

Hoeckema A, Hirsch PR, Hooykaas PJJ, and Schillerport RA (1983) A binary plant vector strategy based on separation of *vir*- and T-region of the *Agrobacterium tumefaciens* Ti-plasmid. Nature, **244**, 360-366.

Hübschmann T. and Borner T. (1998) Characterisation of transcript initiation sites in ribosome-deficient barley plastids. Plant Mol Biol., **36** (3):493-6.

Igloi G.L. and Kössel H. (1992) The Transcriptional Apparatus of Chloroplasts. Critical Reviews in Plant Sciences, **10** (6): 525-558.

Ikeda TM, Gray MW. (1999) Identification and characterization of T3/T7 bacteriophage-like RNA polymerase sequences in wheat. Plant Mol Biol., **40** (4):567-78.

Inada H, Seki M, Morikawa H, Nishimura M, Iba K. (1997) Existence of three regulatory regions each containing a highly conserved motif in the promoter of plastid-encoded RNA polymerase gene (rpoB). Plant J., **11** (4):883-90.

Iratni R, Baeza L, Andreeva A, Mache R, Lerbs-Mache S. (1994) Regulation of rDNA transcription in chloroplasts: promoter exclusion by constitutive repression. Genes Dev., **8** (23):2928-38.

Iratni R, Diederich L, Harrak H, Bligny M, Lerbs-Mache S. (1997) Organ-specific transcription of the rrn operon in spinach plastids. J Biol Chem., **272** (21):13676-82.

Kadowaki, K, Kubo, N., Ozawa, K. and Hirai A. (1996) Targeting presequence acquisition after mitochondrial gene transfer to the nucleus occurs by duplication of existing targeting signals., EMBOJ, **15** (23): 6652-6661.

Kanamaru K, Nagashima A, Fujiwara M, Shimada H, Shirano Y, Nakabayashi K, Shibata D, Tanaka K, Takahashi H. (2001) An Arabidopsis sigma factor (SIG2)-dependent expression of plastid-encoded tRNAs in chloroplasts. Plant Cell Physiol., **42** (10):1034-43.

Kaneko,T., Tanaka,A., Sato,S., Kotani,H., Sazuka,T., Miyajima,N., Sugiura,M. and Tabata,S. (1995) Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis sp. strain PCC6803. I. Sequence features in the 1 Mb region from map positions 64% to 92% of the genome. DNA Res., **2** (4), 153-166.

Kapoor M, Nagai T, Wakasugi T, Yoshinaga K, Sugiura M. (1997) Organization of chloroplast ribosomal RNA genes and in vitro self-splicing activity of the large subunit rRNA intron from the green alga Chlorella vulgaris C-27. Curr Genet., **31** (6):503-10.

Kapoor S. and Sugiura M. (1999) Identification of two essential sequence elements in the nonconsensus type II PatpB-290 plastid promoter by using plastid transcription extracts from cultured tobacco BY-2 cells. Plant Cell., **11** (9):1799-810.

Karnauchov I, Herrmann RG, Klosgen RB. (1997) Transmembrane topology of the Rieske Fe/S protein of the cytochrome b6/f complex from spinach chloroplasts. FEBS Lett., **408** (2):206-10.

Kelly JL, Greenleaf AL, Lehman IR. (1986) Isolation of the nuclear gene encoding a subunit of the yeast mitochondrial RNA polymerase., J Biol Chem., **261** (22):10348-51.

Kelly JL and Lehman IR. (1986) Yeast mitochondrial RNA polymerase. Purification and properties of the catalytic subunit., J Biol Chem., **261** (22):10340-7.

Kempin SA, Liljegren SJ, Block LM, Rounsley SD, Yanofsky MF, Lam E. (1997) Targeted disruption in Arabidopsis. Nature., **389** (6653):802-3.

Krause K, Maier RM, Kofer W, Krupinska K, Herrmann RG. (2000) Disruption of plastid-encoded RNA polymerase genes in tobacco: expression of only a distinct set of genes is not based on selective transcription of the plastid chromosome. Mol Gen Genet., **263** (6):1022-30.

Kuras R, Wollman FA. (1994) The assembly of cytochrome b6/f complexes: an approach using genetic transformation of the green alga Chlamydomonas reinhardtii. EMBO J., **13** (5):1019-27.

Lee HS and Chen ZJ. (2001) Protein-coding genes are epigenetically regulated in Arabidopsis polyploids. Proc Natl Acad Sci U S A., **98** (12):6753-8.

Levy H, Kindle KL, Stern DB. (1999) Target and specificity of a nuclear gene product that participates in mRNA 3'-end formation in Chlamydomonas chloroplasts. J Biol Chem., **274** (50):35955-62.

Liere K. and Maliga P. (1999) In vitro characterization of the tobacco rpoB promoter reveals a core sequence motif conserved between phage-type plastid and plant mitochondrial promoters. EMBO J., **18** (1):249-57.

Lisowsky T and Michaelis G. (1989) Mutations in the genes for mitochondrial RNA polymerase and a second mitochondrial transcription factor of Saccharomyces cerevisiae., Mol Gen Genet., **219** (1-2):125-8.

Liu B, Brubaker CL, Mergeai G, Cronn RC, Wendel JF. (2001) Polyploid formation in cotton is not accompanied by rapid genomic changes. Genome., **44** (3):321-30.

Lohan AJ and Wolfe KH. (1998) A subset of conserved tRNA genes in plastid DNA of nongreen plants. Genetics., **150** (1):425-33.

Lonetto MA, Brown KL, Rudd KE, Buttner MJ. (1994) Analysis of the Streptomyces coelicolor sigE gene reveals the existence of a subfamily of eubacterial RNA polymerase sigma factors involved in the regulation of extracytoplasmic functions. Proc Natl Acad Sci U S A., **91** (16):7573-7.

Madueno F, Bradshaw SA, Gray JC. (1994) The thylakoid-targeting domain of the chloroplast Rieske iron-sulfur protein is located in the N-terminal hydrophobic region of the mature protein. J Biol Chem., **269** (26):17458-63.

Maier RM, Neckermann K, Igloi GL, Kossel H. (1995) Complete sequence of the maize chloroplast genome: gene content, hotspots of divergence and fine tuning of genetic information by transcript editing., J Mol Biol., **251** (5):614-28.

Maier RM, Zeltz P, Kossel H, Bonnard G, Gualberto JM, Grienenberger JM. (1996) RNA editing in plant mitochondria and chloroplasts. Plant Mol Biol., **32** (1-2):343-65. Review.

Maki Y, Tanaka A, Wada A. (2000) Stoichiometric analysis of barley plastid ribosomal proteins. Plant Cell Physiol., **41** (3):289-99.

Malek O, Lattig K, Hiesel R, Brennicke A, Knoop V. (1996) RNA editing in bryophytes and a molecular phylogeny of land plants. EMBO J. **15** (6):1403-11.

Martin, W., Lagrange, T., Li, YF, Bisanz-Seyer, C. and Mache R. (1990) Hypothesis for the evolutionary origin of the chloroplast ribosomal protein L21 of spinach. Curr. Genet., 18(6): 553-556.

Martin W, Mustafa AZ, Henze K, Schnarrenberger C. (1996) Higher-plant chloroplast and cytosolic fructose-1,6-bisphosphatase isoenzymes: origins via duplication rather than prokaryote-eukaryote divergence. Plant Mol Biol., **32** (3):485-91.

Martin W, Stoebe B, Goremykin V, Hapsmann S, Hasegawa M, Kowallik KV. (1998) Gene transfer to the nucleus and the evolution of chloroplasts. Nature., **393** (6681):162-5.

Masters BS, Stohl LL, Clayton DA. (1987) Yeast mitochondrial RNA polymerase is homologous to those encoded by bacteriophages T3 and T7., Cell, **51** (1):89-99.

McCracken S, Fong N, Yankulov K, Ballantyne S, Pan G, Greenblatt J, Patterson SD, Wickens M, Bentley DL. (1997) The C-terminal domain of RNA polymerase II couples mRNA processing to transcription. Nature., **385** (6614):357-61.

Meng BY, Tanaka M, Wakasugi T, Ohme M, Shinozaki K, Sugiura M. (1988) Cotranscription of the genes encoding two P700 chlorophyll a apoproteins with the gene for ribosomal protein CS14: determination of the transcriptional initiation site by in vitro capping. Curr Genet., **14** (4):395-400.

Meurer J, Berger A, Westhoff P. (1996) A nuclear mutant of Arabidopsis with impaired stability on distinct transcripts of the plastid psbB, psbD/C, ndhH, and ndhC operons. Plant Cell., **8** (7):1193-207.

Michel, F, Ferat J.L. (1995) Structure and activities of group II introns. Annu Rev Biochem. 1995;64:435-61. Review.

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** (3):299-307.

Monde RA, Zito F, Olive J, Wollman FA, Stern DB. (2000) Post-transcriptional defects in tobacco chloroplast mutants lacking the cytochrome b6/f complex. Plant J., **21** (1):61-72.

Morton, BR (1998) Selection on the codon bias of chloroplast and cyanelle genes in different plant and algal lineages. J Mol Evol., **46** (4):449-59.

Mould RM, Kapazoglou A, Gray JC. (2001) Assembly of cytochrome f into the cytochrome bf complex in isolated pea chloroplasts. Eur J Biochem., **268** (3):792-9.

Mulligan RM, Leon P, Walbot V. (1991) Transcriptional and posttranscriptional regulation of maize mitochondrial gene expression. Mol Cell Biol., **11** (1):533-43.

Newton KJ, Winberg B, Yamato K, Lupold S, Stern DB. (1995) Evidence for a novel mitochondrial promoter preceding the cox2 gene of perennial teosintes. EMBO J., **1**; 14(3):585-93.

Nickelsen J, Link G. (1991) RNA-protein interactions at transcript 3' ends and evidence for trnK-psbA cotranscription in mustard chloroplasts. Mol Gen Genet., **228** (1-2):89-96.

Oda K, Yamato K, Ohta E, Nakamura Y, Takemura M, Nozato N, Akashi K, Kanegae T, Ogura Y, Kohchi T, et al. (1992) Gene organization deduced from the complete sequence of liverwort Marchantia polymorpha mitochondrial DNA. A primitive form of plant mitochondrial genome. J Mol Biol., **223** (1):1-7.

Pfannschmidt T. and Link G. (1997) A and B forms of plastid DNA-dependent RNA polymerase from mustard (Sinapis alba L.) transcribe the same genes in a different developmental context. Mol Gen Genet., **257**(1) :35-44.

Pfannschmidt T, Ogrzewalla K, Baginsky S, Sickmann A, Meyer HE, Link G. (2000) The multisubunit chloroplast RNA polymerase A from mustard (Sinapis alba L.). Integration of a prokaryotic core into a larger complex with organelle-specific functions. Eur J Biochem. **267** (1):253-61.

Profanter, B. (2002) Analyse differentieller NEP- und PEP-Promotornutzungen am Beispiel des Operons *psa*A/*psa*B/*rps*14 sowie dessen Regulation auf posttranskriptioneller und posttranslationeller Ebene. Diploma Thesis., Ludwig-Maximilian-Universität.

Race HL, Herrmann RG, Martin W. (1999) Why have organelles retained genomes? Trends Genet., **15** (9):364-70.

Rapp WD and Stern DB. (1992) A conserved 11 nucleotide sequence contains an essential promoter element of the maize mitochondrial atp1 gene. EMBO J., **11**(3):1065-73.

Rapp JC, Baumgartner BJ, Mullet J. (1992) Quantitative analysis of transcription and RNA levels of 15 barley chloroplast genes. Transcription rates and mRNA levels vary over 300-fold; predicted mRNA stabilities vary 30-fold. J Biol Chem., **267** (30):21404-11.

Rapp WD, Lupold DS, Mack S, Stern DB. (1993) Architecture of the maize mitochondrial atp1 promoter as determined by linker-scanning and point mutagenesis. Mol Cell Biol., **13** (12):7232-8.

Reed ML, Peeters NM, Hanson MR. (2001) A single alteration 20 nt 5' to an editing target inhibits chloroplast RNA editing in vivo. Nucleic Acids Res., **29** (7):1507-13.

Reed ML, Lyi SM, Hanson MR. (2001) Edited transcripts compete with unedited mRNAs for trans-acting editing factors in higher plant chloroplasts. Gene., **272** (1-2):165-71.

Rivier C, Goldschmidt-Clermont M, Rochaix JD. (2001) Identification of an RNA-protein complex involved in chloroplast group II intron trans-splicing in Chlamydomonas reinhardtii. EMBO J., **20** (7):1765-73.

Rochaix JD, Kuchka M, Mayfield S, Schirmer-Rahire M, Girard-Bascou J, Bennoun P. (1989) Nuclear and chloroplast mutations affect the synthesis or stability of the chloroplast psbC gene product in Chlamydomonas reinhardtii. EMBO J., **8** (4):1013-21.

Rochaix JD. (2001) Assembly,function, and dynamics of the photosynthetic machinery in Chlamydomonas reinhardtii. Plant Physiol., **127** (4):1394-8. Review.

Rolland N, Dorne AJ, Amoroso G, Sultemeyer DF, Joyard J, Rochaix JD. (1997) Disruption of the plastid ycf10 open reading frame affects uptake of inorganic carbon in the chloroplast of Chlamydomonas. EMBO J., **16** (22):6713-26.

Ruf S, Kossel H, Bock R. (1997) Targeted inactivation of a tobacco intron-containing open reading frame reveals a novel chloroplast-encoded photosystem I-related gene. J Cell Biol., **139** (1):95-102.

Sanchez, H, Fester, T., Kloska, S., Schröder, W. and Schuster W. (1996) Transfer of rps19 to the nucleus involves the gain of an RNP-binding motif which may functionally replace RPS13 in Arabidopsis mitochondria. EMBOJ, **15** (9): 2138-2149.

de Santis-Maciossek G, Kofer W, Bock A, Schoch S, Maier RM, Wanner G, Rudiger W, Koop HU, Herrmann RG. (1999) Targeted disruption of the plastid RNA polymerase genes rpoA, B and C1: molecular biology, biochemistry and ultrastructure. Plant J., **18** (5):477-89.

Sasaki Y, Kozaki A, Ohmori A, Iguchi H, Nagano Y. (2001) Chloroplast RNA editing required for functional acetyl-CoA carboxylase in plants. J Biol Chem., **276** (6):3937-40.

Satoh J, Baba K, Nakahira Y, Tsunoyama Y, Shiina T, Toyoshima Y. (1999) Developmental stage-specific multi-subunit plastid RNA polymerases (PEP) in wheat. Plant J., **18** (4):407-15.

Small I, Wintz H, Akashi K, Mireau H. (1998) Two birds with one stone: genes that encode products targeted to two or more compartments. Plant Mol Biol., **38** (1-2):265-77. Review.

Schmitz-Linneweber C, Maier RM, Alcaraz JP, Cottet A, Herrmann RG, Mache R. (2001) The plastid chromosome of spinach (Spinacia oleracea): complete nucleotide sequence and gene organization. Plant Mol Biol., **45** (3):307-15.

Schmitz-Linneweber C, Regel R, Du TG, Hupfer H, Herrmann RG and Maier RM. (2002) The plastid chromosome of Atropa belladonna and its comparison with that of Nicotiana tabacum: the role of RNA editing in generating divergence in the process of plant speciation. Mol Biol Evol. in press.

Sheveleva EV, Giordani NV, Hallick RB. (2002) Identification and comparative analysis of the chloroplast alpha-subunit gene of DNA-dependent RNA polymerase from seven Euglena species. Nucleic Acids Res., **30** (5):1247-54.

Shiina T, Allison L, Maliga P. (1998) rbcL Transcript levels in tobacco plastids are independent of light: reduced dark transcription rate is compensated by increased mRNA stability. Plant Cell., **10** (10):1713-22.

Silhavy D, Maliga P. (1998) Mapping of promoters for the nucleus-encoded plastid RNA polymerase (NEP) in the iojap maize mutant. Curr Genet., **33** (5):340-4.

Silhavy D. and Maliga P. (1998) Plastid promoter utilization in a rice embryogenic cell culture. Curr Genet. **34** (1):67-70.

Sommer S, Kohle A, Yazaki K, Shimomura K, Bechthold A, Heide L. (1999) Genetic engineering of shikonin biosynthesis hairy root cultures of Lithospermum erythrorhizon transformed with the bacterial ubiC gene. Plant Mol Biol., **39** (4):683-93.

Sriraman P, Silhavy D, Maliga P. (1998) The phage-type PclpP-53 plastid promoter comprises sequences downstream of the transcription initiation site. Nucleic Acids Res., **26** (21):4874-9.

SugiuraM, Shinozaki K, Zaita N, Kusuda Mand KumanoM (1986) Clone bank of the tobacco (Nicotiana tabacum) chloroplast genome as a set of overlapping restriction endonuclease fragments: mapping eleven ribosomal genes., Plant Sci., **44**: 211–216.

Sugita M. and Sugiura M. (1996) Regulation of gene expression in chloroplasts of higher plants. Plant Mol Biol., **32** (1-2):315-26. Review.

Summer H, Pfannschmidt T, Link G. (2000) Transcripts and sequence elements suggest differential promoter usage within the ycf3-psaAB gene cluster on mustard (Sinapis alba L.) chloroplast DNA. Curr Genet., **37** (1):45-52.

Takahashi Y, Rahire M, Breyton C, Popot JL, Joliot P, Rochaix JD. (1996) The chloroplast ycf7 (petL) open reading frame of Chlamydomonas reinhardtii encodes a small functionally important subunit of the cytochrome b6f complex. EMBO J., **15** (14):3498-506.

Tanaka K, Oikawa K, Ohta N, Kuroiwa H, Kuroiwa T, Takahashi H. (1996) Nuclear encoding of a chloroplast RNA polymerase sigma subunit in a red alga. Science., **272** (5270):1932-5.

Thomson MC, Macfarlane JL, Beagley CT, Wolstenholme DR. (1994) RNA editing of mat-r transcripts in maize and soybean increases similarity of the encoded protein to fungal and bryophyte group II intron maturases: evidence that mat-r encodes a functional protein. Nucleic Acids Res. **22** (25):5745-52.

Tiller K, Link G. (1993) Sigma-like transcription factors from mustard (Sinapis alba L.) etioplast are similar in size to, but functionally distinct from, their chloroplast counterparts. Plant Mol Biol., **21** (3):503-13.

Tjian R. (1996) The biochemistry of transcription in eukaryotes: a paradigm for multisubunit regulatory complexes. Philos Trans R Soc Lond B Biol Sci., **351** (1339):491-9. Review.

Tracy RL and Stern DB. (1995) Mitochondrial transcription initiation: promoter structures and RNA polymerases. Curr Genet., **28** (3):205-16. Review.

Trifa Y, Privat I, Gagnon J, Baeza L, Lerbs-Mache S. (1998) The nuclear RPL4 gene encodes a chloroplast protein that co-purifies with the T7-like transcription complex as well as plastid ribosomes. J Biol Chem., **273** (7):3980-5.

Trifa Y. and Lerbs-Mache S. (2000) Extra-ribosomal function(s) of the plastid ribosomal protein L4 in the expression of ribosomal components in spinach. Mol Gen Genet., **263** (4):642-7.

Troxler RF, Zhang F, Hu J, Bogorad L. (1994) Evidence that sigma factors are components of chloroplast RNA polymerase. Plant Physiol., **104** (2):753-9.

Unseld M, Marienfeld JR, Brandt P, Brennicke A. (1997) The mitochondrial genome of Arabidopsis thaliana contains 57 genes in 366,924 nucleotides. Nat Genet., **15** (1):57-61.

Zito F, Finazzi G, Delosme R, Nitschke W, Picot D, Wollman FA. (1999) The Qo site of cytochrome b6f complexes controls the activation of the LHCII kinase. EMBO J., **18** (11):2961-9.

Zerges W, Girard-Bascou J, Rochaix JD. (1997) Translation of the chloroplast psbC mRNA is controlled by interactions between its 5' leader and the nuclear loci TBC1 and TBC3 in Chlamydomonas reinhardtii. Mol Cell Biol., **17** (6):3440-8.

Valkova-Valchanova M, Darrouzet E, Moomaw CR, Slaughter CA, Daldal F. (2000) Proteolytic cleavage of the Fe-S subunit hinge region of Rhodobacter capsulatus bc(1) complex: effects of inhibitors and mutations. Biochemistry., **39** (50):15484-92.

Vera A. and Sugiura M. (1994) A novel RNA gene in the tobacco plastid genome: its possible role in the maturation of 16S rRNA. EMBO J., **13** (9):2211-7.

Voelker R. and Barkan A. (1995) Nuclear genes required for post-translational steps in the biogenesis of the chloroplast cytochrome b6f complex in maize. Mol Gen Genet., **249** (5):507-14.

Vogel J, Borner T, Hess WR. (1999) Comparative analysis of splicing of the complete set of chloroplast group II introns in three higher plant mutants. Nucleic Acids Res. **27**(19):3866-74.

Voelker R, Mendel-Hartvig J, Barkan A. (1997) Transposon-disruption of a maize nuclear gene, tha1, encoding a chloroplast SecA homologue: in vivo role of cp-SecA in thylakoid protein targeting. Genetics., **145** (2):467-78.

Wakasugi T, Tsudzuki J, Ito S, Nakashima K, Tsudzuki T, Sugiura M. (1994) Loss of all ndh genes as determined by sequencing the entire chloroplast genome of the black pine Pinus thunbergii. Proc Natl Acad Sci U S A., **91** (21):9794-8.

Weihe A, Hedtke B, Borner T. (1997) Cloning and characterization of a cDNA encoding a bacteriophage-type RNA polymerase from the higher plant Chenopodium album. Nucleic Acids Res., **25** (12):2319-25.

Weihe A, Borner T. (1999) Transcription and the architecture of promoters in chloroplasts.
Trends Plant Sci., **4** (5):169-170.

Westhoff P. and Herrmann RG. (1988) Complex RNA maturation in chloroplasts. The psbB operon from spinach. Eur J Biochem., **171** (3):551-64.

Wilson RJ, Denny PW, Preiser PR, Rangachari K, Roberts K, Roy A, Whyte A, Strath M, Moore DJ, Moore PW, Williamson DH. (1996) Complete gene map of the plastid-like DNA of the malaria parasite Plasmodium falciparum. J Mol Biol., **261** (2):155-72.

Wilson RJ, Williamson DH. (1997) Extrachromosomal DNA in the Apicomplexa. Microbiol Mol Biol Rev., **61** (1):1-16. Review.

Wimberly BT, Brodersen DE, Clemons WM Jr, Morgan-Warren RJ, Carter AP, Vonrhein C, Hartsch T, Ramakrishnan V. (2000) Structure of the 30S ribosomal subunit. Nature., **407** (6802):327-39.

Wolff G, Kück U. (1996) Transcript mapping and processing of mitochondrial RNA in the chlorophyte alga Prototheca wickerhamii. Plant Mol Biol., **30** (3):577-95.

Wolfe KH, Katz-Downie DS, Morden CW, Palmer JD. (1992) Evolution of the plastid ribosomal RNA operon in a nongreen parasitic plant: accelerated sequence evolution, altered promoter structure, and tRNA pseudogenes. Plant Mol Biol., **18** (6):1037-48.

Wolfe KH (1994) Similarity between putative ATP-binding sites in land plant plastid ORF2280 proteins and the FtsH/CDC48 family of ATPases. Curr. Gen., 25(4):379-383.

Wollman FA, Minai L, Nechushtai R. (1999) The biogenesis and assembly of photosynthetic proteins in thylakoid membranes1 Biochim Biophys Acta., **1411** (1):21-85.

Yamaguchi K, Subramanian AR. (2000) The plastid ribosomal proteins. Identification of all the proteins in the 50 S subunit of an organelle ribosome (chloroplast). J Biol Chem., **275** (37):28466-82.

Yamaguchi K, von Knoblauch K, Subramanian AR. (2000) The plastid ribosomal proteins. Identification of all the proteins in the 30 S subunit of an organelle ribosome (chloroplast). J Biol Chem., **275** (37):28455-65.

Yoshinaga K, Iinuma H, Masuzawa T, Uedal K. (1996) Extensive RNA editing of U to C in addition to C to U substitution in the rbcL transcripts of hornwort chloroplasts and the origin of RNA editing in green plants. Nucleic Acids Res. **24** (6):1008-14.

Young DA, Allen RL, Harvey AJ, Lonsdale DM. (1998) Characterization of a gene encoding a single-subunit bacteriophage-type RNA polymerase from maize which is alternatively spliced. Mol Gen Genet., **260** (1):30-7.

9. Appendix

No.	Gene	Strand	Nr. Primermix Julia	PCR-Produkt from		Primer a for PCR-Product	Primer b for PCR-Product	Length of PCR-Product [bp]	Primer for gene spec. primermix for cDNA	Filter position
				cDNA	gen. DNA					
1	psbA	b	2		x	AT7 psbA 5' gta ata cga ctc act ata ggg cta gca tat tgg aag atc aat 3' Tm = 69,4 °C	ApsbA 5' tta ctt ggc gta gct tgt tac a 3* Tm=56,5 °C	433	AT7 psbA	A1
2	matK	b	3		x	T7matK 5' cgc gcg taa tac gac tca cta tag ggg cca gat cat tga tac aaa a 3' Tm = ? °C	K6 5' tta ggg cat ccc att agt aaa 3'	355	T7matK	B1
3	psbK	a	7		x	AT7 psbK 5' gta ata cga ctc act ata ggg ctt tag ttt gat cgg tat ctg 3' Tm = 70,4 °C	ApsbK 5' ttg cca aac aaa ggc taa ga 3' Tm = 53,2 °C	136	ApsbK	C1
4	psbI	a	8		x	AT7 psbI 5' gta ata cga ctc act ata ggg ttc ttc acg tcc agg att acg 3' Tm = 72,4 °C	ApsbI 5'gct tac tct caa act ctt cgt3' Tm=60°C	63	AT7 psbI	D1
	psbK psbI	a			x	AT7psbK s.o.	AT7psbI s.o.	678	(AT7psbI) (ApsbK)	

	Operon									
5	atpA	b	12		x	AT7 atpA 5' gta ata cga ctc act ata ggg aaa tta cat tag tag gaa ta 3' Tm = 66,4°C	atpHF6 5' cgt ata tga gcc tct tca a 3'	585	AT7 atpA	E1
6	atpF	b	13	x		AT7 atpF 5' gta ata cga ctc act ata ggg cta atg gta cgt aaa tgt aac tc 3' Tm = 70,4°C	AatpF 5' ttc gtt tct ttg ggc cac tg 3' Tm = 57,3°C	488	AT7 atpF	F1
7	atpH	b	14		x	AT7 atpH 5' gta ata cga ctc act ata ggg att aaa caa aag gat tcg caa 3' Tm = 68,5°C	AatpH 5' gaa tcc act gat ttc tgc cg 3' Tm = 57,3°C	205	AT7 atpH	G1
8	atpI	b	15		x	AT7 atpI 5' gta ata cga ctc act ata ggg ccc tcc atg gat tca cct a 3' Tm = 70,4°C	AatpI 5' ggg tta tac gat ata tcc ggt g3' Tm = 70,4°C	657	AT7 atpI	H1
9	rpoC2	b	17		x	AT7 rpoC2 5' gta ata cga ctc act ata ggg ctc gct tca gat atg aaa ctt tg 3' Tm = 71,3°C	ArpoC2 5' agc taa gcc tta ttt ggc cac 3' Tm = 57,9 °C	515	AT7 rpoC2	A2
10	rpoC1	b	18		x	AT7 rpoC1 5' gta ata cga ctc act ata ggg cag tat act agg cct tct aa 3' Tm = 70,4°C	ArpoC1 5' gcc taa tgg aga gat agt tg 3' Tm = 55,3°C	305	AT7 rpoC1	B2
11	rpoB	b	19		x	AT7 rpoB	ArpoB	446	AT7 rpoB	C2

						5' cta ata cga ctc act ata ggg atc caa aac ttg aga taa tgg 3' Tm = 69,4°C	5' gtc aaa aga aaa tgc cat ttt g 3' Tm = 52,8 °C			
12	ycf6	a	21		x	AT7 ycf6 5' cta ata cga ctc act ata ggg tta gta gga ccc act aga gtc c 3' Tm = 73,3°C	Aycf6 5' agg gga cat aat tca tat gg 3' Tm = 53,2 °C	117	AT7 ycf6	D2
13	psbM	b	22		x	AT7 psbM 5' cta ata cga ctc act ata ggg tta ata att aat cgt ttg act 3' Tm = 66,5°C	psbM 5' aaa tat tct tgc att tat tgc 3'	62	AT7 psbM	E2
14	psbD	a	27		x	AT7 psbD 5' cta ata cga ctc act ata ggg tag aac ctc ctc agg gga at 3' Tm = 72,4°C	ApsbD 5' tga acc cat ttc ata tga tg 3' Tm = 51,2°C	461	AT7 psbD	F2
15	psbC	a	28		x	psbCR2T7 5'cta ata cga ctc act ata ggg tct tct aaa tcg tcc aca cta ac3'	Hpsbcfor 5' gga gca atg aac cta ttt ga 3' Tm = 53,2°C	502	psbCR2T7	G2
16	orf105	b	30		x	AT7 orf105 5' cta ata cga ctc act ata ggg cat ata gaa tca atg gat tca t 3' Tm = 68,5°C	Aorf105 5' atg aac tat ata cca ttt gat c 3' Tm = 50,9 °C	105	AT7 orf105	H2
17	ycf9	a	31		x	ycf9revT7 5'cta ata cga ctc act ata ggg cat att ctc ttc gga acg at3'	Aycf9 5' tgc cta gat cac gga taa at 3' Tm = 53,2 °C	381	S.ycf9 rev 5' aaa gac taa tcc aat cca taa 3' Tm = 50,1 °C	A3

18	psaB	b	35		x	AT7 psaB 5' cta ata cga ctc act ata ggg aaa cgc tcg gtt tcc att tt 3' Tm = 70,5 °C	psaBfor 5' att tgg ttt ggt att gct ac 3' Tm = 51,2 °C	437	AT7 psaB	B3
19	psaA	b	36		x	AT7 psaA 5' cta ata cga ctc act ata ggg ccc cca agt taa act ggt act 3' Tm = 72,4 °C	HpsaAfor2 5' gat cct aaa gag ata cca ctt 3' Tm = 53,9 °C	830	AT7 psaA	C3
20	ycf3	b	37	x		P9 5' gta atg aca gat cac ggc cat 3' Tm = 57,9 °C	Aycf3 5' tgc cta gat cac gga taa at 3' Tm = 53,2°C	353	P9 5' gta atg aca gat cac ggc cat 3'	D3
21	orf74	b	38		x	AT7 orf74 5' gta ata cga ctc act ata ggg atc tag gca tag gta gca at 3' Tm = 70,4°C	Aorf74 5' gaa atc gaa tcg cct tga aa 3' Tm = 53,2°C	224	AT7 orf74	E3
22	orf70A	a	42		x	AT7 orf 70a 5' gta ata cga ctc act ata ggg act atc ttc tat agg aga cct 3' Tm = 70,4°C	Aorf70a 5' gga atg gtt agt tat aac taa 3' Tm = 50,1°C	218	AT7 orf 70a	F3
23	ndhJ	b	45		x	AT7ndhJ 5' gta ata cga ctc act ata ggg caa tga gca tct tgt att tca 3' Tm = 69,4°C	AndhJ 5' ggt cgt ttg tct gct tgg ct 3' Tm = 59,4°C	469	AT7ndhJ	G3
24	ndhK	b	46		x	AT7 ndhK (JuliaT7) 5' gta ata cga ctc act ata ggg cta att cac taa ttc ggg3'	AndhK 5' ggt caa gac tct cta gtt ta 3'	594	AT7 ndhK	H3

							Tm = 53,2°C			
25	ndhC	b	47		x	AT7 ndhC 5' gat aat acg act cac tat agg gcc cat tgg ttc tat acc cga 3' Tm = 72,4°C	AndhC 5' atg ttt ctg ctt tac gaa tat 3' Tm = 50,1°C	154	AT7 ndhC	A4
26	atpE	b	50		x	AT7 atpE 5' gat aat acg act cac tat agg gtt tgc tga gct tct tgt gga t 3' Tm = 71,3 °C	AatpE 5' gac ctt aaa tct tag tgt act 3' Tm = 52°C	280	AT7 atpE	B4
27	atpB	b	51		x	AT7 atpB 5' gat aat acg act cac tat agg gca aac caa ctc tca tac gag c 3' Tm = 72,3 °C	AatpB 5' ggc tat agc tat gag tgc ta 3' Tm = 55,3 °C	528	AT7 atpB	C4
28	rbcL	a	52		x	AT7rbcL 5' gat aat acg act cac tat agg gat cat ttc ttc gca tgt acc 3' Tm = 70,4 °C	ArbcL 5' gga ctt acc agc ctt gat cg 3' Tm = 59,4 °C	536	AT7rbcL	D4
29	accD	a	53		x	accDrevT7 5'gat aat acg act cac tat agg gta cta gag aga cca tgt ctt c3'	accDfor 5' gct aaa taa atc aat ggg cag 3'	795	accDrevT7	E4
30	psaI	a	54		x	AT7 psaI 5' gat aat acg act cac tat agg gtc cca gcg gat cta aac aat 3' Tm = 71,4 °C	ApsaI 5' gaa cct tcc atc tat ttt tg 3' Tm = 51,2 °C	111	AT7 psaI	F4
31	ycf4	a	55		x	ycf4RT7	Aycf4	346	ycf4RT7	G4

						5'gat aat acg act cac tat agg gag gaa tat gcg acg att3'	5' gcg atc aga aca tat atg gat 3' Tm = 54 °C			
32	ycf10	a	56		x	ycf10R2T7 5'gat aat acg act cac tat agg gta gat att tgt cga gaa g3'	ycf10for 5'ggc aaa aaa gaa agc att ca3'	342	ycf10R2T7	H4
33	petA	a	57		x	petART7 5'gat aat acg act cac tat agg gta ctc gtt aat ggt tga tca3'	petAfor2 5' ctg ttc tta ttt tac cgg agg 3' Tm = 55,9 °C	572	petART7	A5
34	orf99	a	58		x	AT7 orf99 5' gat aat acg act cac tat agg gac tgg aag gat tcc tct ttg g 3' Tm = 72,3 °C	Aorf99 5' gat cag aaa aat gag aaa aat 3' Tm = 48,1 °C	295	AT7 orf99	B6
35	psbJ	b	59		x	AT7 psbJ 5' gat aat acg act cac tat agg gta cta gag gga tga acc caa t 3' Tm = 71,3 °C	ApsbJ (Sorf99) 5'act gga agg att cct ctt tgg 3'	112	AT7 psbJ	C6
36	psbL	b	60		x	AT7 psbL 5' gat aat acg act cac tat agg gtc tta att gaa gaa ata att gga 3' Tm = 67,6 °C	ApsbL 5' tac gac aca atc aaa ccc ga 3' Tm = 55,3 °C	119	AT7 psbL	D6
37	psbF	b	61		x	AT7 psbF new 5' gat aat acg act cac tat agg ggt tta tcg ttg gat gaa ctg 3' Tm = 70,4 °C	ApsbF 5' gcc cta atg act ata gat cg 3' Tm = 55,3 °C	127	AT7psbF new	E6

38	psbE	b	62		x	AT7 psbE 5' gat aat acg act cac tat agg gat cta cta aat tca tcg agt t 3' Tm = 68,5 °C	ApsbE 5' aag cac agg aga acg ttc gt 3' Tm = 57,3 °C	236	AT7 psbE	F6
39	orf103	b	63		x	AT7 orf103 5' gat aat acg act cac tat agg gca cac aat tta agt aga tgc g 3' Tm = 70,4 °C	Aorf103 5' gtt cca cat ttt atc ttt cc 3' Tm = 51,2 °C	256	AT7 orf103	G6
40	petL	a	64		x	AT7 petL 5' gat aat acg act cac tat agg gaa tct caa tga caa tga aca3' Tm = 68,5 °C	ApetL 5' tta gcg gct tta act ata ac 3' Tm = 51,2 °C	176	petLrevx 5' aag tcg tat ctt gct cag acc 3' Tm : 57,9 °C	H6
41	petG	a	65		x	AT7 petG 5' gat aat acg act cac tat agg gtt aat taa tca aag gtc caa3' Tm = 67,5 °C	ApetG 5' gat tga agt ttt tct att tg 3' Tm = 47,1 °C	120	AT7 petG	A6
42	psaJ	a	68		x	AT7 psaJ 5' gat aat acg act cac tat agg gct tct att cta ccc cca ttc c3' Tm = 72,3 °C	ApsaJ 5' gag atc taa aaa cat atc tct 3' Tm = 50,1 °C	329	S.psaJrev 5' aaa cga tta atc tct atc aa 3' Tm : 47,1 °C	B6
43	clpP	b	73	x		AT7 clpP 5' gat aat acg act cac tat agg gcg tga ggg aat gct agga cgt t3' Tm = 73,3 °C	clppx1for 5' gat gca tct tgg gtt gac gta3' Tm = 57 °C	313	AT7 clpP	C6
44	psbB	a	74		x	AT7 psbB 5' gat aat acg act cac tat agg gtt tca ata ttt ccc ata cg 3'	ApsbB 5' ggt ttt tga tcc ttc tga tc 3'	579	AT7 psbB	D6

						Tm = 68,4 °C	Tm = 53,2 °C			
45	psbT	a	75		x	AT7 psbT 5' gat aat acg act cac tat agg gat gaa gaa tta ttt cac ctt ttt a3' Tm = 67,6 °C	ApsbT 5' gga agc att ggt tta tac at 3' Tm = 51,2 °C	123	AT7 psbT	E6
46	psbN	b	76		x	AT7 psbN 5' gat aat acg act cac tat agg gct cag tac ttc aac tag tc3' Tm = 70,4 °C	ApsbN 5' gtc gcc atc ttt ata tct gg 3' Tm =55,3 °C	128	AT7 psbN	F6
47	psbH	a	77		x	AT7 psbH 5' gat aat acg act cac tat agg gaa ttc cat cca ata aaa cgg a3' Tm = 69,4 °C	ApsbH 5' tat ggc tac aca aac tgt tg 3' Tm = 53,2 °C	210	AT7 psbH	G6
48	petB	a	78		x	ADT7petB 5' cgc gcg taa tac gac tac cta tag ggc aaa gaa taa cca gtt ac3' Tm = 73,2 °C	ApetB 5' ata tga ttg gtt cga aga acg 3' Tm= 54 °C	452	ADT7petB	H6
49	petD	a	79		x	AT7 petD 5' gat aat acg act cac tat agg gat tta tca ata ggt aat gtt gc3' Tm = 68,5 °C	ApetD 5' caa aaa aac ctg act tga atg 3' Tm= 52 °C	450	AT7 petD	A7
50	rpoA	b	80		x	AT7 rpoA 5' gta ata cga ctc act ata ggg att caa aag gtc caa caa tgt 3' Tm= 69,4 °C	ArpoA 5' gga tta caa ata gag aga aat 3' Tm= 50,1 °C	457	AT7 rpoA	B7
51	ycf2	a	93		x	AT7 ycf2	Aycf2	503	AT7 ycf2	C7

						5' gat aat acg act cac tat agg gaa aca tgc tca ata tca t3' Tm= 67,4 °C	5' ggt tga ttg ttt cgc tcc tg 3' Tm= 57,3 °C			
52	ycf15	a	94		x	AT7 ycf15 5' gta ata cga ctc act ata ggg ata gac gcc tgt tgg cat tcc 3' Tm= 73,3 °C	ycf15for 5' cta ctg ctg aaa cat gga ag 3' Tm= 55,3 °C	188	AT7 ycf15	D7
53	orf92	a	95		x	AT7 orf92 5'gat aat acg act cac tat agg gag agc tcg gat cga atc ggt3'	Aorf92 (Sorf115) 5'tat att gaa ttt acc cga tag tc3'	329	AT7 orf92	E7
54	orf115	b	96		x	AT7 orf115 5'gat aat acg act cac tat agg gta tat tga att tac ccg ata gtc3' Tm= 60 °C	Aorf115 (Sorf92) 5'gag agc tcg gat cga atc ggt3' Tm= 68 °C	329	AT7 orf115	F7
55	orf79	a	98		x	AT7 orf79 5' gat aat acg act cac tat agg gat tag gaa tac gcg taa tc 3' Tm= 69,4 °C	Aorf79 5' gtg atg tat gga ata tat ga 3' Tm= 49,1 °C	225	AT7 orf79	G7
56	ndhB	b	99	x		AT7 ndhB 5' taa tac gac tca cta tag gga aat ata ggc ctg cct gc 3' Tm= 70,5 °C	p12 5' ggt cta atg agg cta cta tg 3' Tm= 55,3 °C	770	AT7 ndhB	H7
57	orf70B	a	102		x	AT7 orf70b 5' gat aat acg act cac tat agg gaa ata gga ttg act acc g 3' Tm= 69,5 °C	Aorf70b 5' ggt agg gtt acc att atc ct 3' Tm= 55,3 °C	170	AT7 orf70b	A8

58	orf131	b	103		x	AT7 orf131 5' gat aat acg act cac tat agg gat tct ggc aat gta gtt gg 3' Tm= 70,4 °C	Aorf131 5' gga aat agg att gac tac cg 3' Tm= 55,3 °C	333	AT7 orf131	B8
59	orf75	b	113		x	AT7 orf75 5' gat aat acg act cac tat agg gga att ctt taa ttc aga atc aat 3' Tm= 67,6 °C	Aorf75 5' ggg agc cac tac gaa gaa gtt 3' Tm= 59,8 °C	207	AT7 orf75	C8
60	ndhF	b	115		x	AT7 ndhF 5' gat aat acg act cac tat agg gtt caa gtt gat tca aat tat tcg 3' Tm= 68,5°C	AndhF 5' tac ggt tgg aat tat ggt tc 3' Tm= 53,2°C	699	AT7 ndhF	D8
61	ycf5	a	119		x	AT7 ycf5 5' gat aat acg act cac tat agg gta acc cta ttc cta aaa gat3' Tm= 56°C	Aycf5 5'aag tat gat gat att ggg ct3' Tm= 54°C	487	AT7 ycf5	E8
62	ndhD	b	120		x	AT7 ndhD 5'gataatacgactcactatag gga att aat tct aac tcc cac3' Tm= 56°C	AndhD 5'acg aat tat ttt cct tgg tt3' Tm= 52°C	439	AT7 ndhD	F8
63	psaC	b	121		x	AT7 psaC 5'gataatacgactcactatagg gat tca cat ctc tta caa cc3' Tm= 56°C	ApsaC 5'gt ata gac cca atg tca cat 3' Tm= 56°C	177	AT7 psaC	G8
64	ndhE	b	122		x	AT7 ndhE 5'gataatacgactcactataggg cta att att caa caa att tga t3' Tm= 54°C	sibenik1 5' tga gca caa gtc gaa ata tgg 3' Tm= 54 °C	245	AT7 ndhE	H8

65	ndhG	b	123		x	AT7 ndhG 5'gataatacgactcactataggg caa ttg cac cta tta aag caa3' Tm= 56°C	AndhG 5'gat tta tct gaa cca ata ca3' Tm= 52°C	510	AT7 ndhG	A9
66	ndhI	b	124		x	AT7 ndhI 5'gataatacgactcactatag ggt aaa ttc gaa att gtt cga a3' Tm= 58°C	AndhI 5'gct ccc tat gat aac tga at3' Tm= 56°C	482	AT7 ndhI	B9
67	ndhA	b	125	x		AT7 ndhA 5'gataatacgactcactatagg gac gag gtt gtc aat aat aga t3' Tm= 60°C	MunndhAex2for 5' tat cta aca gtt caa gta cag ttg ata tag 3' Tm= 78°C	800	AT7 ndhA	C9
68	ndhH	b	126		x	AT7 ndhH 5'gataatacgactcactataggg cat aaa gtt ctt gtt tcg aca a3' Tm= 58 °C	AndhH 5'ggt atg cga atg atg cat aa3' Tm= 56 °C	534	AT7 ndhH	D9
69	ycf1	b	128		x	AT7ycf1 5'gataatacgactcactatagg gtg aca ttt ccg ttc tta tgg3' Tm= 60 °C	Aycf1 5'tta gtc tgg ata cgg caa aat 3' Tm= 60 °C	672	AT7ycf1	E9
70	psbJoperon	b	129		x	AT7psbJ 5'gataatacgactcactataggg tac tag agg gat gaa ccc aat3' Tm= 60 °C	Aorf103 5'gtt cca cat ttt atc ttt cc3' Tm= 54 °C	1319		F9
71	ycf1.3’	b			x	AT7ycf1.3’ 5'gataatacgactcactatagg gtg aca ttt ccg ttc tta tgg3'	Aycf1.3’ 5'tta gtc tgg ata cgg caa aat 3'	544	AT7ycf1.3’	G9

						Tm= 60 °C	Tm= 58 °C			
72	ycf2.3’	a			x	AT7 ycf2.3’ 5'gataatacgactcactatagg gtt cta agc cat ctc tga cg3' Tm= 60 °C	y#6 5' cga tgt ggt aga aag c 3' Tm= 49,2 °C	744	AT7 ycf2.3’	H9
73	rps16	b	5	x		rps16revx 5' tta gga ttg att agg acg aag 3' Tm = 54 °C	Hrps16for 5' cga tgt ggt aga aag c 3' Tm=48 °C	237	rps16revx	A10
74	rps2	b	16		x	AT7 rps2 5' gta ata cga ctc act ata ggg att gag gat atg gcg tca t3' Tm = 70,5 °C	Arps2 5' gtc atg gta cta gga aat gg3' Tm = 55,3 °C	581	AT7 rps2	B10
75	rps14	b	34		x	AT7 rps14 new	Arps14 5' atg gca agg aaa agt ttg at 3' Tm=51,2 °C	304	AT7 rps14 new s.o.	C10
76	rps4	b	40		x	AT7 rps4 5' gta ata cga ctc act ata ggg ata ata ttc tac gac tag caa ttc3' Tm = 69,4 °C	Arps4 5' gac ctc gtt tca aaa aaa tac 3' Tm = 52 °C	574	AT7 rps4	D10
77	rpl33	a	69		x	AT7rpl33rev 5' gta ata cga ctc act ata ggg aca gat act caa atc tag a3' Tm = 54 °C	Arpl33for 5' ggg atc gaa ttg att ata ga 3' Tm = 54 °C	153	S.rpl33rev	E10
78	rps18	a	70		x	AT7 rps18 5' gat aat acg act cac tat agg gac acg ata ctc aaa tct aga3'	Arps18 5' ggg atc gaa ttg att ata ga 3'	290	rps18rev 5' aat att ttt cct tgt tca cta a3'	F10

						Tm = 69,4 °C	Tm = 51,2 °C		Tm = 54 °C	
79	rpl20	b	71		x	AT7 rpl20 5' gta ata cga ctc act ata ggg tcg ttc gaa atc ata taa ag3' Tm = 68,4 °C	Arpl20 5' ggg ata tat agc tcg gag ac 3' Tm = 57,3 °C	282	AT7 rpl20	G10
80	5'rps12	b	72	x		AT7 3'rps12 5' gat aat acg act cac tat agg gtt att ttg gct ttt tta ccc cat3' Tm = 64 °C	A5'rps12 (rps12for) 5' cct ctg gat ttg aaa tca ctg 3' Tm = 60 °C	140	AT7 3'rps12	
81	rps11	b	81		x	AT7 rps11 5' gat aat acg act cac tat agg gtc ttt ttt tcg gag gtc tac3' Tm = 70,4 °C	Arps11 5' tac cga aaa tta gtt cgc gt 3' Tm = 53,2 °C	395	AT7 rps11	A11
82	rpl36	b	82		x	AT7 rpl36 5' gat aat acg act cac tat agg gca agt ctg att atc cct gt3' Tm = 70,4 °C	Arpl36 5' gaa aat aag agc ttc cgt tc 3' Tm = 53,2 °C	122	AT7 rpl36	B11
83	rps8	b	84		x	AT7 rps8 5' gat aat acg act cac tat agg gcc aga tat aac aca aaa tct c 3' Tm = 69,4 °C	rps8rev 5' ggt tcg aat agc atc tac taa 3' Tm = 58 °C	313	AT7 rps8	C11
84	rpl14	b	85		x	AT7 rpl14 5' gat aat acg act cac tat agg gag cta atg aaa cta ttt tag3' Tm = 67,5 °C	Arpl14 5' tga ttc aac ctc aga ccc at3' Tm = 55,3°C	357	AT7 rpl14	D11
85	rpl16	b	86		x	AT7 rpl16	Arpl16	367	AT7 rpl16	E11

						5' gat aat acg act cac tat agg gtc gta tag gca ttt tag atg c3' Tm = 70,4 °C	5' gaa cca gat tcc gta aac aac3' Tm = 55,9°C			
86	rps3	b	87		x	AT7 rps3 5' gat aat acg act cac tat agg gat caa ttt tcg ctc gaa tcg t3' Tm = 70,4 °C	Arps3 5' gga caa aaa ata aat cca ctt g3' Tm = 52,8°C	578	AT7 rps3	F11
87	rpl22	b	88		x	AT7 rpl22 5' gat aat acg act cac tat agg gat gac agg tcg atc ttt tta t3' Tm = 69,4 °C	Arpl22 5' gtc tgc tga caa agc acg aa3' Tm = 57,3 °C	270	AT7 rpl22 s.o.	G11
88	rps19	b	89		x	AT7 rps19 5' gat aat acg act cac tat agg gag atc tat tat cgc ttt ttg3' Tm = 68,5 °C	Arps19 5' cct ttg tag cca atc att ta3' Tm = 51,2 °C	245	AT7 rps19	H11
89	rpl2	b	90	x		T7rpl2for 5' gat aat acg act cac tat agg gat gta gaa gat cat aac a3' Tm = 54°C	P18 5' tc aaa cgg acc tcc cca gaa gg c 3' Tm = 66 °C	284	p18	A12
90	rpl23	b	91		x	AT7 rpl23 5' gat aat acg act cac tat agg gaa gag gtg gaa tag aat aac c3' Tm = 70,4 °C	rpl23for 5' ata tga aca gta aga act 3' Tm = ? °C	257	AT7 rpl23	B12
91	rps7	b	100		x	AT7 rps7 5' gat aat acg act cac tat agg gaa aat gtg caa aag ctc tat3' Tm = 68,5 °C	Arps7 5' gtc acg tcg agg tac tgc ag3' Tm = 61,4 °C	460	AT7 rps7	C12

92	3'rps12	b	101	x		AT7 3'rps12 5' gat aat acg act cac tat agg gtc tgc agt acc tcg acg tga c3' Tm = 62 °C	A3'rps12 (rps12in2for) 5' cct ctg gat ttg aaa tca ctg 3' Tm = 60 °C	735		D12
93	rrn16	a	105		x	AT7 rrn16 5' gat aat acg act cac tat agg gtt cca aac tca acg atg g 3' Tm = 70,5 °C	Arrn16 5' ggt atc tgg gga ata agc at 3' Tm = 55,3 °C	665	AT7 rrn16	E12
94	rpl32	a	116		x	AT7 rpl32 5' gta ata cga ctc act ata ggg t tca agt tga ttc aaa tta ttc g 3' Tm = 52 °C	rpl32for 5' atg gca gtt cca aaa aaa cg 3' Tm = 56 °C	197	rpl32rev 5' cgt att cgt aaa aat att tgg 3'	F12
95	rps15	b	127		x	AT7 rps15 5' gta ata cga ctc act ata ggg ttt tgt ctc tcg aat gtc caa 3' Tm = 60 °C	Arps15 5' gga tct ctt gaa ttt caa gta tt 3' Tm = 60 °C	206	AT7 rps15 s.o.	G12
1	trnH-gug	b	1			AT7trnH-gug 5' gta ata cga ctc act ata ggg cga acg acg gga att gaa 3' Tm = 71,6 °C	AtrnH-gug 5' aga act tgt ttc tct tct tg 3' Tm = 51,2 °C	295	AT7trnH-gug	AA1
2	trnK-uuu	b	4	x		trnKx2rev 5' tgg gtt gcc cgg gat tcg aa 3' Tm = 61,4 °C	trnKex1f 5' ggt tgc taa ctc aac ggt a 3' Tm = ? °C	71	trnKx2rev	BB1
3	trnQ-uug	b	6		x	AT7trnQ-uug 5' gta ata cga ctc act ata ggg atc aga aac aac aat cac att c 3' Tm = 69,4 °C	AtrnQ-uug 5' tcg caa aat cct agt cca at 3' Tm = 53,2 °C	274	trnQ-UUGrevx 5' ctg gga cgg aag gat tcg aac c 3' Tm = 64 °C	CC1

4	trnS-gcu	b	9			AT7trnS-gcu 5' gta ata cga ctc act ata ggg aaa aaa tca agt cat caa cgg 3' Tm = 69,4 °C	AtrnS-gcu 5' atg gga gag atg gct gag tg 3' Tm = 59,4°C	108	trnSgcurev	DD1
5	trnG-ucc	a	10	x		AT7trnG-ucc 5' gta ata cga ctc act ata ggg att gta gcg ggt ata gtt ta 3' Tm = 69,4 °C	trnGuccrev 5' gga atc gaa ccc gca tcg t 3'	61	trnGuccrev	EE1
6	trnR-ucu	a	11			AT7-trnR-ucu 5' gta ata cga ctc act ata ggg aaa tat tga tca ctt ttg tct ta 3' Tm = 69,4 °C	AtrnR-ucu 5' gcg tcc att gtc taa tgg at 3' Tm = 55,3 °C	193	trnRucurevx 5' atc cat tag aca atg gac gc3' Tm = 58 °C	FF1
7	trnC-gca	a	20		x	AT7trnC-gca 5' gta ata cga ctc act ata ggg agt ttt ttg ttg atc agg cg 3' Tm = 60°C	AtrnC-gca 5' ttt gta gca ttt tgg cga ca 3' Tm = 53,2 °C	100	trnC-GCArevx 5' agg cga cac ccg gat ttg aac 3' Tm = 61,8 °C	GG1
8	trnD-guc	b	23		x	AT7trnD-guc 5' gta ata cga ctc act ata ggg ata tat tta ttg gat cca tcg g 3' Tm = 69,4 °C	AtrnD-guc 5' att tcc ctg gga ttg tag tt 3' Tm = 53,2 °C	101	trnD-GUCrevx 5' cgg gac tga cgg ggc tcg aac 3' Tm = 67,6 °C	HH1
9	trnY-gua	b	24			AT7-trnY-gua 5' gta ata cga ctc act ata ggg att ttt ggg ccg agc tgg att 3' Tm = 72,4 °C	AtrnY-gua 5' aat tga ttc ttc ttg ggt cga 3' Tm = 54 °C	102	AT7-trnY-gua	AA2
10	trnE-uuc	b	25			S.AT7trnE-uuc 5' gta ata cga ctc act ata ggg tca att tta gac tta ttg gta atc 3'	AtrnE-uuc 5' ttg gtc aag cag gcc ccc ta 3'	104	AT7trnE-uuc 5' gta ata cga ctc act ata ggg tac ccc cag ggg aat tcg aat 3'	BB2

						Tm = 69,4 °C	Tm = 61,4 °C		Tm = 73,3 °C	
	trnE,trnY, trnD Operon	b				trnD-GUCrevx 5' gta ata cga ctc act ata ggg gat ata ttt att gga tcc atc gg 3' Tm = 62 °C	AtrnE-uuc 5' gga att aaa tca aaa agg ccc t 3' Tm = 54,7 °C			
11	trnT-ggu	a	26		x	AT7trnT-ggu 5' gta ata cga ctc act ata ggg agt ttt aca aag ccc ctt at 3' Tm = 69,4 °C	AtrnT-ggu 5' gga att aaa tca aaa agg ccc t 3' Tm = 60 °C	100	trnT-GGUrevx 5' gcc ctt tta act cag tgg tag 3' Tm = 57,9 °C	CC2
12	trnS-uga	b	29		x	AT7trnS-uga 5' gta ata cga ctc act ata ggg ctt ggc tag gtg gga tag c 3' Tm = 73,6 °C	AtrnS-uga 5' gtc ttt cgg aga gat ggc tg 3' Tm = 59,4 °C	194	S.trnS-uga rev 5' agg aga gag agg gat tcg a 3' Tm = 56,7 °C	DD2
13	trnG-gcc	a	32		x	trnG revT7 5' gta ata cga ctc act ata ggg tat att taa tgc tat tta tg 3' Tm = 52 °C	AtrnG-gcc 5' aca aaa aaa tgc gga tat gg 3' Tm = 51,2 °C	131	HtrnGrev	EE2
14	trnfM-cau	b	33		x	AT7 trnfM-cau 5' gta ata cga ctc act ata ggg aga tgt tgc gga gac agg att 3' Tm = 62 °C	AtrnfM-cau 5' gtt ttc agt tct tat ctt tcg 3' Tm = 56 °C	99	AT7 trnfM-cau	FF2
15	trnS-gga	a	39		x	AT7trnS-gga 5' gta ata cga ctc act ata ggg att aag cga agg tac gga aag 3' Tm = 71,4 °C	AtrnS-gga 5' ggc gta gca ttg gaa ctg ct 3' Tm = 59,4 °C	77	trnS-GGArevx 5' cgg aaa gag agg gat tcg aac 3' Tm = 59,8 °C	GG2
16	trnT-ugu	b	41		x	AT7trnT-ugu	AtrnT-ugu	202	AT7trnT-ugu	HH2

						5' gta ata cga ctc act ata ggg cta tcg gaa tcg aac cga t 3' Tm = 71,5 °C	5' aaa tcg aaa ttc tat atc gc 3' Tm = 49,1 °C			
17	trnL-uaa	a	43	x		AT7trnL-uaa 5' gta ata cga ctc act ata ggg ata gag gga ctt gaa ccc t 3' Tm = 71,5 °C	AtrnL-uaa 5' ggg ata tgg cga aat cgg ta 3' Tm = 57,3 °C	101	AT7trnL-uaa s.o.	AA3
18	trnF-gaa	a	44		x	AT7trnF-gaa 5' gta ata cga ctc act ata ggg cat aca aat taa tca tgt gcc ag 3' Tm = 70,4 °C	AtrnF-gaa 5' ggg ata gct cag ttg gta ga 3' Tm = 57,3 °C	86	HtrnF rev 5' tgc cag gaa cca gat ttg 3'	BB3
19	trnV-uac	b	48	x		AT7trnV-uac 5' gat aat acg act cac tat agg gta tag ggc tat acg gac tcg a 3' Tm = 72,3 °C	AtrnV-uac 5' gca cta agg gct ata gct ca 3' Tm = 57,3 °C	73	AT7trnV-uac	CC3
20	trnM-cau	a	49		x	AT7trnM-cau 5' gat aat acg act cac tat agg gtg aat aca atc aaa tac aa 3' Tm = 66,4 °C	trnM.P1 5' gct ttc ata cgg cgg gag c 3' Tm = 62 °C	199	trnM-CAUrevx 5' tac cta cta ttg gat ttg aac 3' Tm = 52 °C	DD3
21	trnW-cca	b	66		x	AT7trnW-cca 5' gta ata cga ctc act ata ggg cac gct ctg tag gat ttg aac 3' Tm = 72,4 °C	AtrnW-cca 5' gct att cta gta tat gca tag 3' Tm = 56 °C	157	AT7trnW-cca	EE3
22	trnP-ugg	b	67		x	AT7trnP-ugg 5' gat aat acg act cac tat agg gat gac agg att tga acc t 3' Tm = 69,5 °C	AtrnP-ugg 5' att gat ttt tac gac ggg tt 3' Tm = 51,2 °C	296	AT7trnP-ugg	FF3

23	j inf A	b	83		x	AT7j inf A 5' gat aat acg act cac tat agg gta cct taa ttc tga atc tat ttc 3' Tm = 68,5°C	Aj inf A 5' gga gtc atg aag ctt taa tt 3' Tm = 51,2 °C	306	AT7j inf A	GG3
24	trnI-cau	b	92		x	AT7trnI-cau 5' gat aat acg act cac tat agg gcg tgc atc cag tag gaa tt 3' Tm = 71,4 °C	AtrnI-cau 5' gga tcc ccg cta agc atc ca 3' Tm = 61,4 °C	90	AT7trnI-cau	HH3
25	trnL-caa	b	97		x	AT7trnL-caa 5' gat aat acg act cac tat agg gta tta tgc ctt gaa gag gac t 3' Tm = 70,4 °C	AtrnL-caa 5' tgg tga aat ggt aga cac gcg 3' Tm = 64 °C	81	AT7trnL-caa	AA4
26	trnV-gac	a	104		x	AT7trnV-gac 5' gat aat acg act cac tat agg gct tag gga taa tca ggc tc 3' Tm = 71,4 °C	AtrnV-gac 5' ggg aag gga tat aac tca gc 3' Tm = 57,3 °C	79	AT7trnV-gac	BB4
27	trnI-gau	a	106	x		AT7trnI-gau 5' gat aat acg act cac tat agg gcc atc ctg gac ttg aac c 3' Tm = 72,5 °C	AtrnI-gau 5' ggg cta tta gct cag tgg ta 3' Tm = 57,3 °C	58	AT7trnI-gau	CC4
28	trnA-ugc	a	107	x		AT7trnA-ugc 5' gat aat acg act cac tat agg gac tcg aac cgc tga cat cc 3' Tm = 73,4 °C	AtrnA-ugc 5' ggg ata tag ctc agt tgg ta 3' Tm = 55,3 °C	50	AT7trnA-ugc	DD4
29	rrn 23	a	108		x	AT7rrn 23 5' gat aat acg act cac tat agg gtt ccc tta acc aag cca c 3' Tm = 71,5 °C	Arrn 23 5' ggg agt tga aaa taa gca ta 3' Tm = 51,2 °C	547	AT7rrn 23	EE4

30	rrn 4.5	a	109		x	AT7rrn 4.5 5' gat aat acg act cac tat agg gct tcc atc cct tgg ata gat a 3' Tm = 71,3 °C	Arrn 4.5 5' gag ccg ttt atc att acg at 3' Tm = 53,2 °C	243	S.rrn4,5rev	FF4
31	rrn 5	a	110		x	AT7rrn 5 5' gat aat acg act cac tat agg gct ggc gtc gag cta ttt ttc c 3' Tm = 73,3 °C	Arrn 5 5' ttc gat att ctg gtg tcc ta 3' Tm = 53,2 °C	122	AT7rrn 5	GG4
32	trnR-acg	a	111		x	AT7trnR-acg 5' gat aat acg act cac tat agg gtc tca gat ata cta gca ctg 3' Tm = 70,4 °C	AtrnR-acg 5' gga tta gag cac gtg gct acg 3' Tm = 61,8 °C	418	trnR-acgrevx 5' tgg gcg agg agg gat tcg aac cc3' Tm = 74 °C	HH4
33	trnN-guu	b	112		x	AT7trnN-guu 5' gat aat acg act cac tat agg gaa gta gga ttc gaa cct acg a 3' Tm = 64 °C	AtrnN-guu 5' tcc tca gta gct cag tgg tag a 3' Tm = 66°C	173	AT7trnN-guu	AA5
34	psbJ		114			AT7psbJ 5' gat aat acg act cac tat agg gta cta gag gga tga acc caa t 3' Tm = 62 °C	ApsbJ 5' act gga agg att cct ctt tgg 3' Tm = 62 °C	112		BB5
35	spr A	a	117		x	AT7spr A 5' gat aat acg act cac tat agg gtg gaa ata gaa tgg aaa tat aga 3' Tm = 62 °C	Aspr A 5' gat ttc cct ttc ata ttt gat 3' Tm = 52 °C	242	sprA 5' tgg aaa tag aat gga aat ata ga3' Tm = 58 °C	CC5
36	trnL-uag	a	118		x	AT7trnL-uag	AtrnL-uag(Zagreb3)	171	trnLUAG	DD5

						5' gat aat acg act cac tat agg gaa atc tta aaa aag aag agt ccc 3' Tm = 64 °C	5' ggt gaa att ggt aga cac gc 3' Tm = 60 °C		5'ccg cca ctc gga ctc gaa c3' Tm = 64 °C	

Oligonucleotides used for ycf2 transcript analysis (not inclueded in Table above):

	ycf2a	AT7ycf2a 5' gat aat acg act cac tat agg gcaagaagaaaacctttaggc3' Tm = 56 °C	#8 5' gaacaaccgggagcaatt 3' Tm = 54 °C	355
	ycf2b	AT7ycf2b 5' gat aat acg act cac tat agg gttttttttaccgacgcac3' Tm = 54 °C	#5 5' tttgataacccggattccta3' Tm = 60 °C	286
	ycf2.2.3'	AT7ycf2.2.3' 5'gat aat acg act cac tat a gggtccggataaagaccaaag3' Tm = 64 °C	ycf2.2.3' 5' atcgatattgatgatagtgac3' Tm = 56 °C	800
	ycf2.3'	AT7ycf2.3' 5' gat aat acg act cac tat agg ggttctaagccatctctgacg3' Tm = 64 °C	#6 5' gatcacttcttatggactc3' Tm = 54 °C	746
		ycf15-5'for 5' agtggaacagaattgactgg3' Tm = 58 °C	ycf15for2 5' tcttagatcaaaacactatg3' Tm = 52 °C	-
		PM13 5' 3' Tm = 64 °C	Tm = 60 °C	

Oligonucleotides used for rpoB characterisation:

PCR mutant	cpoBleft 5' gatccaaaacttgagataatgg3' Tm = 56 °C	aadAoutfor 5' tatcagcccgtcatacttgaagc3' Tm = 54 °C
Primer extension	rpoBpe 5' tgtatctgattaaatccaggtattg3'	
	rpoBpe2 5' cttagacttattaattaagttata3'	

Danksagung

Diese Arbeit wurde am Lehrstuhl I für Botanik, Department für Biologie der Ludwig-Maximilian Universität München in der Arbeitsgruppe von Herrn Prof. Dr. Reinhold G. Herrmann durchgeführt. Für die Möglichkeit, ein Teil der Arbeitsgruppe zu sein bedanke ich mich gerne. Ein kleines Dankeschön auch für die Zeit, die er mir gegeben hat, um mir überlegen zu könen, weswegen, warum und wie für meine Arbeiten nötig waren.

Den Prof. Dr. Hans Ulrich Koop danke ich für Bereitschaft, das Zweitgutachten zu übernähmen.

Herrn Dr. Boris Hedtke (Berlin) hatte die Idee die RpoT-Gene des Tabaks zu isolieren "ansteckend" weiter gegeben. Sein aktieves Arbeiten, Mitfiebern und sein immenser Enthusiasmus, haben mir geholfen, das Projekt zu beenden. Das hat mir auch ermöglicht Berlin kennenzulernen. Herrn Prof. Dr. Thomas Börner danke ich für seine grose unterstützung.

Mein besonderes Dank geht an Herrn PD Dr. Rainer M. Maier der mich motiviert hat, mit ihm weiter zu arbeiten, nachdem mein erster Wissenschaflicher Betreuer sich verabschiedet hatte. Ich bewundere sein Nervenkostüm und Fähigkeiten um mit grosse Ruhe mit allen Problemen umzugehen zu können, selbst so einen wie ich es binn.

Dr. Christian Schmitz-Lin(n)eweber danke ich für seine Hilfe bei der Feiabformung der Arbeit. Viele Ideen, manche Musikrichtungen und sein Lachen haben mir geholfen viele "Tiefs" zu überstehen und viele "Hochs" besser geniessen zu können. Danke für die Möglichkeit, die Breite des Sprachfeldes die ich während der Zeit der Doktorarbeit gelernt habe zu nutzen und die Möglichkeit sich richtig austoben zu können ohne ermahnt zu werden, und das ist nicht alles....

Ein Dank gilt Michael Tillichm der mit Mocca, Espresso and Co. angefangen hat, meine Morgenstunden zu verbessern. Das hat dazu geführt, dass ich ein Kaffe-"junckie" geworden bin. Danke für seine interessen und Ideen, die er durch seine „Ge'ma eine Rauchen“ mit mir getauscht hat.

Sabine Kemp hat im wahrsten Sinne des WORTES MIR GEZEIGT WAS DEUTSCHE Ordnung bedeutet. Ich habe auf diesen Gebiet sicherlich noch viel zu lernen. Ein Dankeschön für die unterstützung die sie während der Makroarray-Technik "Entwicklung" geleistet hat, für ihre Fragen, die ich schwer beantworten konnte (oder gar nicht), und für ihre super logische Art und Weise, wie auch schwierigste Probleme angegangen werden können.

Danke auch Anina Neumann, die mich in der Charakterisierung der Promotormutanten unterstützt hat, Martina Reymers, die mir bei der Kultivierung des transplastomischen materials von rpo- Pflanzen und der Cytochrommutanten geholfen hat,

Claudia Nickel die micht in die Protein-biochemie eingewiesen hat und Irmi Reiber, die sich die ganze Zeit um meine Pflanzen im Gewächshaus gekümmert hat; Danke auch an alle Mitglieder des Lehrstuhls "upstairs and downstairs" die für eine nette Atmosphäre während der Arbeit gesorgt haben: Peter Poltnigg, Helena Funk, Holger Hupfer, Anja Drescher, Lina Lehsneeva, Cristina Dal Bosco, Birgit Profanter, Moni Wimmer, Martin Lentsch, Petra Weber, Magdalena Swianteck, Dr. Anna Sokolenko, Giusy Canino, Gabi Burkhard, Ralph Regel, Giusy Canino, Lena Pojidaeva, Elli Gericke, Jarda Mraceck, Gisela Nagy, Martina Silber, Manfredo Gödel, Katrin Ahmann.

Hvala mojoj majci koja me je, neumorno radeći podržavala cijelo vrijeme za vrijeme mog studija kako bi ja ostvarila moje želje. Hvala ti majko, za sve, i za ono šta si nekada pokušavala svojom snagom pokazati mi da si mi ujedno **i otac i majka**. Hvala ti.

Danke an Reinhard Reitmayer, der mich unterstüzte, meine arbeiten abzuschliessen, und der es mit seiner sensiblen Art verstand, Hilfe zu geben.

Hvala cijeloj mojoj obitelji koja je uvjek u želji da učini nešto dobro za mene davala mi ono najbitnije, a to je osijećaj da nisam sama i da se njima kako god se ja osijećala uvjek mogu obratiti.

Curriculum vitae

Personal

Home Address:
Langerhansstr. 14
80999 München
(089) 8189-5091 (Home)
(089) 17861-183 (work)
e-mail: julia@botanik.biologie.uni-muenchen.de

Date and Place of Birth
August 14th 1971, Wolfratshausen, Germany

Education

Abitur 1990: Secondary Scool in Novi Sad, Vojvodina, Yugoslavia.

Diploma Sudies: 1990-1996

1990-1996 at the University for Food Science and Biochemical Industry in Zagreb, Croatia.

Major: Biochemical Engineering

Theses: The toxical impact of herbicide Dikuran In prokaryotes (*Salmonella tiphymurium*) and eukaryotes (*Saccharomices cerevisiae*)

Ph. D. studies: 1997-2002

Botanisches Institut, Ludwig-Maximilians Universität, München.

Thesis: Gene expression in plastids of higher plants: evolutionary and functional aspects of different RNA polymerases – coordinated assembly of multiprotein-complexes.

EXTRAMURAL ACTIVITIES

Research sojourns in the laboratory of Prof. Dr. Thomas Börner at the Humbold University in Berlin, Germany: Identification and molecular characterisation of nuclear-encoded T3/T7 phage-like polymerases in tobacco (June-September 1998, March 1999, September 1999, April 2000, September 2000).

Short research sojourn in the laboratory of Prof. Igloi Gabor at the Albert-Ludwigs University in Freiburg, Germany: technical approach in isolation of tRNA fractions from plastids of higher plants (July 2002).

ORAL/POSTER PRESENTATIONS

Meeting of the German Section of the International Society for Endocytobiology (ISE-G) in Kohren-Salis, Germany.Transcription of the plastid genome in PEP-lacking material. 25-27.09. 2000. Oral Pres.

FEBS Advanced course on Origin and Evolution of Mitochondria and Chloroplasts. Julia Legen, Sabine Kemp, Reinhold G. Herrmann and Rainer M. Maier: Comparative analysis of plastid transcription profils attributed to differential activity of the plastid-encoded (PEP) and nuclear-encoded (NEP) RNA polymerases. Poster.

Publications List

Börner T., Hedtke B., Hess W.,R., Legen JU., Herrmann R.G., and Weihe A. (1999) Phage-type RNA polymerases in higher plants. In: The chloroplast from Molecular Biology to Biotechnology (Argyroudi J.H. and Senger H., eds.) Dordrecht, Germany: Kluwer Academics Publisher: pp. 73-78.

Legen J., Misera S., Herrmann R.G., Meurer J. (2001) Map position of 69 Arabidopsis thaliana genes of all known nuclear encoded cinstituent polypeptides and various regulatory factors of the photosynthetic membrane: a case study.; DNA Res., 8(2):53-60.

Legen J., Schmitz-Linneweber C., Drescher A., Hupfer H., Tillich M., Herrmann R.G., and Maier R.M. (2001) Decoding of the ndhH operon from spinach: An example for the complexity of plastid gene exprssion in higher plants. Endocytobiosis and Cell Res., 14, 11-20.

Hedtke B., Legen J., Weihe A., Herrmann, R.G., Börner T.(2002) Six active phage-type RNA polymerase genes in Nicotiana tabacum.; Plant J., 30(6): 625-37.

Legen J., Kemp S., Krause K., Profanter B., Herrmann, R.G., and Maier R.M. (2002) Comparative analysis of plastid transcription profiles of entire plastid chromosomes from tobacco attributed to wild-type and PEP-deficient transcription machineries.; Plant J. 31(2): 171-188.