The molecular architecture of translational regulatory elements within the 5' untranslated region of the chloroplast *psbC* mRNA in *Chlamydomonas reinhardtii*

Mir Munir Ahmed Rahim

A thesis

In

The department

Of

Biology

Presented in Partial fulfillment of the Requirements

For the Degree of Master of Science in Biology at

Concordia University

Montreal, Quebec, Canada

May 2003

Your file Votre référence
ISBN: 0-612-83850-1
Our file Notre référence
ISBN: 0-612-83850-1

ABSTRACT

The molecular architecture of translational regulatory elements within the 5' untranslated region of the chloroplast *psbC* mRNA in *Chlamydomonas reinhardtii*

Mir Munir Ahmed Rahim

Translation of specific mRNAs of the semiautonomous genetic systems of mitochondria and plastids is controlled by the products of nuclear genes. Previous studies have characterized a particularly complex example of this control pertaining to translation of the chloroplast *psbC* mRNA in the unicellular eukaryotic alga *Chlamydomonas reinhardtii*. *psbC* translation is controlled by interactions between its 547 nucleotides 5' untranslated region and the product of the nuclear loci TBC1, TBC2 and TBC3.

This thesis extends our understanding of *psbC* translational control from the genetic level to the biochemical level. RNA structures involved in translation and its control by the nuclear gene products were identified using chemical and enzymatic probing experiments, oligonucleotide hybridization probing experiments, circular dichroism, and reverse transcription inhibition experiments. Results strongly support a large stem-loop secondary structure that was proposed over 13 years ago. Additional evidence indicates this structure forms a pseudoknot structure by hybridization of its apical loop with a proximal sequence.

A mutation that removes two bulges from the stem of this structure drastically alters its structure. Evidence supporting 6 smaller stem-loop structures was obtained.

Sequence elements and RNA structures in the *psbC* 5' untranslated region that interact with factors encoded or controlled by TBC1, TBC2 and TBC3 were identified by *in vivo* dimethyl sulfate protection analysis. Control by TBC1 and TBC3 is mediated by specific sequences that are dispersed throughout the 5' untranslated region. The TBC2 mutant has a DMS protection within the coding sequence. Models of translational control based on these findings are proposed.

Acknowledgements

I express my deepest gratitude to my supervisor, Dr. William Zerges for his guidance and advice throughout the course of this investigation. I thank my committee members, Dr. M. Herrington and Dr. J. Turnbull for their timely advice. I also thank Dr. P. Albert for helping with photography. I extend my thanks to my lab-mates, Frederick, Dana and Hamid, and all my friends in the department for their help and moral support. My special thanks go to my family for their love and encouragement.

List of Figures

List of Tables

1.1. Chloroplasts

Chloroplasts are the developmental fate of the plastids that are specialized in photosynthesis. They are also involved in the synthesis of amino acids, lipids and pigments, and in the assimilation of nitrogen, sulfur and phosphorus. Chloroplasts have evolved from an endosymbiosis of an aerobic eukaryotic and an oxygen-evolving photosynthetic prokaryote of the cyanobacterial lineage (reviewed by Gray, 1992; Bhattacharya, 1998). From these combined cells have descended the modern eukaryotic algae and plants.

Chloroplasts proliferate by binary division of existing plastids and are transmitted to daughter cells during cell division. All known plastids have semiautonomous genetic systems, which have remnants of the genetic system of their ancestral prokaryote. As such, plastid genomes and gene expression machinery are quite different from the nuclear genome and genetic system of the nuclear-cytosolic compartments (reviewed by Rochaix,1996). There has been a selective pressure to maintain some of the ancestral eubacterial genes in the chloroplast during the evolution (reviewed by Zerges, 2002). Chloroplasts contain multiple copies of relatively small genomes. For example, land plant chloroplast genomes vary form 120 kb to 160 kb and in only few cases reaching 220 kb (reviewed by Palmer, 1990). Most endosymbiont genes have been either lost or transferred to the nucleus during the evolution (Gillham *et al.*, 1994). Indeed as much as 13% of the *Arabidopsis* nuclear genome may have been derived from

the cyanobacterial genome of the endosymbiont (Martin *et al.*, 2002). Phylogenetic reconstruction of changes in plastid genome content revealed an accelerated rate of gene loss in *Chlamydomonas/Chlorella* lineage, relative to other eukaryotic photosynthetic organisms (Maul *et al.*, 2002).

Model organisms used for photosynthesis research include land plants (tobacco, maize, spinach, barley), green algae (*Chlamydomonas spp.*), euglenoid algae (*Euglena spp*) and cyanobacteria. Spinach, tobacco, maize and barley allow isolation of large quantities of chloroplasts for *in vitro* and *in organellar* studies. *Arabidopsis thaliana* and maize are amenable to genetic approaches. However, genetic crosses require considerable resources and time, especially for Maize. Of the vascular plants, only the chloroplast genome of tobacco has been transformed with exogenous DNA, and this is hampered by the requirement to transform all copies of the chloroplast genome in the approximately 100 chloroplasts of every cell of the plant (Svab and Maliga, 1993).Cyanobacteria are advantageous for many *in* vitro and in *vivo* studies, however isolation of membranes and multisubunit protein complexes therein are difficult.

The eukaryotic green alga *Chlamydomonas* has been widely used for diverse studies of chloroplast biology and photosynthesis because it is amenable to molecular biological, biochemical, genetic and physiological approaches (reviewed by Mets and Rochaix, 1998 and Harris, 1998). Most studies have used *Chlamydomons reinhardtii. C. reinhardtii* nucleus is partially surrounded by a cup-shaped chloroplast, which occupies most of the cytosol. A prominent pyreniod is situated within the chloroplast distal to the nucleus and specializes in starch

biosynthesis. Vegetative cells are typically haploid and reproduce by mitosis. The sexual cycle is a response to nitrogen starvation, an indicator of adverse conditions. Upon nitrogen deprivation, vegetative cells differentiate to gametes, in a complex developmental process that involved drastic changes in gene expression and physiology. Gametes of opposite mating-type, *mt-* and *mt+*, adhere via recognition proteins on their flagella and then fuse to form the diploid zygote. Zygotes can either become stable diploid vegetative cells or undergo meiosis to form four haploid spores, contained within an ascus. Spores germinate to form vegetative progeny. Thus, tetrad analysis is routinely used for genetic mapping experiments.

Nuclear genes are inherited in a Mendelian fashion while chloroplast genes are inherited uniparentally from the *mt+* parent. *C. reinhardtii* has a short life cycle and is easily manipulated in the laboratory, maintained on minimal nutrients and grows rapidly. Unlike vascular plants, photosynthesis deficient mutants of *C. reinhardii* are viable on media containing a fixed carbon source (Harris, 1989). *C. reinhardtii* grows heterotrophically on fixed carbon source such as acetate without photosynthetic activity.

The genetics of *C. reinhardtii* is one of the most advanced of any model organism. Uniparental inheritance and the basis thereof were identified by Dr. Ruth Sager in her classic work of the 1950-1970s (see attached paper for additional reference, Sanger and Ramanis, 1968). The first genetic maps for the nuclear genome of *C. reinhardtii* were published by Levine and Goodenough in 1970. Estimates of the DNA content of the haploid genome range from $1.0\text{-}1.6\text{x}10^5$ kbp

(Harris, 1989). The genetic map consists of 17 linkage groups on which nearly 200 molecular markers have been mapped. The genome has a high GC content (62%) and contains numerous classes of repetitive sequences (reviewed by Silflow, 1998).

The chloroplast genome of *C. reinhardtii* has recently been sequenced completely (Maul *et al.*, 2002). It is circular, composed of 203,395 bp, and contains 99 genes. The chloroplast genome has 34.6% GC content. This genome is gene-poor compared to the sequenced chloroplast genomes of other plants and photosynthetic algae; more than 20% is composed of repetitive sequences (Maul *et al.*, 2002). Like most plastid genomes, the *C. rienhardtii* chloroplast genome possesses two copies of an inverted repeat sequence of 2,221 bp which are separated by two nearly equal sized unique regions of 80,873 bp and 78,100 bp. Genes carried on *C. reinhardtii* chloroplast genome include genes for photosynthesis, transcription, ribosomal proteins, tRNAs, rRNAs, protease and some genes of unknown function.

A powerful collection of genetic, molecular, biochemical and microbiological techniques have been developed for *Chlamydomonas* research over the past century (Harris, 1989; Harris, 2001; Rochaix, 1998). Transformation of all three genomes: in the nucleus, chloroplast and mitochondria is readily achieved with techniques such as particle bombardment, electroporation and by agitation of cells in presence of DNA and glass beads (reviewed by Kindle, 1998 and Goldschmidt-Clermont, 1998). Nearly all integration events of transforming DNA result from random integration into chromosomal DNA, with less than 1% occurring by

homologous recombination (Sodeinde, 1993). However, in the chloroplast transforming DNA is usually integrated by homologous recombination and only in exceptional cases is it maintained extrachromosomally (reviewed by Heifetz, 2000)

1.2. RNA Structural and Function

RNAs form specific structures which are necessary for their functions. RNA molecules form stable helical secondary structure through base paring. These can be further packed together or associated by tertiary interactions giving the RNA an irregular shape. Molecular interactions in RNA that form stable structures include hydrogen bonding between irregular complementary surfaces as in tRNA helices, helical stacking of P4-P6 helices in *Tetrahymena* group 1 intron, and hydrogen bonds involving 2' hydroxyl group on ribose sugar (Murphy *et al.,* 1994; reviewed by Draper 1996; reviewed by Stobel and Doudna, 1997). Monovalent and divalent ions, and in particular Mg^{2+} ions, bind to RNAs to stabilize both secondary and tertiary structures. Diffusely bound Mg^{2+} ions at non-specific sites stabilize secondary structures, in part by neutralizing electrostatic repulsion between phosphate groups on opposite strands of a helix (Draper and Misra, 1998; Misra and Draper, 1998). A second class of Mg^{2+} ions bind to specific sites on an RNA to stabilize specific tertiary structures (Misra and Draper, 1998). In some cases Mg^{2+} binding can destabilize an RNA structure (Misra and Draper, 1998).

Several RNA structural motifs have been identified (Conn *et al.*, 1998). A pseudoknot forms by base pairing of a sequence within the loop of a stem-loop secondary structure to a complementary sequence located near the base of the

stem, forming a second stem (Conn *et al.*, 1998). Pseudoknot structures in the 5' UTRs of mRNAs have been shown to regulate translation in both prokaryotes and eukaryotes. Another structural motif was identified from the 3 dimensional structure of the P4-P6 domain of the self-splicing group1 intron in *Tetrahymena* (Costa and Michel, 1995). A short stem with a four base loop, called a "tetra loop" having the consensus GNRA (N is any ribonucleotide and R is a pyrimidine) binds a receptor consisting of three consecutive As that protrude from the stem of a second stem-loop structure, called an "A platform" (Costa and Michel, 1995).

RNA structure is difficult to determine at the atomic level because RNAs do not readily crystallize for X-ray diffraction analysis (reviewed by Draper and Misra 1995). The structure of tRNAs were the first to be deciphered and they are known at atomic resolution. tRNA structures melt with three melting transitions corresponding to the three stems (Hilbers *et al.*, 1976). The tertiary interactions are usually less stable than secondary structures but in presence of sufficient concentration of Mg^{2+} tertiary interactions become one of the most stable components of $tRNA^{fmet}$ structure (Stein and Crothers, 1976). Other parts of the molecule cannot melt independently when the tertiary structure is intact (Stein and Crothers, 1976). Mg^{2+} plays an indispensable role in the higher order assembly of large RNAs such as *Tetrahymena* group 1 intron, for which structural details are available at atomic resolution (reviewed by Doudna and Doherty, 1997). Large RNAs form structures that involve extensive helical packaging. UV cross-linking experiments have shown that helices P2 and P2.1 of the group 1 self-splicing intron in *Tetrahymena* are coaxially stacked involving sequences within the core

structure. These sequences are necessary for catalytic activity and that cross-linking and catalysis require a common tertiary structure (Downs *et al.*, 1990). The helices P4 and P6, part of the active core of the ribozyme, fold independently of other domains and are also coaxially staked in the native molecule (Murphy *et al.*, 1994). The remaining structures in the catalytic core (P3-P9) retain their secondary structure but long-range tertiary interactions and complete folding is achieved only in presence of P4-P6 domain (Doherty and Doudna, 1997).

The precise characterization of RNA structure is important for understanding RNA function. Several approaches used for RNA structural analysis include covariance analyses of homologous RNAs, site-specific mutations, thermodynamic parameters (using computer programs), chemical and enzymatic probing of secondary structures, UV induced RNA-RNA cross-linking, oligonucleotide hybridization to single-stranded regions, temperature dependent melting, circular dichroism, NMR and X-ray crystallography for high resolution structure analysis (reviews by Ehresmann *et al.*, 1987, Krol and Carbon, 1989, Sosnick *et al.*, 2002, Branch *et al.*, 1989). Computer predictions are based on energy minimization to construct secondary structure models of a given RNA sequence. One such program commonly used is *mfold* (Zuker, 1994). However, pseudoknot and tertiary interactions are difficult to predict by energy-based models (Lyngso and Pedersen, 2000). Chemical and enzymatic probes can provide accurate information on RNA structure and their interactions with protein factors. (reviewed by Ehresmann *et al.*, 1987). Secondary and some tertiary interactions in RNA can also be identified by physical methods such as UV induced RNA-RNA

cross-linking, temperature dependent RNA melting and circular dichroism couple with temperature melting (reviewed by Branch *et al.*, 1989, Draper *et al.*, 1996, Sosnick *et al.*, 2000). Melting of RNA is accompanied by increase in absorbance of 260nm light, a phenomenon called hyperchromicity. Tertiary interactions are weaker and melt before secondary structures (reviewed by Draper *et al.*, 1996). Although these methods yield valuable information about the tertiary interactions, RNA structure stability and folding pattern, large RNAs unfold in a complex fashion and resolving their structure becomes difficult. NMR and X-ray crystallography provide structural details at atomic resolution, but their use has been limited due to requirement for large amounts of highly purified RNA. Only three classes of RNA structures are known in atomic detail. These are tRNA, *Tetrahymena* group 1 intron and hammer-head ribozyme (reviewed by Doudna and Doherty, 1997).

Computer predictions based on free energy comparison coupled with chemical and enzymatic probing could yield valuable data for RNA structure analysis. Chemical probe such as dimethyl sulfate (DMS) specifically methylates guanine (G) at N7 position, and cytosine (C) and adenine (A) at positions N3 and N1 respectively, when they are not involved in Watson-Crick base-pairing (reviewed by Ehresmann *et al.*, 1987). Modified As and Cs in the DMS treated RNA molecule can be detected in ^{32}P end-labeled RNA following cleavage with chemical treatment and or by primer extension reaction using 5' ^{32}P-labeled oligonucleotide complementary to the RNA of interest. Modified As and Cs generate stops of reverse transcription. RNA when used as template generates reverse transcription products that correspond in length to the distance between

the oligonucleotide primer and modified A or C and can be resolved on denaturing polyacrylamide gel electrophoresis and then revealed by autoradiography (reviewed by Ehresmann *et al.*, 1987).

Ribonucleases (RNase) can also be used to determine whether or not specific nucleotides in an RNA are base-paired (reviewed by Ehresmann *et al.*, 1987). RNase T1 and V1 have specificity towards unpaired G and paired or stacked nucleotides respectively, and are commonly used to probe RNA structure. RNase digestion of 5' or 3' end-labeled RNA yields fragments that can be resolved by denaturing polyacrylamide gel electrophoresis and autoradiography (reviewed by Ehresmann *et al.*, 1987). The advantage of chemical probes (e.g. DMS) over enzymatic probes is that they are less sensitive to steric hindrance due to their small size. Chemical probes can be used both *in vitro* and *in vivo* and under different conditions giving information about the structure and regions protected by protein binding or structures formed specifically *in vivo* (reviewed by Ehresmann *et al.*, 1987; Higgs *et al.*, 1999; Klaff *et al.,* 1997).

Translation initiation is an important regulated step of gene expression in all organisms in which it has been studied. Translational control is typically brought about by sequences within the 5' untranslated region (UTR) which bind regulatory proteins, adopt specific RNA structures, or both. The 5' UTRs of chloroplast mRNAs are predicted to form structures required for translation regulation (reviewed by Zerges, 2000). Unlike in prokaryotes, where regulatory RNA structures and protein binding sites are located close to the initiation codon and have a repressive role, structures in the 5' UTRs of chloroplast mRNA are

further away form the initiation site on the primary sequence and are required for translation initiation (reviewed by Zerges, 2000). For example, the central region (positions 224-319) of 5' UTR of *psbC* mRNA in the *C. reinhardtii* chloroplast has the potential to form a stem-loop structure (Rochaix *et al.*, 1989). (*psbC* encodes the P6 subunit of the photosystem II reaction center, the homologue of CP43 of plants.) This region is required for translation (Zerges *et al.*, 1997): its deletion abolishes translation from this 5' UTR. The mutation, *psbC-FuD34* (deletion of C315 and insertion of two Ts between A309 and A310), removes two bulges from this stem that is predicted to stabilizes the structure and abolishes translation from the 5' UTR (Rochaix *et al.,* 1989; Zerges *et al.*, 1997). The translational defect caused by this mutation could result from an alteration of RNA structure, of a protein binding site, or both effects. Recently, it was found that the *psbC-FuD34* 5' UTR does indeed form an altered structure, in the absence of Mg^{2+} ions, which is not detected for the wild-type 5' UTR (F. Vigneault and Zerges, unpublished data). Although this structure was proposed over 10 years ago, no study has addressed whether or not it forms .

Sequences in the 5' UTR of the *psbA* mRNA of spinach chloroplasts form a stem-loop structure, as determined by DMS modification studies, and several proteins interact with this UTR (Klaff and Guissem, 1995). RNA-protein complex formation is dependent on the structure of the UTR (Klaff *et al.*, 1997). Deletions within the stem-loop and destabilization of its structure abolish protein binding (Klaff *et al.* 1997). Sequences within the stem-loop were found to be necessary for protein binding (Alexander *et al.*, 1998). A stem-loop has also been described in

the 5' UTR of the *psbA* mRNA of *C. reinhardtii* chloroplasts, which is distinct from the one in spinach, and not entirely within the mature mRNA (Mayfield *et al.*, 1994).

In the 5' UTR of the *petD* mRNA in *C. reinhardtii* chloroplasts three elements were identified: one adjacent to the 5' terminus is required for mRNA stability and two elements are required for translation (Sakamoto *et al.*, 1993; Higgs *et al.*, 1999). Element I, the mRNA stability element at the 5' end of the mRNA, forms a stem-loop, based on DMS modification and RNase mapping. Element II, located in the center of the UTR, is protected from DMS modification by protein binding. Element III, located at the 3' end of the 5' UTR, forms a stem-loop structure (Higgs *et al.*, 1999).

The importance of RNA structure for translation was also found for the mutations in the 5' UTR of *rps7* mRNA (encoding a protein of the chloroplast ribosomal small subunit) that affect the stability and expression of reporter genes in *Escherachia coli* and *C. reinhardtii*. Several of these mutations affect predicted secondary structures in the 5' UTR. Complementary nucleotide changes to reconstitute the wild-type stem-loop structures also restore reporter gene expression. Additional mutations to either strengthen or weaken the reconstituted stems result in reduced reporter expression (Fargo *et al.*, 1999; Fargo *et al.*, 2000).

1.3. Chloroplast Translation

Genes encoding the majority of proteins involved in translation on chloroplast ribosomes (e.g. ribosomal proteins, aminoacyl tRNA synthetases, initiation factors) are absent from the chloroplast genomes of land plants and are assumed to be encoded by the nuclear genome of these organisms (reviewed by Hauser *et al.*, 1998). About two thirds of chloroplast ribosomal proteins and all the chloroplast tRNA synthases in land plants and most algae are known to be encoded in the nuclear genome, synthesized in the cytosol and imported into the organelle (*Gillham et al.*, 1994; Harris *et al.*, 1994).

An *in vitro* system for chloroplast translation has been difficult to develop, with only four publications describing this approach (Hirose and Sugiura, 1996; Hirose et al., 1998). Nevertheless, it is assumed that translation in chloroplasts is carried out in three phases; initiation, elongation and termination (reviewed by Hauser *et al.*, 1998, Zerges, 2000 and Stern *et al.*, 1997). Initiation brings about the assembly of the ribosomal subunits and translations factors at initiation codon on the mRNA. During elongation, the step-wise polymerization of amino acids elongates the polypeptide based on the codon-anticodon interactions. Finally, termination involves the release of the ribosomal subunits from each other and the mRNA at the first stop codon in the coding sequence in a process involving translation termination factors (reviewed by Zerges, 2000).

Translation in chloroplasts is carried out on a 70S ribosome that is similar to that of the eubacteria in protein and rRNA composition (reviewed by Harris, 1994). Recently, proteomic approaches have been used to characterize the complete

complement of proteins in both subunits of the chloroplast ribosome of spinach (Yamaguchi and Subramanian, 2000; Yamaguchi *et al.*, 2000). Of the 25 proteins identified in small subunit (30S), 21 were orthologues of all *E. coli* 30S ribosomal proteins and 4 were plastid-specific ribosomal proteins (Yamaguchi *et al.*, 2000). Out of these 25 proteins, 12 are encoded in the plastid genome, whereas the remaining 13 are encoded by the nuclear genome. In the large subunit (50S) of spinach chloroplast, 33 proteins were identified, of which 31 were orthologues of *E. coli* ribosomal proteins and two were plastid specific (Yamaguchi and Subramanian, 2000). The nuclear genome encodes 25 of the plastid 50S ribosomal proteins whereas eight are encoded in the plastid organelle genome. Out of 20 proteins in the 30S subunit of *Chlamydomonas* chloroplast ribosome 19 were found to be orthologes of *E. coli* ribosomal proteins and a homologue of spinach plastid-specific ribosomal protein-3. In addition a novel S1 domain-containing protein was identified (Yamaguchi *et al.*, 2002). A homologue of the ribosomal protein S1, which directs the small subunit to the translation initiation region by binding to a pyrimidine-rich sequence of the mRNA, was reported in the chloroplast of spinach (Franzetti *et al.*, 1992). Bacterial homologues of initiation factor IF1 in spinach, and, IF2 and IF3 in *Euglena gracilis* have been found (Stern *et al.*, 1997; Gold and Spermulli, 1985; Kraus and Spermulli, 1986).

Although translation in chloroplasts is closely related to eubacteria, some similarities with the eukaryotic cytosolic system have been observed, as shown in Figure 1.1 (reviewed by Zerges, 2000).

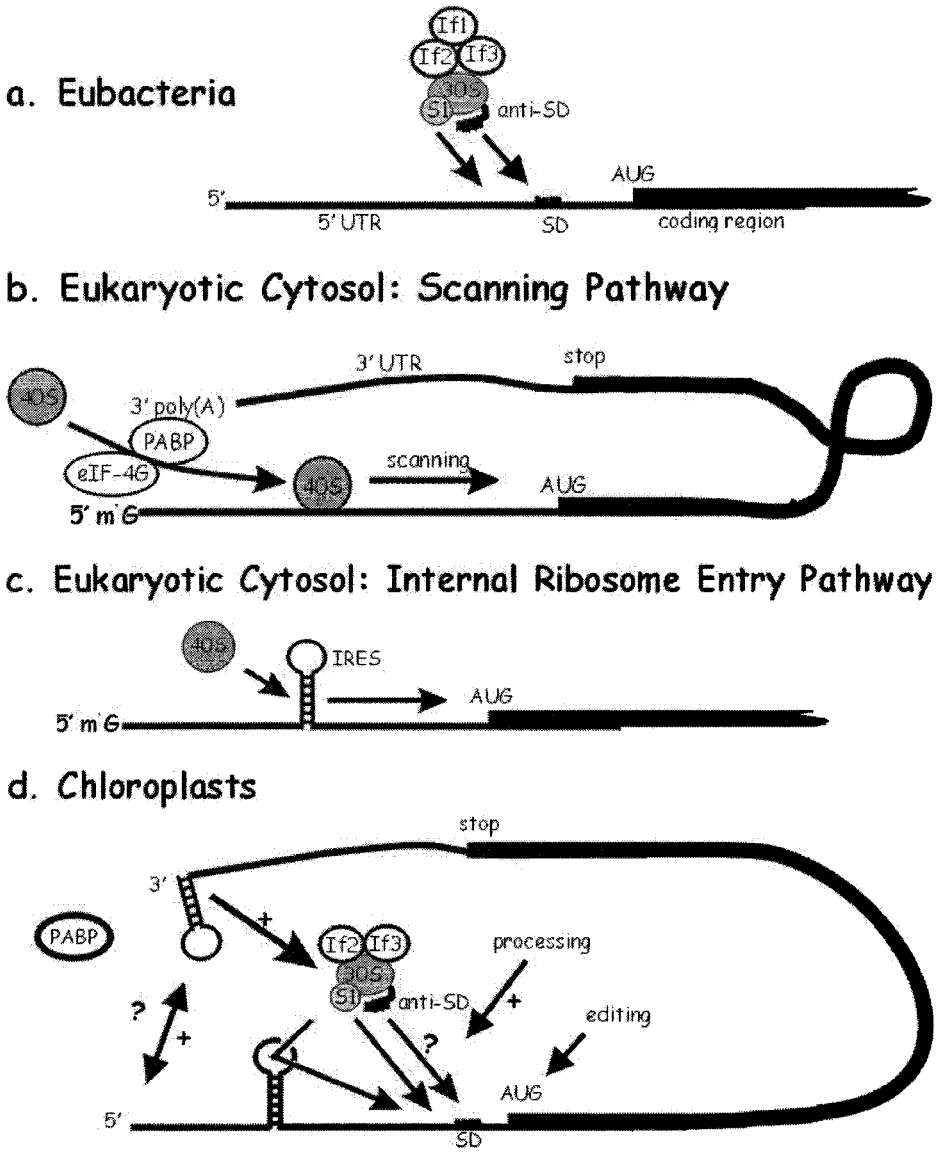


Figure 1.1. Differences and similarities in initiation phase of mRNA translation in eubacteria (a), cytosol of eukaryotes (b and c) and chloroplasts (d) (from Zerges, 2000).

In prokaryotes (Fig.1.1a), the initiation phase begins with the interaction of 30S ribosomal subunit with the mRNA, initiation factors, GTP and fmet-$tRNA_{fmet}$ (McCarthy and Brimacmbe, 1994). Two types of RNA-RNA interactions are known to occur: codon-anticodon (usually AUG), and base paring between Shine Dalgarno sequence (GGAGG), 10-15 nt upstream of the initiation codon of the mRNA, and anti-Shine Dalgarno sequence near the 3' end of the 16S rRNA in small ribosomal subunit. Other interactions involve the S1 protein that plays a role in mRNA recognition, for which a chloroplast homologue exists (Yamaguchi *et al.*, 2000).

The initiation phase of translation in eukaryotes (Fig.1.1b and c) is much more complex than in prokaryotes (reviewed by Pestova *et al.*, 2001). It starts with the binding of eIF2/GTP/Met-$tRNA_i$ ternary complex and other eIFs to the 40S subunit to form a 43S preinitiation complex. Binding of the 43S complex to the 5' terminus m^7G cap is mediated by eIF4F. The bound complex translocates along the 5' non-translated region (5' NTR) scanning for the initiation codon where it forms a 48S complex with initiator tRNA. This is mediated by eIF1A. The final step in initiation is the displacement of factors from 48S complex and binding of 60S subunit to form the 80S ribosome ternary complex. These recruitment steps also require the interaction of the 5' m^7Gcap and poly(A) binding protein (PABP) bound to the 3' poly(A) tail (reviewed by Zerges, 2000).

Initiation on a few cellular and viral mRNAs (Fig.1.1c) is m^7G cap-independent and occurs by internal ribosome entry (reviewed by Pestova *et al.*, 2001 and by Zerges, 2000). Encephalomyocarditis virus (EMCV) and hepatitis C

virus epitomize distinct mechanisms of internal ribosome entry sites (IRES)-mediated initiation (reviewed by Pestova *et al.*, 2001).

In chloroplasts (Fig.1.1d), both similarities and differences with prokaryotic and eukaryotic translation initiation have been observed (reviewed by Zerges, 2000). The Shine Dalgarno (SD) sequence is absent from the 5' UTR of many chloroplast mRNAs, and when present, varies extensively both in position and nucleotide composition (Harris *et al.*, 1994; Fargo *et al.*, 1997; reviewed by Zerges, 2000). The ability to transform the chloroplast genome of *C. reinhardtii* has permitted detailed studies of translational control sequences by site-directed mutagenesis followed by in vivo assessment of their effects on translation. Replacement mutagenesis of SD-like sequence 10 nucleotides upstream of ATG codon of *petD* gene (ATG→TTA), coding for subunit IV of cytochrome b6/f complex, had no effect on *petD* translation (Sakamoto *et al.*, 1994). The 362 nt *petD* 5' leader was shown to be necessary and sufficient to drive expression of GUS gene in transgenic chloroplasts (Sakamoto *et al.*, 1993). Replacement of SD-like sequence of *psbD* (GGAG) by AAAG had no effect on translation while CCUC reduced translation to 25% of wild-type level (Nickelsen *et al.*, 1999). Similarly, replacement of SD-like sequence (GGAGG) of the *psbC* mRNA reduced translation (Zerges *et al.*, 2003). Replacement mutagenesis of SD-like sequences in two ribosomal protein mRNAs (*rsp4* and *rsp7*) and two photosynthetic protein mRNAs (*atpB* and *atpE*) had no effect on expression of a reporter gene in transgenic chloroplasts (Fargo *et al.*, 1997). Similar results were obtained when these chimeric reporter genes were introduced in to *E. coli* (Fargo *et al.*, 1997). In

contrast to these, deletion of SD-like sequence of *psbA* RNA, encoding D1 subunit of the photosystem II reaction center, abolished its translation, but also reduced the level of the mRNA (Mayfield *et al.*, 1994). In an *in vitro* tobacco chloroplast translation system, mutation of SD-like sequence of tobacco chloroplast *psbA* mRNA reduced its translation by less than 15% while SD-like sequence of tobacco chloroplast *rps14* mRNA was found to be necessary for its translation (Hirose *et al.*, 1996).

Although AUG is the initiation codon of most chloroplast mRNAs, as in bacteria, GUG can also be used. This is the case for *psbC* in *C. reinhardtii* (Rochaix *et al.,* 1989). Individual chloroplast mRNAs differ in their requirement for AUG initiation codon. Translation of *petD* mRNA was reduced to 10-20% of the wild-type level when AUG initiation codon was replaced with AUU or AUC (Chen *et al.*, 1993). Five mutations were created in the AUG initiation codon of *petA*, encoding cytochrome f of the cytochrome b6/f complex (Chen *et al.*, 1995). AUG replacement by AUU, ACG, ACC and ACU accumulated cytochrome f to 20%, 2-5%, 1-2% and 0.8% of wild-type levels, respectively, while replacement with UUC completely abolished translation. Only AUU codon mutant could grow phototrophically. Thus, the initiation codon in the *C. reinhardtii* chloroplast appears to be important for translation efficiency but not as the initiation site (Chen *et al.*, 1995).

Chloroplast mRNAs have 5' and 3' untranslated regions (UTRs). Translational *cis*-elements within the 5' UTRs and *trans*-acting protein factors have been identified by site-directed or random mutations, *in vitro* RNA-protein binding

assays and genetic experiments (reviewed by Zerges, 2000). The 5' UTRs of chloroplast mRNAs bear little homology to each other and most are predicted to be highly structured using energy-based models (Fargo *et al.*, 1999). Structures within the 5' UTRs resemble the internal ribosome entry sites (IRES) of viruses (Hepatitis C and Encephalomyocarditis virsus) and few eukaryotic cytosolic mRNAs that are required to recruit small ribosomal subunit independent of m^7G cap and 3' poly(A) tail (reviewed by Pestova *et al.*, 2001; Zerges, 2000).

Poly(A)-rich sequences have been reported in some chloroplast mRNAs (Lisitsky *et al.*, 1996). A homologue of poly(A)-binding protein (PABA) has been proposed to regulate translation of *psbA* mRNA in response to light (Yohn et al., 1998; Danon *et al.*, 1991). 5' UTR of *psbA* mRNA binds to the native complex containing the RB47 protein with higher affinity in presence of 3' UTR in *cis* than the 5' UTR alone (Yael *et al.*, 2002). Replacement of 3'UTR of tobacco chloroplast *psbA* mRNA with the 3'UTR of *rpl32* mRNA decreases the synthesis of fused reporter gene (Eibl *et al.*, 1999). These results indicate the involvement of 3'UTRs in translation regulation and a possible interaction between the 3' and the 5'UTRs as seen in case of eukaryotic translation system (reviewed by Zerges, 2000).

Stability and translation of chloroplast mRNA in *C. reinhardtii* are also affected by RNA processing (reviewed by Stern *et al.*, 1998). RNA processing includes splicing, editing, 3' and 5' termini formation. mRNAs of many genes in the chloroplasts exist in two forms, one being longer at the 5' terminus and being less abundant than the shorter form (reviewed by Rochaix , 1996 and Zerges, 2000). The longer form of *psbA* mRNA is not associated with poly-ribosomes and in the

mutants defective in translation of this mRNA there is an increase in the level of longer form (Bruick *et al.*, 1998). In mutants that are defective in the stability of the *psbD* or *psbB* mRNAs in the *C. reinhardtii* chloroplast, only the short major form is absent (Nickelsen *et al.*, 1999; Vaistij *et al.*, 2000). These results suggest that only the shorter form is translatable (reviewed by Rochaix, 1996). Processing at the 3' termini is required because transcription termination is inefficient or absent in *C. reinhardtii* chloroplasts (reviewed by Stern *et al.*, 1998). Exonucleases and/or endonucleases are involved in the processing and maturation of the 5' and 3' termini of mRNAs in the chloroplasts (reviewed by Stern *et al.*, 1998).

1.4. Nuclear Genes Involved in Chloroplast Translation Regulation

The products of many nuclear loci are involved in expression of chloroplast genes. As illustrated in Figure 1.2, many act in a chloroplast target gene-specific manner, mostly at post transcriptional steps such as RNA stability, processing and splicing, and in translation (reviewed by Rochaix, 1996). In *S. cerevisiae* analogous genetic interactions between nuclear and mitochondrial genome has been described (Fox, 1996). Studies involving nuclear and chloroplast mutants that are affected in chloroplast translation are providing valuable information regarding the molecular mechanisms of chloroplast translation regulation.

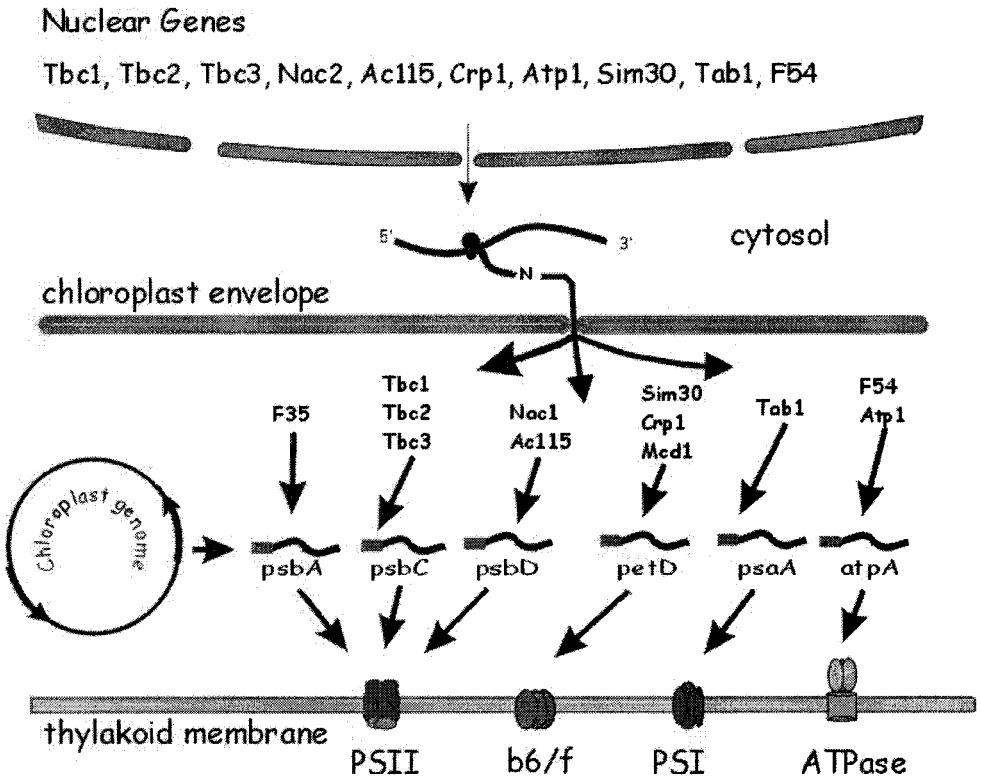


Figure 1.2. The products of nuclear genes control translation of mRNAs encoded by specific chloroplast genes. These nucleus-encoded proteins are synthesized on cytosolic ribosomes and interact directly or indirectly with their target mRNAs transcribed from circular chloroplast genome. The inverted repeats of the chloroplast genome are indicated by arrows. The 5' UTRs of the chloroplast mRNAs are shown as red boxes. (from Zerges, 2000).

Translation of *psbD* mRNA, encoding the D2 subunit of photosystem II, depends on at least two nuclear genes *NAC1* and *AC115* (Kucka *et al.*, 1988; Kuchka *et al.*, 1989). Mutations at *NAC1* and *AC115* loci reduce D2 synthesis, without having corresponding effects on the *psbD* mRNA level. This affect is suppressed by a dominant nuclear mutation *sup4b* by, probably, bypassing requirement for wild-type *NAC1* and *AC115* (Wu and Kuchka, 1995). Another nuclear gene, *NAC2* appears to be required for *psbD* mRNA stability (see below) (Kuchka *et al.*, 1989). Two sequence elements within the 74 nucleotide 5' UTR of *psbD* mRNA are required for translation (Nickelsen *et al.*, 1999). These are the SD-like sequence and a stretch of 11 consecutive U residues; deletion of either abolished *psbD* mRNA translation.

The nuclear mutant F35 is defective in translation of *psbA* mRNA, encoding for D1 subunit of photosystem II (Girard-Bascou *et al.*, 1992). The F35 mutation reduces association of *psbA* mRNA with chloroplast ribosomes and decreases its half-life (Yohn *et al.*, 1996). A 47 kDa protein (RB47) binds to a 36 bp stem loop in the 5'UTR of *psbA* mRNA and has been proposed to mediate the positive light regulation of its translation. (Danon *et al.*, 1991). RB47 is a homologue of poly(A) binding protein (PABP) (Yohn *et al.*, 1998) and its binding is modulated by the levels of ADP and reduced thioredoxin in chloroplast stroma (Danon *et al.*, 1994).

A mutation of the nuclear *TAB1* gene (*tab1*-F15) eliminates translation of *psaA* and *psaB*, encoding the core subunits of the photosystem I reaction center (Stampacchia *et al.*, 1997). A suppressor mutation in the *psaA* 5' UTR, *SuF15*, corresponds to a G→A transition adjacent to a putative SD sequence in *psaB*

5'UTR and restores translation. This suppressor also restores expression of *psaA* showing that the loss of *psaA* expression is a consequence of absence of *psaB* expression (Stampacchia *et al.*, 1997).

Deletion analysis of the 362 nucleotide 5' UTR of *petD*, coding for subunit IV of cytochrome b6/f identified two regions required for its translation (+152-302 and +312-320) (Sakamoto *et al.*, 1994). An unstructured 16 nt region in the middle of the UTR appears to interact with protein factors and a 14 nucleotide region close to the AUG initiation codon. The later region forms a stem-loop RNA secondary structure as indicated by the inability of dimethylsulfate (DMS) to modify A and C residues in the predicted stem (Higgs *et al.*, 1999). An 8 nucleotide region at the 5' end of the UTR was shown to be required for stability of *petD* mRNA.

The nuclear mutant *F54* is defective in translation of *atpA* mRNA, encoding α-subunit of the ATP synthase F1 complex (Drapier *et al.*, 1992).

As described above, a class of nuclear mutants that transcribe but do not accumulate the short form of the mRNA of a single chloroplast gene in *C. reinhardtii* have been proposed to be defective in 5' end processing. This processing, and the wild-type function of the nuclear gene involved, has been proposed to be required for stabilization and translation. Only the shorter "mature" mRNA form is thought to be translated. This proposed form of translational activation by 5' end processing has been seen for *psbB* (Vaistij *et al.*, 2000), *psbA* (Bruick *et al.*, 1998), *psbD* (Nickelsen *et al.*, 1999), and *petD* (Sakamoto *et al.*, 1994). While dual 5' ends have been detected for three of the

four chloroplast mRNAs encoding core subunits of photosystem II, a longer minor form of *psbC* mRNA has not been reported.

One of the best characterized systems of translational control in chloroplasts involves the *psbC* mRNA in the *C. reinhardtii* chloroplast. Previous studies have identified three nuclear genes; *TBC1*, *TBC2* and *TBC3* (for Translation of *psbC*) that function in *psbC* translation. *TCB1* and *TBC2* appear to be specifically required for *psbC* translation; nuclear mutants fail to translate the chloroplast mRNA. *TBC3* could have more general functions in translation in chloroplasts and it is described further below. Translation is controlled by interactions between the *psbC* 547 nt 5' UTR and the products of *TBC1*, *TBC2*, and *TBC3* because the 5' UTR can confer all the genetic interactions to expression of a heterologous reporter gene *in vivo* (Zerges sand Rochaix, 1994; *Zerges et al.*, 1997). A *cis*- acting suppressor mutation, *psbC-F34su1*, reverses translational block caused by mutation in *TBC1* but not *TBC2* (Rochaix *et al.*, 1989).

TBC3 was identified by a dominant nuclear suppressor mutation, which partially reverses the translational defects caused by mutations of the central region; *psbC-FuD34* and a deletion of this region and of the nuclear mutation *tbc1-F34* (Zerges *et al.*, 1997). This suppressor mutation has a weak phenotype on its own and induces a two-fold decrease in the rate of synthesis and in the accumulation of the *psbC* encoded polypeptide (Zerges *et al.*, 1997). Like the *cis*-acting suppressor mutation *psbC-F34sul*, the *tbc3-rb1* suppressor mutation

does not reverse the translational block caused by a mutation in *TBC2* (Zerges *et al.*, 1997).

While *TBC1* and *TBC3* have not yet been characterized at the molecular and biochemical levels, the *TBC2* coding sequence predicts a protein of 1115 residues with nine copies of a novel degenerate 39 amino acid repeat with a quasi conserved PPPEW motif near its C-terminal end (Auchincloss *et al.*, 2002). The central region of the protein displays partial amino acid sequence identity with Crp1, a protein in *Zea mays* that is implicated in the processing and translation of the chloroplast *petA* and *petD* mRNAs (Fisk *et al.*, 1999). The TBC2 protein cofractionates with chloroplast stroma and is associated with a 400 kDa protein complex (Auchincloss *et al.*, 2002).

In a recent study, a series of site-directed mutations in the *psbC* 5' UTR was generated and the ability of these constructs to drive expression of a reporter gene was assayed in chloroplast transformants that are wild-type or mutant at these nuclear loci (Zerges *et al.*, 2003). Two regions within the *psb*C 5' UTR are required for translation initiation (Zerges *et al.*, 1997; Zerges *et al.*, 2003). (The 5' end of the mRNA is +1.) One is located 5' to the GUG initiation codon (Figure 1.3). The other is in the center (+224-320) and, as described above, it has the potential to form a stable stem-loop secondary structure (Rochaix *et al.*, 1989). However, no study has tested its existence. Two bulges in the predicted stem, caused by two sites of non-complementarity between the strands, are required for translation because a mutation that eliminates them, *psbC-FuD34*, completely abolishes translation of the mRNA (Rochaix *et al.*, 1989; Zerges *et al.*, 1997). A

point mutation at position 227 within this region, *psbC-F34sul*, partially reverses the translation block caused by the *TBC1* mutation, but not that of a *TBC2* mutation. This mutation would increase the size of an internal loop at the base of the stem (Rochaix *et al.*, 1989, Zerges and Rochaix, 1994).

Regions with target sites for *TBC1* and *TBC2* were identified by estimating the residual translation activity in the respective mutant backgrounds. *TBC1* targets include the middle and most of the 3' part of the 5' UTR whereas the only detected *TBC2* targets are in the 3' part. The 5'-most 93 nt of the leader are required for wild-type levels of transcription and/or mRNA stabilization. The results indicate that *TBC1* and *TBC2* function independently and raise the possibility that *TBC1* acts together with *TBC3*.

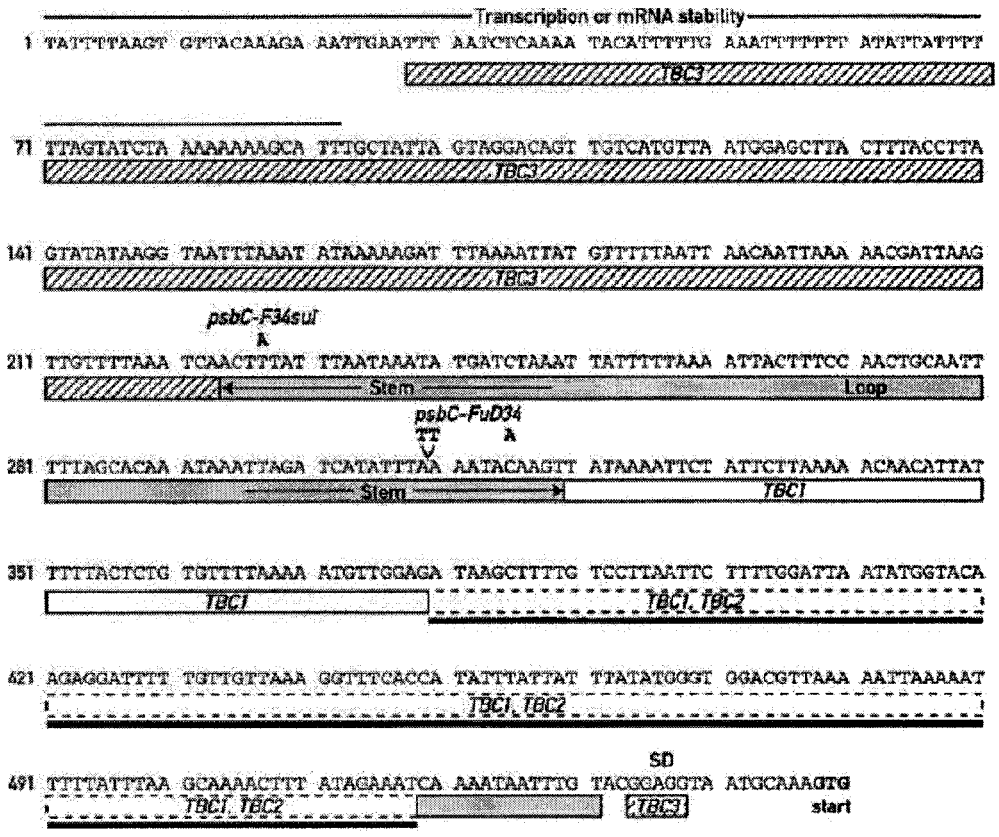


Figure 1.3. Identification of functional regions within the *psbC* 5' leader. The sequence of the 547 nt *psbC* 5' leader (Rochaix *et al.*, 1989) is shown with the relevant regions indicated as follows. Horizontal arrows indicate the complementary sequences that result from two inverted repeats in the DNA from which the 5' UTR is transcribed. Regions that are partially required for translation are indicated with striped boxes: (27-222), and open boxes (321-378). The regions that are important for translation are indicated with a darkened solid box: 223-320, 519-532. "*TBC1*"and "*TBC3*" indicate the target regions of the factors that are encoded or controlled by these loci. The region that is dispensable for translation and partially required for the interactions with *TBC1* and *TBC2* is indicated by a box with a broken border and is underlined. The GUG initiation codon is in bold text. The SD-like sequence (GGAGG) is indicated "SD". The mutations *psbC-FuD34* and *psbC-F34sul* are indicated (from Zerges *et al.*, 2003)

The organization of the 5' UTRs of *psbD and psbC* share several features. Both have a region immediately downstream of the 5' terminus which is required for stability of the mRNA (Zerges *et al.*, 2003; Sakamoto *et al.*, 1994; Sakamoto *et al.*, 1993). Both 5' UTRs have sequence elements in the center and just 5' to the initiation codon that are required for translation (Zerges *et al.*, 2003; Sakamoto *et al.*, 1994; Higgs *et al.*, 1999)

This thesis addresses the structure and function of the *psbC* 5' UTR relating to its ability to drive translation and interact with *TBC1*, *TBC2* and *TBC3*. First, chemical and enzymatic probing experiments addressed whether or not the predicted stem-loop structure forms *in vivo* and *in vitro*. These approaches were used to identify other structures or structured regions of the 5' UTR. Second, primer extension was used to determine whether or not *psbC* mRNAs exist in two forms; a longer minor form and a shorter major form, as has been found for the mRNAs encoding the other photosystem II core subunits, Third, DMS mapping experiments were used to characterize the *psbC* 5' UTR in wild type and the various mutant strains affected in *psbC* translation in order to identify how these nuclear gene products and the *cis*-acting RNA sequence elements with which they interact control translation.

2.1. *Chlamydomonas reinhardtii* Strains and Culture Conditions

All *C. reinhardtii* strains were cultured on tris-acetate-phosphate medium (TAP) under a light intensity of c.a. 100 $\mu Em^{-2}s^{-1}$ and with gentle agitation on an orbital shaker (Harris, 1989). The photosystem II (PSII) deficiency of the mutant strains was verified by testing for the inviability on a medium lacking a reduced carbon source, high salt minimal medium (HSM), under bright light conditions. *C. reinhardtii* strains used are described in Table 2.1 :

Table 2.1. *C. reinhardtii* strains (Rochaix *et al.*, 1989; Zerges *et al.*, 1997).

		Genotype	
Strain	Phenotype	Nuclear	Chloroplast
JVD.4A+	Wild type (WT)	*tbc1, tbc2, tbc3*	*psbC-WT*
F34.3+	PSII deficient	*tbc1-F34, tbc2, tbc3*	*psbC-WT*
F64+	PSII deficient	*tbc1, tbc2-F64, tbc3*	*psbC-WT*
RB1.10A	Partially active PSII	*tbc1, tbc2, tbc3-rb1*	*psbC-WT*
FuD34.3	PSII deficient	*tbc1, tbc2, tbc3*	*psbC-FuD34*
F34sul	Partially active PSII	*tbc1-F34, tbc2, tbc3*	*psbC-F34sul*
RB1.4A	Partially active PSII	*tbc1, tbc2, tbc3-rb1*	*psbC-FuD34*

2.2. RNA Extraction from *Chlamydomonas* Cells

Cultures of 15ml of *Chlamydomonas* in liquid TAP medium were grown to a cell density of 1-5x10^6 cells/ml and then centrifuged at 2988xg (Sorval SS-34 rotor) for 5 minutes. Cell pellets were suspended in 1ml TE (10mM Tris.Cl [pH 8], 1mM EDTA [pH8]) in eppendorf tubes. Cells were pelleted again by centrifugation at 5000 rpm in a microfuge for 5 minutes and kept at –80 °C. To the frozen cell pellet 0.5-0.6 ml of Tri Reagent (Sigma) was added and vortexed twice for 15 seconds.

One scoop of sterile glass beads (equivalent to 40-50 mg) was added to this, vortexed for 15 seconds and maintained at room temperature for 5 minutes. Samples were then centrifuged at 12000 rpm for 5 minutes at 4°C and maintained at room temperature for another 5 minutes. To this, 0.1ml chloroform (isoamyl alcohol free) was added, vortexed well and maintained at room temperature for 5 minutes. Samples were centrifuged at 12000 rpm for 5 minutes at 4°C and aqueous phase was extracted with phenol-chloroform-isoamyl alcohol (24:24:1) and then with chloroform-isoamyl alcohol alone. One volume (0.25ml) of 2-propanol was added to the aqueous phase. These samples were vortexed and maintained at room temperature for 5 minutes. They were then centrifuged at 7500 rpm for 5 minutes and the RNA pellet was washed with 75% ethanol. RNA pellet was resuspended in 20μl of DEPC treated HPLC grade H_2O and stored at –20°C. Total RNA concentration was determined by measuring absorbance at 260nm (Sambrook and Russell, 2001).

RNA concentration (μg / ml) = A_{260} x (40μg / ml) x Dilution factor

2.3. Dimethyl sulfate (DMS) Treatment

In vivo treatment was done in TAP medium (modification of Higgs *et al.*, 1999). *Chlamydomonas* cells (2 x 10^6 cells/ml) were treated with DMS (7.9M, Aldrich) in 15ml TAP (5μl DMS/1ml TAP) for 5 minutes at room temperature with gentle agitation. Reactions were quenched by addition of 50μl β-mercaptoethanol/ 1ml TAP. Cells were pelleted at 2988xg (Sorval SS-34 rotor) for 5 minutes and

washed once with 800μl TE [pH 8]. Cells were stored at –80°C for RNA extraction as described above.

For *in vitro* DMS treatments of RNAs, 75μg RNA in 200μl *in vitro* nuclease buffer (10mM HEPES [pH 7.9], 230mM KCl, 81mM $MgCl_2$, 0.05mM EDTA, 41mM DTT, 8% glycerol) was treated with 1μl DMS at room temperature for 5 minutes (Higgs et al., 1999). These reactions were quenched by addition of 10μl β-mercaptoethanol. RNA was precipitated with 2.5 volume 95% ethanol and washed with 75% ethanol. The RNA pellet as resuspended in 20μl DEPC-treated HPLC grade H_2O to a concentration of approximately 3.5μg/μl.

2.4. Primer Extension

Primers used for primer extension reactions are shown in Table 2.2 below:

Table 2.2. Primers used in primer extension reactions (BioCorp Inc.)

Primer Name	Primer Sequence
Cod100*	5' TAGCACCTAAAGTTTACCT 3'
OligoA**	5' *ACGAGCCATGGA*CACTTTGCATTACCTCCG 3'
utr425	5' ACCTTTAACAACAAAAATC 3'
H4404**	5' *GAGCT*CATTTTTAAAACACAGAGT 3'
LP264	5' AATTGCAGTTGGAAAGT 3'
oligo191	5' ACTTAATCGTTTTTAATTGTT 3'
5'ORF1	5' CCATTAACATGACAACTGTCCTACTAAT 3'

* Hybridizes to the coding sequence downstream of the 5' UTR.
** Sequences italicized do not hybridize to the 5' UTR and contain restriction site.

Primers were 5' end labeled with $[^{32}P]PO_3$ by T4 phage polynucleotide kinase (PNK). This reaction mixture (20μl) contained 10pmoles oligonucleotide,

PNK buffer A (50mM Tris-HCl [pH 7.6], 10mM $MgCl_2$, 5mM DTT, 0.1mM spermindine, 0.1mM EDTA), 10pmoles of [γ -^{32}P] ATP (5000Ci/mmol, Amersham), 10 U of PNK (10U/μl, Fermentas). The reaction mixture was incubated at 37°C for 60 minutes and then inactivated at 65°C for 10 minutes. Labeled primers were purified from [γ -^{32}P] ATP and unlabelled primer on a 12% polyacrylamide gel (Sambrook and Russell, 2001). When the bromophenol blue had ran 2/3 of the gel length the run was stopped. One of the glass plates was removed leaving the gel on the other plate. The gel was wrapped with saran wrap, exposed to X-ray film (Fuji RX) for 5 minutes and the film was developed. The ^{32}P labeled band was excised from the gel and soaked in 300μl of elution buffer (0.5M ammonium acetate, 1mM EDTA [pH 8]) for 4 hours at room temperature with agitation on a rocking-platform shaker (Sambrook and Russell, 2001). These samples were centrifuged at maximum speed in microfuge for 1 min, supernatant removed and gel pellet was washed with 100 μl of elution buffer and added to the remaining supernatant. Eluted primers were precipitated with 2.5 volumes of 95% ethanol and washed with 75% ethanol. The pellet was resuspended in 20μl of DEPC-treated HPLC grade H_2O. 2μl of this (50,000-200,000 cpm) was hybridized to DMS treated RNA for primer extension as described below.

Hybridizations contained 5ug RNA, 50,000-200,000 cpm [^{32}P] primer, which were mixed and adjusted to 100μl with DEPC-treated HPLC grade H_2O. Nucleic acid in these samples was precipitated with 0.1 volume 3M sodium acetate and 2.5 volume 95% ethanol at –20°C for 45 minutes, washed with 75% ethanol, and

resuspended in 30μl of hybridization buffer (150mM HEPES, 1M NaCl, 1mM EDTA). Samples were incubated at 85°C for 5 minutes (in a dry bath) and then allowed to cool slowly to room temperature by removing the dry bath block from the heating device. Nucleic acids were then precipitated by addition of 0.1 volume 3M sodium acetate and 2.5 volumes 95% ethanol, and washed with 75% ethanol. Pellets were resuspended in 25μl reaction mixture containing 0.5mM each or dNTPs, reverse transcriptase buffer (50mM Tris-HCl [pH 8.3], 75mM KCl, 3mM $MgCl_2$ [Invitrogen]), 16mM $MgCl_2$, 12U RNasin (RNase inhibitor,40U/μl, Promega), 0.125μg BSA and 8mM DTT. Extension reactions were carried out with 100U of Superscript II (200U/μl, Invitrogen) with a temporal temperature gradient from 25°C – 55°C (10 minutes at 25°C, 10 minutes at 35°C, 40 minutes at 45°C, 10 minutes at 55°C) and the enzyme was finally inactivated at 75°C for 15 minutes. The nucleic acids were precipitated by addition of 0.1 volume 3M sodium acetate and 2.5 volumes 95% ethanol at -20°C for 60 minutes, followed by centrifugation at maximum speed for 15 minutes. Pellets were washed with 75% ethanol, and then dissolved in 10μl RNA loading dye (95% formamide, 0.025% bromophenol blue, 0.025% xylene cyanol FF, 5mM EDTA, 0.025% SDS). These samples were incubated at 95°C for 5 minutes and cooled on ice before loading on 6% denaturing polyacrylamide gel containing 8M urea.

To determine the precise positions of reverse transcriptase stops in the primer extension reactions, dideoxy termination sequencing ladders were prepared with primers used in primer extension reactions and 5 μg of double-stranded pKS

or pBS containing *psbC* 5' UTR as template using UBS sequencing kit following manufacturer's protocol.

Sequencing and primer extension reaction products were electrophoresed on 6% polyacrylamide:8 M urea gel. For electrolyte gradient gels the anode buffer use for electrophoresis was 0.5X TBE (45mM Tris-borate, 1mM EDTA) while cathode buffer was 0.75X TBE (60mM Tris-borate, 1.3mM EDTA) containing 1M sodium acetate [pH8] (Sambrook and Russell, 2001) . The gel was pre-run for I hour at 50W. Wells were flushed with buffer before loading the samples. Samples were loaded and then subjected to electrophoresis at 50W, 50mA until the bromophenol dye had left the gel. The gel was fixed in 10% acetic acid and 10% methanol for 15 minutes, transferred to chromatography paper, and then dried on gel drier. The dried gel was exposed to X-ray film (Fuji RX or Kodak Biomax MS) and autoradiogram was developed using a Kodak developer.

2.5. *In vitro* Transcription of *psbC* 5'UTR

psbC-FuD34 5'UTR was cloned in pUC19 downstream to T7 promoter for *in vitro* transcription. *psbC-WT* and *psbC-F34sul* 5'UTRs had been cloned previously by F. Vigneault. These templates were designed to have the T7 initiation site as close as possible to the 5' end of the major *psbC* transcript. *In vitro* transcribed *psbC* 5' UTR from these templates will have three extra Gs at the 5' end and a sequence of 6 nucleotides (TCCATG) at the 3' end.

pBS/pKS containing cloned *psbC-Fud34* 5'UTR was used as template to amplify the 5'UTR with PCR. The primer at the distal end of the 5' UTR was FT7m

(5'CGGAATTCG*TAATACGACTAAGGG*ATTTTAAGTGTTACA3') and contained T7 phage promoter (italicized). The primer at proximal end of the 5' UTR was oligoAm (5'CGGAATT*CCATGG*ACACTTTGCATTACCTCCG3'). It hybridizes to a position +550 to +533 relative to the 5' end of the message. Both the primers contained the *EcoRI* restriction site for cloning (underlined). OligoAm also contained *NcoI* restriction site (italicized) described later. The PCR reaction (100µl volume) contained PCR buffer (10mM Tris-HCl [pH 8.8], 50mM KCl, 0.08% Nonidet P40 [Fermentas]), 0.2mM each of dNTPs , 0.25µM each of the primers, 2.5mM $MgCl_2$, 25 µg BSA, 1ng template DNA , 2.5U Taq DNA polymerase (5U/µl, Fermentas). PCR conditions used were as below:

First denaturation		94 °C, 3 minutes
25 cycles	Denaturation:	94 °C, 1 minute
	Annealing:	37 °C, 1 minute
	Extension:	72 °C, 1 minute
Final extension		72 °C, 10 minutes
Stop		4 °C, ∞

PCR reactions were extracted with phenol:chlorform:isoamyl alcohol (24:24:1) once and with chloroform:isoamy (24:1) alcohol twice, ethanol precipitated in the presence of 0.3M sodium acetate, washed with 75% ethanol and resuspended in distilled H_2O.

The PCR amplified fragment and the vector (pUC19) were digested with *Eco*R1 (Fermentas). The digested vector ends were dephosphorylated with shrimp alkaline phosphatase (Fermentas) following manufacturer's protocols. Digested

PCR fragments and vector were agarose gel purified by phenol-freeze-fracture method (Bewsey *et al.*, 1991). Ligation reactions were set up with T4 phage DNA ligase (Fermentas) following manufacturer's protocol. A control reaction contained vector but not insert was set up to access the background frequency of colony. Ligation reactions (2μl of the total volume) were used to transform electrocompetent *E. coli DH5*α cell by electroporation (Sambrook and Russell, 2001). Transformed cells were plated on Luria-Bertani (LB) agar medium containing 50μg/ml of ampicillin, 5-bromo-4-chloro-3-indolyl-β-D-galacto-pyranoside (X-gal) and isopropyl-β-D-thiogalactoside (IPTG). Recombinant colonies (white) were picked, cultured in LB broth with 50 μg/ml of ampicillin and plasmids were purified by the alkaline lysis mini-prep method (Sambrook and Russell, 2001). To confirm the presence of the *psbC* 5' UTR insert and determine the orientation of this insert, these plasmids were digested with *Hind*III (New England Biolabs) following manufacturer's protocol. Selected recombinant plasmids were sequenced with a CEQ analyzer at the Center for Structural and Functional Genomics, Concordia University, to eliminate clones with errors during PCR amplification.

Recombinant plasmids, containing either the wild-type, *psbC-F34sul* or the *psbC-FuD34* 5' UTR, were linearlized by digestion with *Nco*I (Fermentas), which cuts at a site introduced at the 3' end of the UTR. Complete digestion was confirmed by agarose gel electrophoresis and ethidium bromide staining (Sambrook and Russell, 2001). Linearized plasmids were extracted once with phenol:chloroform:isoamyl alcohol (24:24:1), extracted twice with chloroform:

isoamyl alcohol (24:1), followed by ethanol precipitation and washing with 75% ethanol. *In vitro* transcription reactions with T7 phage RNA polymerase (Promega) were prepared at room temperature (to prevent precipitation of the template DNA by spermidine in the buffer) following the manufacturer's protocol. The reactions (100μl) contained transcription optimized buffer (40mM Tri-HCl [pH7.9], 10mM NaCl, 6mM $MgCl_2$, 2mM spermidine), 10mM DTT (100mM), 100 units RNasin (40U/μl, Promega), 0.5mM each of rNTPs, 40U T7 phage RNA polymerase (20U/μl, Promega), 1-3μg of template DNA (linearized plasmid). Reactions were incubated at 37°C for 2 hours and then treated with RQ1 DNase (1U/μl, Promega) at 37°C for 15 minutes (1U DNase/1μg DNA template). The reactions were extracted with phenol:chloroform:isoamyl alcohol, purified on a G25 sephadex spin column (Sambrook and Russell, 2001), ethanol precipitated with sodium acetate and resuspended in DEPC-treated HPLC grade H_2O. RNA size homogeneity was checked on 2% agarose gel.

2.6. RNase Mapping

In vitro transcribed *psbC* 5' UTRs with the wild type sequence, and carrying the mutations *psbC-F34sul* or psbC-FuD34, were labeled at 3' or 5' ends before treatment with RNase. For 5' end labeling, approximately 50 pmoles of RNA was dephosphorylated with 5U of shrimp alkaline phosphatase (Fermentas) in presence of 20U RNasin (40U/μl, Promega) in 50μl reaction volume at 37°C for 30 minutes. RNAs were phenol:chloroform extracted and ethanol precipitated. 5' end-labeling was carried out with T4 phage polynucleotide kinase (Fermentas) and [γ

-^{32}P] ATP as described above. RNA 3' labeling was done with T4 phage RNA ligase (Fermentas) and cytidine 3',5'-(^{32}P)bisphosphate triethylammonium salt (^{32}P-pCp, Amersham). 20 pmoles of RNA was labeled with 40 pmoles of pCp (3000Ci/mmol) in 30μl reaction volume containing 1X RNA ligase buffer (50mM HEPES-NaOH [pH 8.0], 10mM $MgCl_2$, 10mM DTT [Fermentas]), 1mM ATP, 10% dimethyl sulfoxide (DMSO), 0.3μg BSA and 3U T4 phage RNA ligase. Incubation was overnight at 4°C. All labeled RNAs were ethanol precipitated with ammonium acetate and purified on 6% polyacrylamide:8M urea gels (Sambrook and Russell, 2001).

RNA structure was probed with RNase T1 (Fermentas) and RNase V1 (Ambion). RNase treatment was performed with approximately 30,000 cpm RNA in 10μl 1X buffer (10mM Tris [pH 7], 100mM KCl, 10mM $MgCl_2$ [Ambion]) and either 1μg yeast sheared RNA (10mg/ml, Ambion) and 1mU of RNase V1 (1U/μl, Ambion) for 5 minutes at room temperature or 4μg yeast sheared RNA and 500mU of RNase T1 (100U/μl, Fermentas) for 2 minutes at room temperature. Reactions were stopped with 20μl of precipitation/inactivation buffer (provided by Ambion) and incubated at -20°C for at least 1 hour. The samples were centrifuged in a microfuge at 14,000 rpm, for 15 minutes at 4°C and the RNA pellet was washed with 75% ethanol. The RNA pellet was dissolved in formamide RNA loading dye as described above. Samples were incubated at 95°C for 5 minutes and then cooled on ice before electrophoresis on a 6% polyacrylamide gel under denaturing conditions (8M urea) .

As a molecular weight standard for RNA gel electrophoresis, the 100 bp RNA ladder (Fermentas) was labeled with ^{32}P using T4 polynucleodie kinase and and [γ -^{32}P] ATP as described above. A molecular weight "ladder" was prepared by heat-treating labeled RNA (10,000 cpm) with 1μg yeast sheared RNA in 5μl alkaline hydrolysis buffer (50mM sodium carbonate / sodium bicarbonate [pH 9.2], 1mM EDTA) for 2 minutes followed by incubation on ice and addition of 10μl of formamide RNA loading dye. This sample was incubated at 95°C for 5 minutes and then on ice to be electophoresed along with the RNase treated RNAs.

2.7. UV Induced RNA-RNA Cross-linking

In vitro transcribed *psbC* 5' UTR with wild type, *F34su1* and *FuD34* sequence was 5' end-labeled with [γ32P] ATP as described above. Approximately 10,000 cpm RNA was suspended in 10μl buffer containing 10mM Tris [pH 7], 100mM KCl, 10mM $MgCl_2$, and was incubated at 75°C for 5 minutes. These were allowed to cool down to room temperature and were exposed to short wave-length UV radiation for 2 minutes on ice (Branch *et al.*, 1989). RNA was ethanol precipitated in presence of 0.3M sodium acetate, washed with 75% ethanol and resuspended in formamide RNA loading dye. Samples were incubated at 95°C for 5 minutes, cooled on ice and electrophoresed on 6% polyacrylamide gel containing 8M urea. 5' end-labeled RNA treated under similar conditions without exposure to UV was used as control. The gel was fixed, dried and exposed to photographic film as described before.

2.8. Circular Dichroism (CD)

CD spectra were measured for *in vitro* transcribed 5'UTRs (WT, *F34sul* and *FuD34*) with Jasco J700 CD spectrophotometer. CD spectra were recorded from 300 nm to 210 nm at 25°C in 1 ml buffer (6 mM HEPES (pH7.5), 50 mM KCl, 2.6 mM MgCl2, 0.1 mM EDTA, 6% glycerol) containing 20 μg of RNA (Schmid *et al.*, 1996). Date acquisition was every 0.2 in 1cm cell path length. Scans were repeated 4 times.

3.1. Structural probing experiments provide evidence for seven stem-loop structures in *psbC* 5' UTR

The *psbC* 5' UTR was predicted to form several secondary structures with *mfold* computer program (Mathew *et al.*, 1999; Zuker, 1994). Of them, the seven structures supported by the data presented below are shown in Figure 3.1. They are named SL1-SL7, for stem-loop. The largest and most stable stem-loop can be formed by two complementary sequences in the central region of *psbC* 5' UTR (positions 223-320) (Rochaix *et al.*, 1989). SL1 could be formed by sequences between positions 54 and 86, SL2 between positions 134 and 152, SL3 between positions 223 and 320, SL4 between positions 347 and 373, SL5 between positions 375 and 394, SL6 between positions 394 and 411, and SL7 between positions 477 and 495.

To determine whether or not each of the predicted stem-loop structures forms *in vitro*, RNase structural mapping experiments were performed with 5' [^{32}P] end-labeled and 3' [^{32}P] end-labeled *psbC* 5' UTR. The entire 547 base 5' UTR was used to ensure that any structures involving distant or dispersed sequences might form. The T7 promoter was positioned to generate a 5' end close to that of the endogenous mRNA (see Materials and Methods).

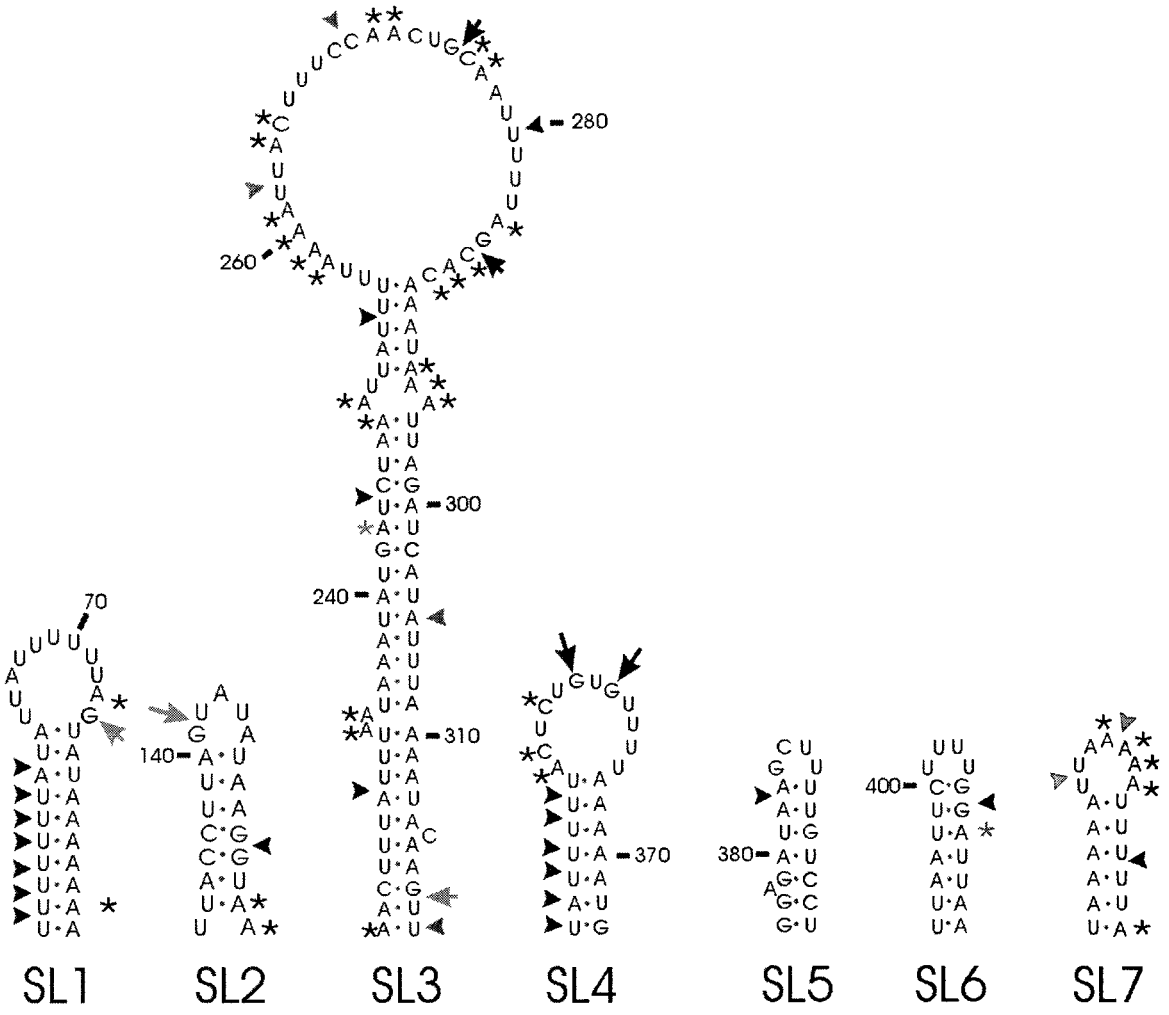


Figure 3.1. Seven stem-loop secondary structures are formed by the *psbC* 5' UTR. The seven stem-loop structures that are supported by the results are illustrated. Asterisks (*) indicate A and C residues that are methylated by DMS *in vitro*. Dots indicate base-pairing interactions. Arrows indicate G residues that are substrates for RNase T1 and thus are not paired. Arrowheads indicated cleavage sites for RNase V1: double-stranded sequences. Most frequently methylated or cleaved positions are indicated by black astericks, arrows and arrowheads. Less frequently methylated or cleaved sites are indicated in grey.

^{32}P-labeled RNA was treated with RNase T1 and RNase V1 in separate reactions. RNase T1 has specificity for unpaired Gs while RNase V1 cleaves at base-paired or stacked nucleotides (Ehresmann *et al.*, 1987). RNase treatment generates fragments that can be resolved by denaturing polyacrylamide gel electrophoresis and visualized by autoradiography. Because cleavage products can form new and irrelevant RNA structures, it was important to ensure that digestion was mild. In every experiment approximately 50% of the RNA substrate was not cleaved to minimize this problem. The positions of RNase cleavage can be determined from the ladder generated with alkaline hydrolysis of end-labeled RNA and ^{32}P-labeled RNA molecular weight markers.

Figure 3.2 (panels a and b) shows results of RNase T1 and RNase V1 digestion of the 3' end-labeled *psbC* 5' UTR with wild type sequence. The RNase cleavages that support structures predicted with *mfold* are shown on these structures in Figures 3.1. No digestion products were observed with the 5' UTR treated under similar conditions but without the addition of RNase, used as control (lane 3 of Figure 3.2). RNase T1 susceptible Gs were found at positions 210, 213, 275, 285, 360, 362, 424, 425, 441, 442 and 513 and are indicated on Figure 3.12 (with the black arrows above the sequence). Other Gs are weakly susceptible to RNase T1 digestion (indicated by white arrows above the sequence). These include Gs at positions 204, 318, 415, 416, 435, 471 and 472. Many regions in *psbC* 5' UTR are susceptible to RNase V1 digestion, and hence, are paired. As summarized in Figure 3.12, these cleavage sites occur at and around U254, U263,

C270, U280, A305, U320, U330, U350, U381, U404-G406, C419-A420, U437-A439, U443-U444, A459-U464, U483-U484, A486-A487, U491-U493, A504-U508, U510-A511, A516-A517, and A520-A522.

Results beyond nucleotides residues downstream to 525 and upstream to 200 could not be obtained with 3' end-labeled UTR because the products resulting from cleavage close to the 5' end comigrated with the intact RNA substrate and cleavages close to the initiation codon generated products that were too short to be seen on the gels (i.e. they ran off with the dye front). Analysis of the translation initiation region should be performed with a shorter RNA substrate. The region upstream of position 200 was mapped with the 5' end-labeled UTR (Figure 3.3). (RNase mapping of *psbC-F34su1* and *psbC-FuD34* 5' UTR in Figure 3.3 are described in later sections). These analyses revealed a similar pattern of cleavages as shown in Figure 3.2 (panels a and b), and described above in Wild type. RNase V1 cleavages were detected at and around positions U54-A61, U91-U92, U110, A121, G150, G168-A174, G180, and A201. RNase T1 cleaved at G74, G88, G94, G103, G124, G126. It was not possible to determine the RNase susceptibility at either extremity of the 5' UTR (5' to position 54 and 3' to 525) because these products resulting for cleavage close to the labeled end ran off the gel with RNAs. RNAs cleaved close to the non-labeled end comigrated with the non-digested RNA. Nevertheless, results described below provided information regarding the structure of the extremities.

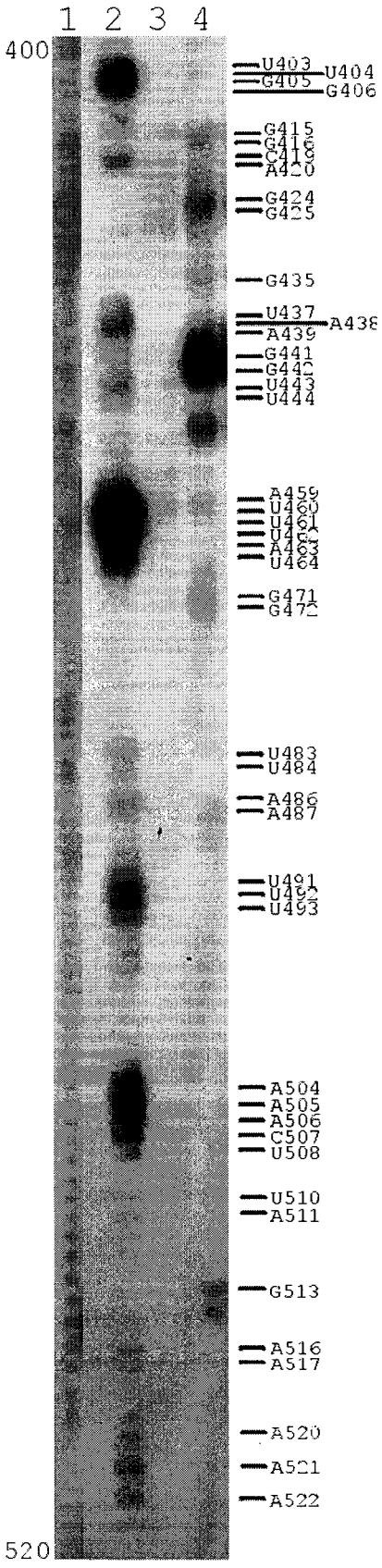
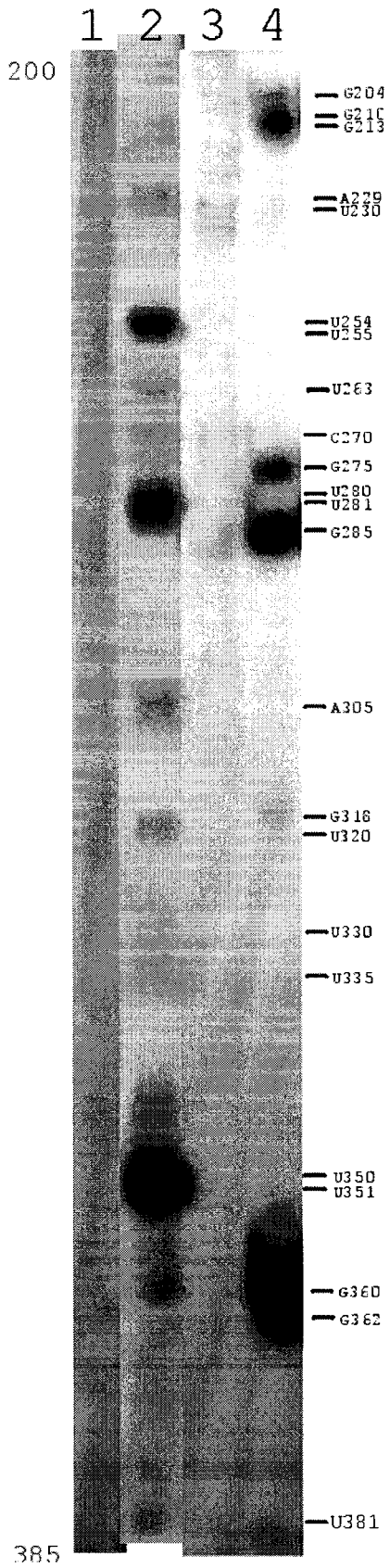


Figure 3.2 Panel a. RNA structure mapping with the enzymatic probes RNase V1 and RNase T1 using 3' end-labeled *psbC* 5' UTR. 3' end-labeled *psbC* 5' UTR was treated with RNase V1 (lane 2) or RNase T1 (lane 4), and compared with untreated 5' UTR (lane 3). Molecular weight markers were (lane 1) was prepared by alkaline hydrolysis of end-labeled RNA. Nucleotides susceptible to RNase digestion are indicated at the right side of the figure. Nucleotides between position 400 to 520 are shown (with respect to the 5' end of the UTR at position 1)



Panel b. Legend is same as Figure 3.2a, except nucleotides between positions 385 to 200 are shown (5' end of the UTR is 1).

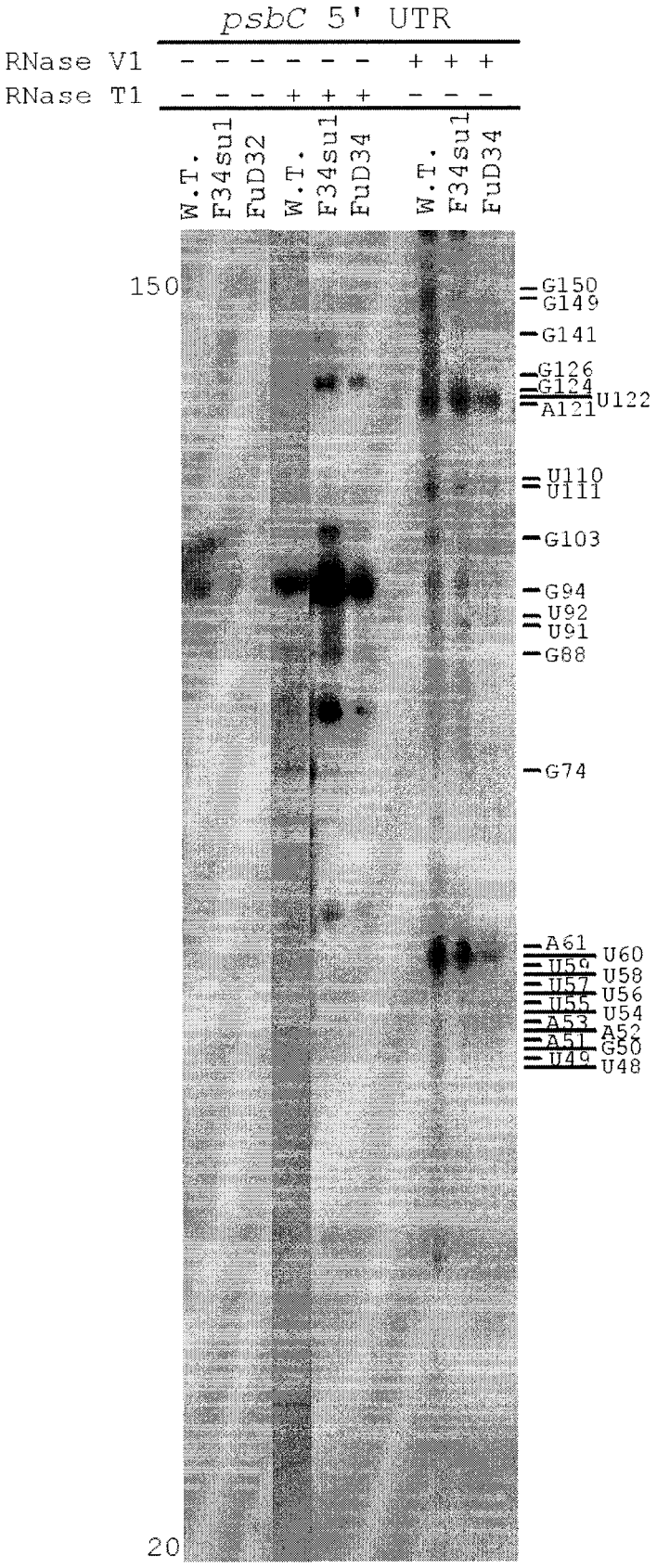
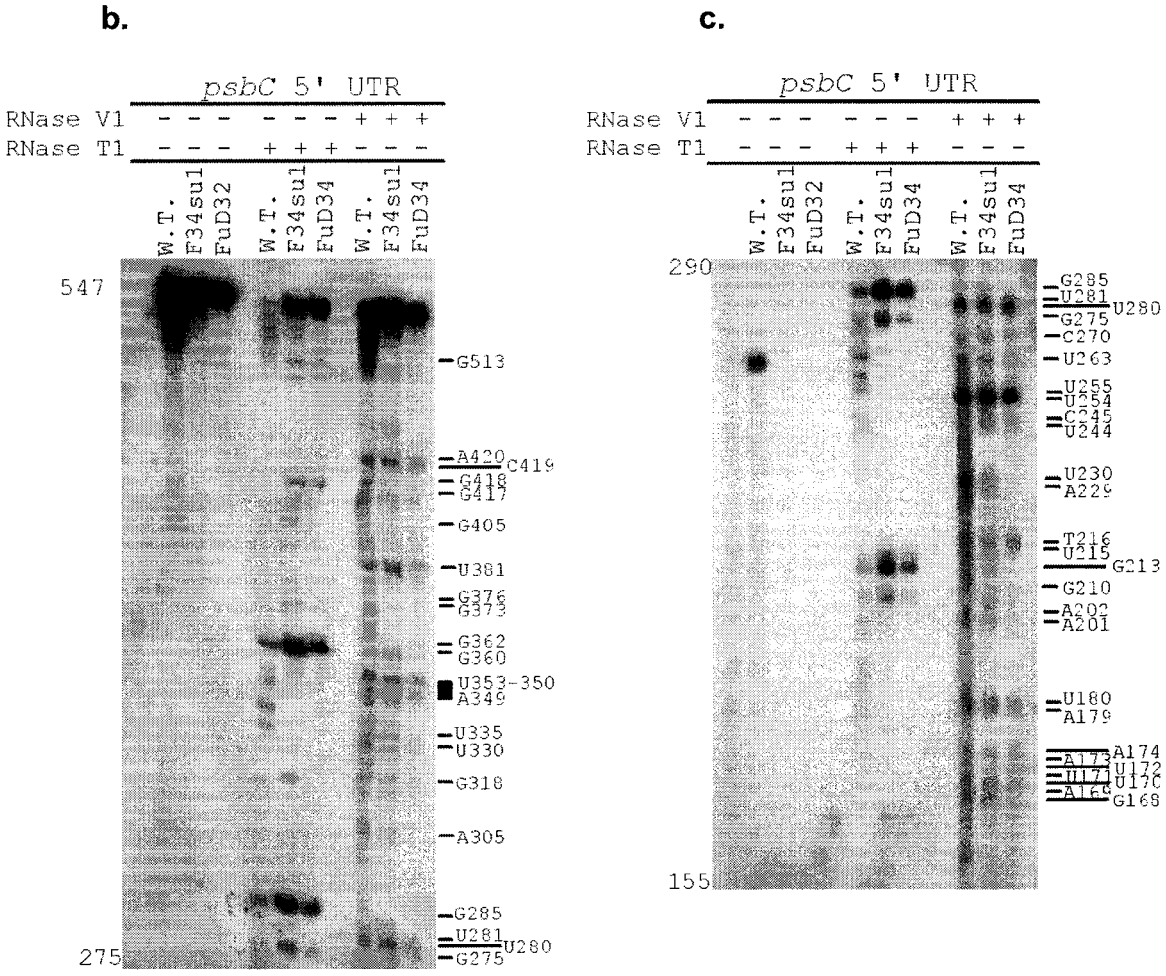


Figure 3.3 Panel a. RNA structure mapping with the enzymatic probes RNase V1 and RNase T1 using 5' end-labeled *psbC* 5' UTR. RNase treatment was performed with 5' end-labeled wild-type and mutant 5' UTRs and compared with untreated 5' UTR. Positions that are susceptible to cleavage are indicated on the right side of the figure. (only positions 20 to 150 are shown).



Panels c and b. Legend is same as Figure 3.3a, except nucleotides between position 1 to 275 (b) and 275 to 547 (c) are shown (5' end of the UTR is 1).

Enzymatic probes are bulky and often cannot access sites which they could otherwise cleave (Ehresmann *et al.*, 1987). For this reason dimethysulfate (DMS) was also used to probe the structure of the *psbC* 5' UTR. DMS treatment of RNA results in methylation of As and Cs that are not base-paired, and thus reveals single-stranded positions (Ehresmann *et al.*, 1987). These modified bases arrest reverse transcriptase (RT) and can be detected by performing primer extension on DMS treated RNA. RNA isolated from wild-type and mutant cells was treated with DMS and modified bases in *psbC* 5' UTR were detected by primer extension of primers that hybridize to *psbC* 5' UTR. Since this UTR is 547 nucleotides long, seven primers hybridizing to different regions of the UTR were used to characterize DMS accessibility throughout the entire UTR (Figure 3.4). Some oligonucleotide primers were unable to hybridize to the UTR due to presence of stable structures (Figure 3.4). For *in vitro* DMS methylation, total RNA was treated with DMS and the methylated As and Cs in *psbC* 5' UTR were detected by primer extension (Higgs *et al.*, 1999). Primer extension cDNA products generated due to RT stops caused by DMS modified As and Cs were electrophoresed on denaturing polyacrylamide gel along with sequencing ladder to determine the positions of these modified bases.

Figure 3.4. The oligonucleotide primer hybridization sites are indicated on the map. Sequence of *psbC* 5' UTR and the 5' part of the coding sequence are represented with the localization of the primers (underlined) used for the primer extension (position 1 is the 5' end of the mRNA). Only sequences complementary to the sequence of the primer are underlined. Primers that hybridized to the 5' UTR are represented by thick underlining and primers that did not hybridize are represented by thin underlining. The coding sequence is shown in italics. The GUG initiation codon is shown in bold characters and the SL3 (223-320) is boxed.

At many regions in the UTR, A and C residues were methylated by DMS indicating that these are unpaired *in vitro*. RNA that had not been treated with DMS was used as control. Many RT stops were observed on the untreated UTR. These will be called "premature RT stops" to distinguish them from RT stops due to methylated A and C residues. RT is known to stall at modified and unusual bases and stable structures. Similar RT stops were observed when primer extension was done on *psbC* 5' UTR generated from an *in vitro* transcription reaction (data not shown). This rules out the possibility that these RT stops are due to unusual bases since the enzymes required to modify bases were not present in the *in vitro* transcription reaction. Moreover, no RT stops were detected in the coding sequence (with primer cod100), which is unlikely to be structured (not shown). Thus, although these premature RT stops significantly hampered analysis of the DMS accessibility of many bases, they did confirm that the 5' UTR is highly structured and provide some information regarding the locations of structured and unstructured regions (see Discussion).

The analyses of the *in vitro* DMS methylation patterns revealed strong premature RT stops at many positions in the 5' UTR. No premature RT stop is seen between positions 475 and 547 (3' end of the UTR) and most of A and C residues at this region are methylated (Figure 3.5a). (In vitro DMS methylation data are summarized in Figure 3.14). The putative Shine Dalgarno sequence (GGAGG) at position 534-538 is unstructured; A536 of this sequence is DMS methylated *in vitro*. On the non-treated RNA many premature RT stops are detected between positions 370 and 490 (shown in Figure 3.5b). Some of the

stops are stronger than the others. These stops are at positions 471-474, 391-394 and 381-385. Nucleotide residues at and close to position 371-375 are predicted to be involved in a pseudoknot interactions with nucleotides in the loop of SL3 (see Discussion). Very few As and Cs are methylated at this region (Figure 3.5b and Figure 3.14). The presence of RT stops make it difficult to map all the methylated As and Cs. Presence of RT stops could be due to stable structures at this region. Premature RT stops are also seen between positions 310 and 380 (shown in Figure 3.5c). The RT stop at position 371-375 (Figure 3.5b) can also be seen. A second RT stop is observed at position 339-344. Not many As and Cs are methylated by DMS at this region (Figure 3.5c and Figure 3.14). Between positions 215 and 305, many As and Cs are methylated by DMS (Figure 3.5d and Figure 3.14). A premature RT stop is observed at position 302-305 in SL3. Many As and Cs that were predicted in the bulges and the loop of SL3 are methylated by DMS (Figure 3.1). Our *in vitro* DMS methylation and RNase mapping data strongly supports this structure. This is the first biochemical evidence that SL3 forms in *psbC* 5' UTR. Some premature RT stops were observed in *psbC-F34su1* mutant that are not present in wild type (lane10 of Figure 3.10). These stops are seen at positions 268-271 and 227-229 in *psbC-F34su1* mutant. Most of the As and Cs between positions 140 and 220 are methylated by DMS (Figure 3.5e). Two premature RT stops are observed at positions 141-145 and 149-151. In Figure 3.11b, a premature RT stop is seen at position A243-C244. This stop is exactly on the opposite strand of the RT stop at C302, A303 and A305 on SL3. This stop is not seen in Figure 3.5d and Figure 3.10 because once the reverse transcriptase

has passed through this site, the stop on the opposite strand no longer exists. This reverse transcriptase stop is also much stronger in *FuD34* mutant compared to the wild type *psbC* 5' UTR. Between positions 40 to 140, shown in Figure 3.5f, many RT stops can be seen. Three strong premature RT stops are observed at positions 91-95, 111-115 and 122-124. Few As and Cs are methylated by DMS and are difficult to determine due to RT stops at this region. Regions close to the 5' end of *psbC* 5' UTR seem to be relatively unstructured. Between positions 1 and 40, most of As and Cs are methylated by DMS and no strong RT stop is present (shown in Figure 3.5g). The strong band at position 1 is the 5' end of the UTR, where all the primer extension eventually stop.

Based on these data, *psbC* 5' UTR seem to be structured. Some of the structures predicted by *mfold* computer programme is supported while others are not (Figure 3.1). The predicted stem-loop structure at position 224-319 (Rochaix *et al.*, 1989) is well supported. This structure could be more stable in *psbC-FuD34* mutant due to strong RT stops in the stem. Many similar strong RT stops are occur throughout the UTR. These stops could represent stable structure like the one seen above or peudoknots and other stable long-range interactions. These interactions could be necessary for the proper folding of the UTR required for translation of *psbC* mRNA. Besides proper folding, *psbC* 5' UTR is involved in interaction with nuclear encoded factor required for translation of the message. Information on this could be obtained from the *in vivo* DMS methylation experiments.

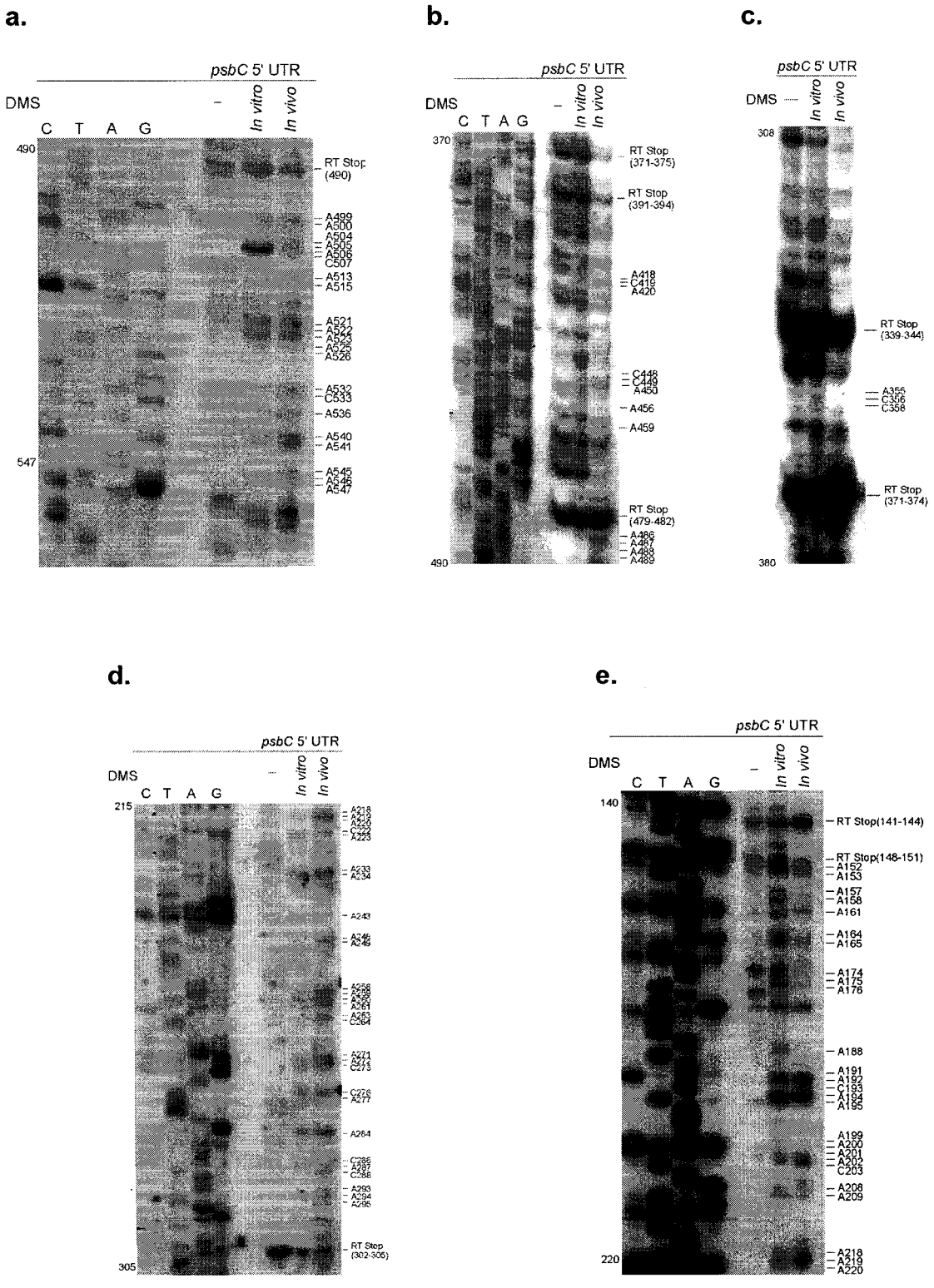


Figure 3.5. Legend on next page.

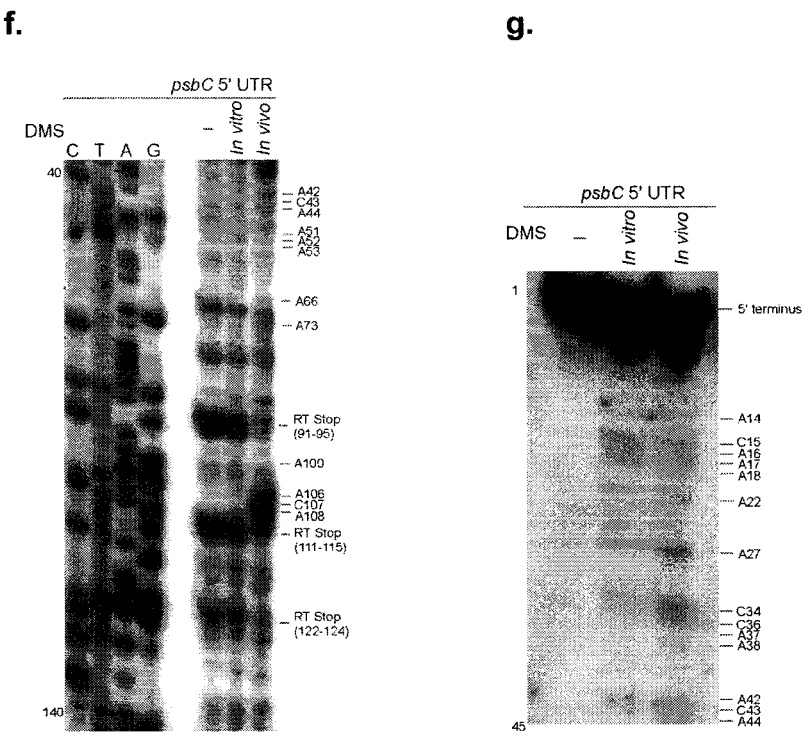


Figure 3.5. Analysis of DMS methylation patterns of the *psbC* 5' UTR in wild-type *in vitro* and *in vivo*. DMS treated RNA (*in vitro* and *in vivo*) and untreated total RNA was analyzed in *Wild Type*. Primer extensions were performed with primers, cod100 (panel a), oligoA (panel b), utr425 (panel c), H4404 (panel d), LP264 (panel e), oligo191 (panel f), and 5'ORF1 (panel g). (For primer positions see Figure 3.4). *In vivo* and *in vitro* DMS methylated nucleotides and strong premature reverse transcription (RT) stops are indicated on the right side of the figure. The number of nucleotides represented by each figure is indicated on the left side of the figure. Dideoxynucleotide mediated chain termination sequencing reactions are also shown (as C T A G above the figure). For a summary of these results see Figure 3.14.

3.2. *In vivo* DMS probing experiments revealed a footprint in the coding sequence in the *tbc2* mutant

Because DMS can enter cells and modify A and C residues *in vivo*, it was used to probe the architecture of the *psbC* 5' UTR in Wild type and the mutant strains affected in *psbC* translation. In *vivo* DMS treatment of the wild type and mutant strains (Table 2.1) was performed to characterize *cis*-acting sequence elements in *psbC* 5' UTR where nuclear encode factor (TBC1, TBC2 AND TBC3) interact. DMS accessibility to bases *in vivo* is inhibited by base-pairing *and* proteins bound to the RNA, or both effects. Distinguishing whether structure or bound proteins are responsible for protection from DMS modification requires comparison of the *in vivo* and *in vitro* patterns, since RNA binding proteins were not present in the latter experiments. In addition, *in vivo* DMS protection of sequences that probably do not form RNA structures, as determined by computer-based predictions and RNase probing experiments, are candidates for protein RNA-binding sites.

These *in vivo* probing experiments also addressed the extent to which the i*n vitro* structural analyses reflect the structure of the 5' UTR *in vivo.* In this regard, it was reassuring to find only minor differences between the patterns of DMS accessibility of the wild-type and mutant 5' UTRs *in vitro* and *in vivo.*

No effect of the *F64-tbc2* mutation on the DMS methylation pattern of the *psbC* 5' UTR was detected. However, for reasons that are unclear, the DMS methylation data could not be obtained for some of the regions of *F64-tbc2 psbC* 5' UTR (Figure 3.14). With the *tbc2* mutant, primer extension signals were

weaker in all experiments, even though many different preparations of RNA were used. It might be relevant that the *tbc2* mutant grows much slower on TAP medium than other strains and reproducibly yielded less RNA.

However, a comparison of *in vivo* DMS methylation of A and C residues in the coding region, i.e. downstream of GUG initiation codon, revealed a reduced or absence of methylation of bases between positions -53 to -65 (relative to the first position of the GUG initiation codon, see Figure 3.6 and 3.14). Almost all As and Cs at this region of *psbC* mRNA are modified by DMS in wild type and other mutant strains. This hypomethylated region in the *tbc2* mutant could extend further 3' and further experiments should address this possibility.

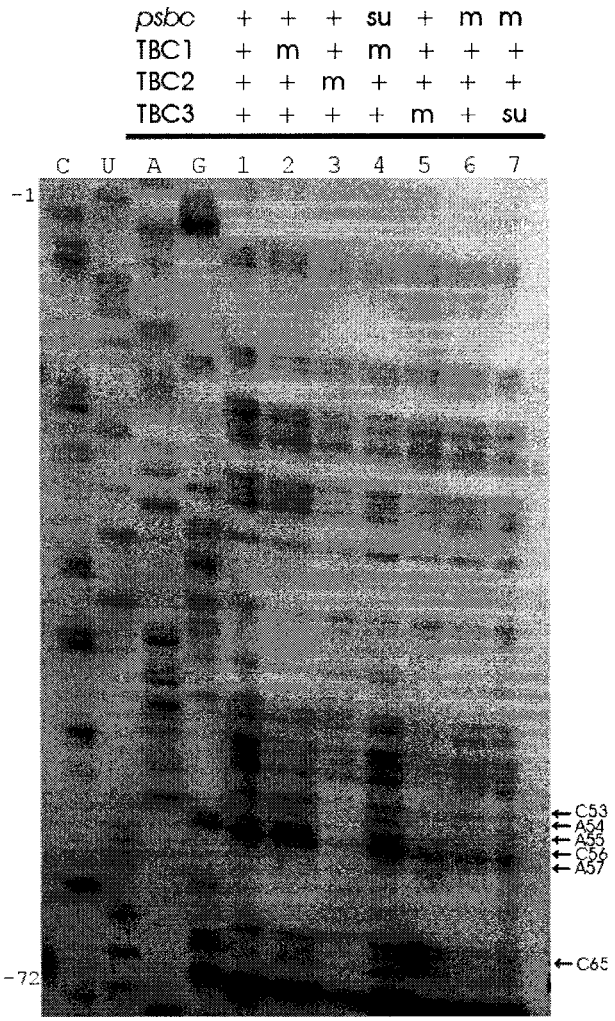


Figure 3.6. Analysis of DMS methylation patterns of the 5' end of the *psbC* coding region (positions -1 to –72, relative to the GUG initiation codon) in wild-type and mutant strains *in vivo*. DMS treated RNA (*in vivo*) was analyzed in the following strains; *Wild Type* (Lane 1), a *tbc1* mutant strain (Lane 2), a *tbc2* mutant strain (Lane 3), a double mutant strain for *tbc1* and the suppressor mutation in the *psbC* 5' UTR (*psbC-F34sul*) (Lane 4), a strain with a wild-type *psbC* allele and the suppressor mutation in *tbc3* (Lane 5), a mutant for the *psbC* 5' UTR (with *psbC-FuD34*) (Lane 6) and a double mutant for the *psbC* 5' UTR (*psbC-FuD34*) and the suppressor mutation in *tbc3* (Lane 7). Primer extensions were performed with primer 'cod100" (Figure 3.4). Differences in *in vivo* DMS methylation pattern in the *tbc2* mutant are indicated by arrows. For a summary of these results see Figure 3.14.

3.3. Specific sequence elements that mediate translational control by *tbc1* and *tbc3* were revealed by results of *in vivo* DMS probing experiments.

Many differences are observed between the *in vivo* DMS methylation patterns of wild-type, and mutant strains for *tbc1* and *tbc3*. Recall that the *psbC* 5' UTR sequence is identical in these strains and thus, differences in DMS accessibility should reflect alterations in the 5' UTR structure, bound proteins, or both effects. Changes in the DMS methylation pattern in the *tbc1* mutant that are most directly related to the translational defect should be reversed by the *psbC-F34sul* mutation in the stem of SL3. Thus, the *in vivo* DMS methylation pattern of the *psbC 5'* UTR was determined in a double mutant strain: with the *tbc1-F34* and *psbC-F34sul* mutations.

In the *tbc1* mutant, C544, close to located between the SD-like sequence (see Introduction) and the GUG initiation codon is accessible to DMS (Figure 3.7, lane 6). This base is not methylated in the wild-type strain or in other mutant strains (Figure 3.7). The A and C residues between positions 459 and 426 are not methylated in the *tbc1* mutant, although they are methylated in all other strains (Figure 3.6, lane 6). There is also hypermethylation of bases A407 (Figure 3.8, lane 6), C269, A271, A272 and C273, present in the loop region of SL3 (Figure 3.10, lane 6), AA42, C43, A44, A67, A87, C89, A90, A125 (Figure 3.12, lane 6), A22, A26, C34 and C36 (Figure 3.13, lane 6). Base A278 is hypomethylated in *tbc1* mutant (Figure 3.10). Many of these effects were reversed by *F34su1* mutation (see Figure 3.14).

Many effects were seen in the *tbc3* mutant and there was some correspondence of these alterations with the *tbc1*. In one location opposite effects were seen in the *tbc1* and *tbc3* mutants. In the *tbc1* and *tbc3* mutants A407 is hypermethylated (Figure 3.8). Hypermethylation of these bases was also detected in the strain carrying the *psbC-FuD34* mutation. Hypomethylation of bases A278, C222 and A223 is observed in both the *tbc1* and *tbc3* mutants (Figure 3.10). C89 and A90 are hypermethylated in both *tbc1* and *tbc3* mutants (Figure 3.12). In the *tbc3* mutant, bases A309, A310, A311 and A312 are hypermethylated (Figure 3.9, lane 12). These base are also hypermethylated in *psbC-F34su1* and *psbC-FuD34.* All the As and Cs between positions 152 and 179 are hypermethylated in *tbc3* and *psbC-F34su1* mutants (Figure 3.11a). A120 and A121 are hypermethylated in t*bc3* mutant (Figure 3.12). The A42, C43 and A44, which are hypermethylated in the *tbc1* mutant, are hypomethyated in *the tbc3* mutant (Figure 3.12).

As described above a strong premature RT stop was detected at positions 302-304 on the mutant *psbC-FuD34* 5' UTR (Figure 3.10). This premature RT stop is weaker on RNA from DMS-treated wild-type cells, but it is clearly observable (Figure3.10, lane 3). Even though the *tbc1* mutant has the wild-type *psbC* 5'UTR sequence, this premature RT stop was consistently stronger on RNA isolated from this strain (Figure 3.10, lane 4-6). It is difficult to explain how the *tbc1* mutation could increase the intensity of this premature RT stop on the RNA after it has been isolated from all proteins. Some protein remain tightly bound to RNAs, even following several rounds of phenol

extraction. Future experiments should address this possibility. It is unlikely that this premature RT stop is caused by a modified base (which could be more frequent in the *tbc1* mutant) because, as described above, it was clearly detected on the 5' UTR generated in an *in vitro* transcription reaction (not shown). It would seem that RNA isolated from the *tbc1* mutant has a more stable RNA structure involving this region, even though the RNA sequence of the entire 5' UTR is wild-type.

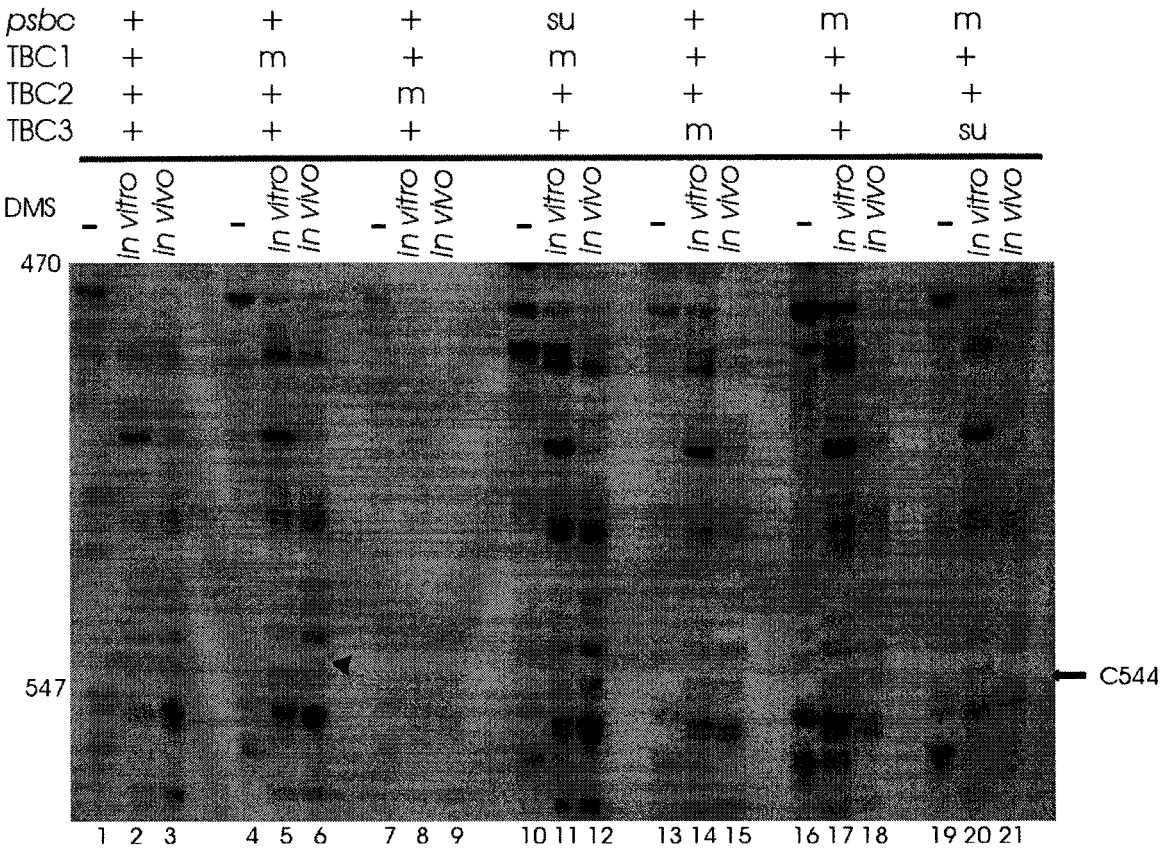


Figure 3.7. DMS methylation analysis of the translation initiation region (470-547) *in vitro* and *in vivo* for wild-type and mutant strains. DMS treated RNA (*in vitro* and *in vivo*) and untreated total RNA was analyzed in the following strains; *Wild Type* (Lane 1-3), a *tbc1* mutant strain (Lane 4-6), a *tbc2* mutant strain (Lane 7-9), a double mutant strain for *tbc1* and the suppressor mutation in the *psbC* 5' UTR (*psbC-F34sul*) (Lane 10-12), a strain with a wild-type *psbC* allele and the suppressor mutation in *tbc3* (Lane 13-15), a mutant for the *psbC* 5' UTR (*psbC-FuD34*) (Lane 16-18) and a double mutant for the *psbC* 5' UTR (*psbC-FuD34*) and the suppressor mutation in *tbc3* (Lane 19-21). Primer extensions were performed with primer 'cod100' (Figure 3.4) Genotypes are indicated at the top by '+' as wild type, 'm' as mutant and 'su' as suppressor. These symbols are the same in the following figures, as otherwise indicated. Differences in *in vivo* DMS methylation pattern are indicated by arrow-heads and shown on the right side of the Figure. For a summary of these results see Figure 3.14.

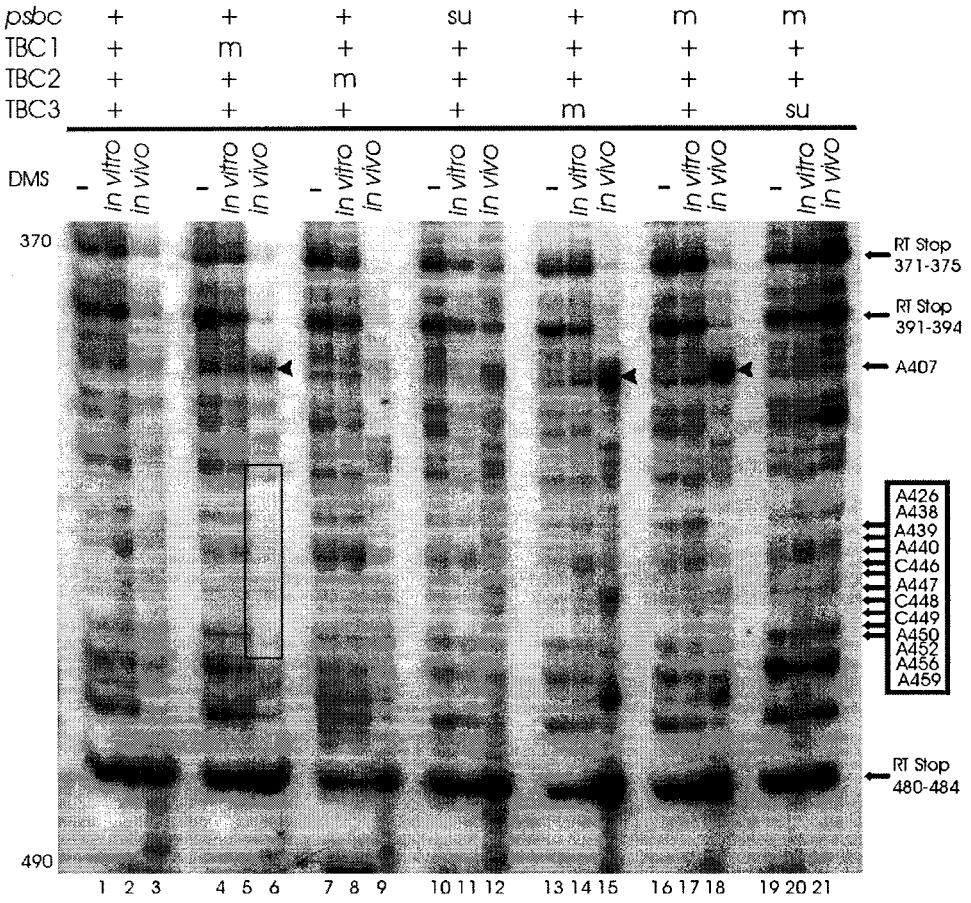


Figure 3.8. Analysis of DMS methylation patterns of the *psbC* 5' UTR positions 370-490 in wild-type and mutant strains *in vitro* and *in vivo*. DMS treated RNA (*in vitro* and *in vivo*) and untreated total RNA was analyzed in the following strains; *Wild Type* (Lane 1-3), a *tbc1* mutant strain (Lane 4-6), a *tbc2* mutant strain (Lane 7-9), a double mutant strain for *tbc1* and the suppressor mutation in the *psbC* 5' UTR (*psbC-F34sul*) (Lane 10-12), a strain with a wild-type *psbC* allele and the suppressor mutation in *tbc3* (Lane 13-15), a mutant for the *psbC* 5' UTR (with *psbC-FuD34*) (Lane 16-18) and a double mutant for the *psbC* 5' UTR (*psbC-FuD34*) and the suppressor mutation in *tbc3* (Lane 19-21). Primer extensions were performed with primer 'oligoA" (Figure 3.4). Differences in *in vivo* DMS methylation pattern are indicated by arrow-heads. The region of lane 6 that revealed the hypomethylated bases in the *tbc1* mutant are indicated by the box. Strong premature reverse transcription (RT) stops are indicated as "RT Stop". For a summary of these results see Figure 3.14.

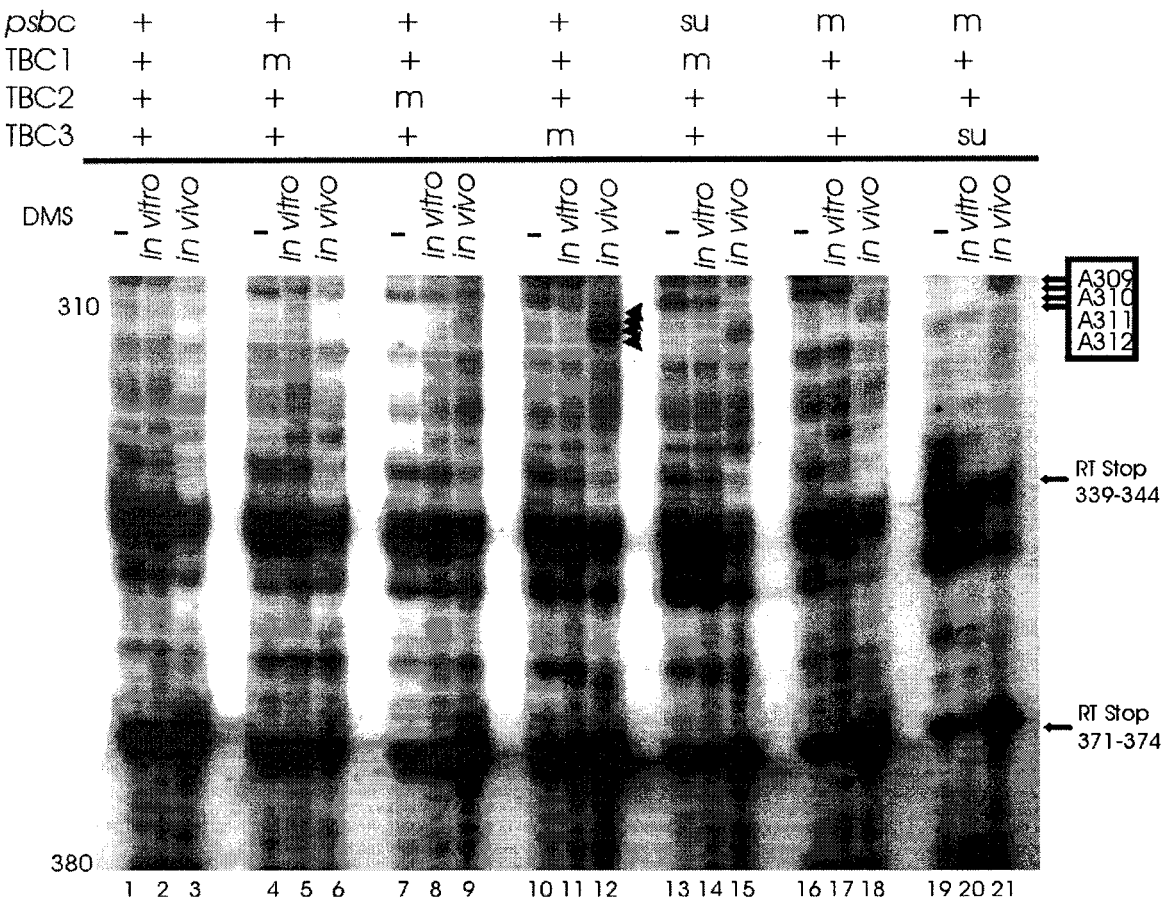


Figure 3.9. Analysis of DMS methylation patterns of the *psbC* 5' UTR positions 310-380 in wild-type and mutant strains *in vitro* and *in vivo*. DMS treated RNA (*in vitro* and *in vivo*) and untreated total RNA was analyzed in the following strains; *Wild Type* (Lane 1-3), a *tbc1* mutant strain (Lane 4-6), a *tbc2* mutant strain (Lane 7-9), a strain with a wild-type *psbC* allele and the suppressor mutation in *tbc3* (Lane 10-12), a double mutant strain for *tbc1* and the suppressor mutation in the *psbC* 5' UTR (*psbC-F34sul*) (Lane 13-15), a mutant for the *psbC* 5' UTR (with *psbC-FuD34*) (Lane 16-18) and a double mutant for the *psbC* 5' UTR (*psbC-FuD34*) and the suppressor mutation in *tbc3* (Lane 19-21). Primer extensions were performed with primer 'utr425" (Figure 3.4) Differences in *in vivo* DMS methylation pattern is indicated by arrow-heads. Bands that reveal strong premature RT stops are indicated as "RT Stop". For a summary of these results see Figure 3.14.

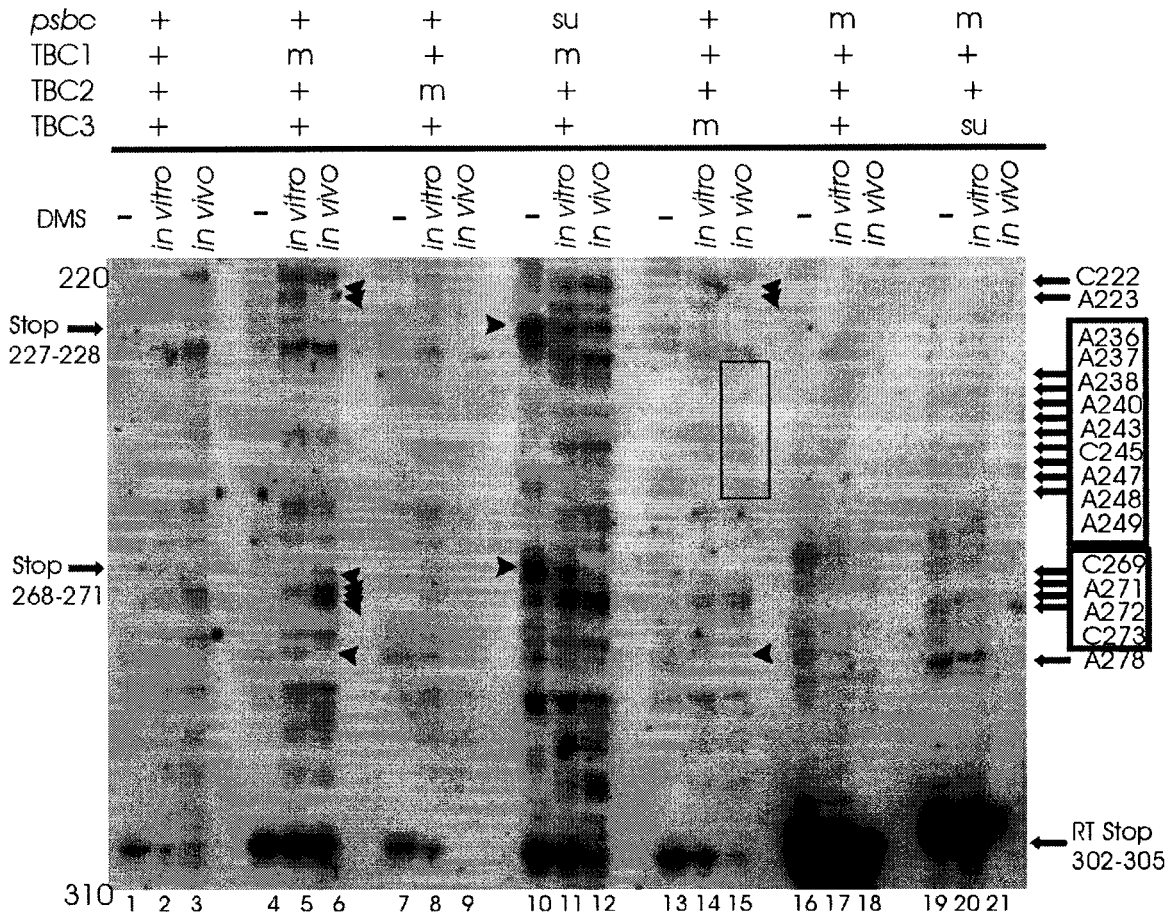


Figure 3.10. Analysis of DMS methylation patterns of the *psbC* 5' UTR positions 200-310 in wild-type and mutant strains *in vitro* and *in vivo*. DMS treated RNA (*in vitro* and *in vivo*) and untreated total RNA was analyzed in the following strains; *Wild Type* (Lane 1-3), a *tbc1* mutant strain (Lane 4-6), a *tbc2* mutant strain (Lane 7-9), a double mutant strain for *tbc1* and the suppressor mutation in the *psbC* 5' UTR (*psbC-F34sul*) (Lane 10-12), a strain with a wild-type *psbC* allele and the suppressor mutation in *tbc3* (Lane 13-15), a mutant for the *psbC* 5' UTR (with *psbC-FuD34*) (Lane 16-18) and a double mutant for the *psbC* 5' UTR (*psbC-FuD34*) and the suppressor mutation in *tbc3* (Lane 19-21). Primer extensions were performed with primer 'H4404' (Figure 3.4). Differences in *in vivo* DMS methylation pattern is indicated by arrow-heads. The region of lane 15 that reveals a hypomethylated sequence in the mutant for *tbc3* is indicated by the box. Bands corresponding to strong premature reverse transcription (RT) stops are indicated as "RT Stop". For a summary of these results see Figure 3.14.

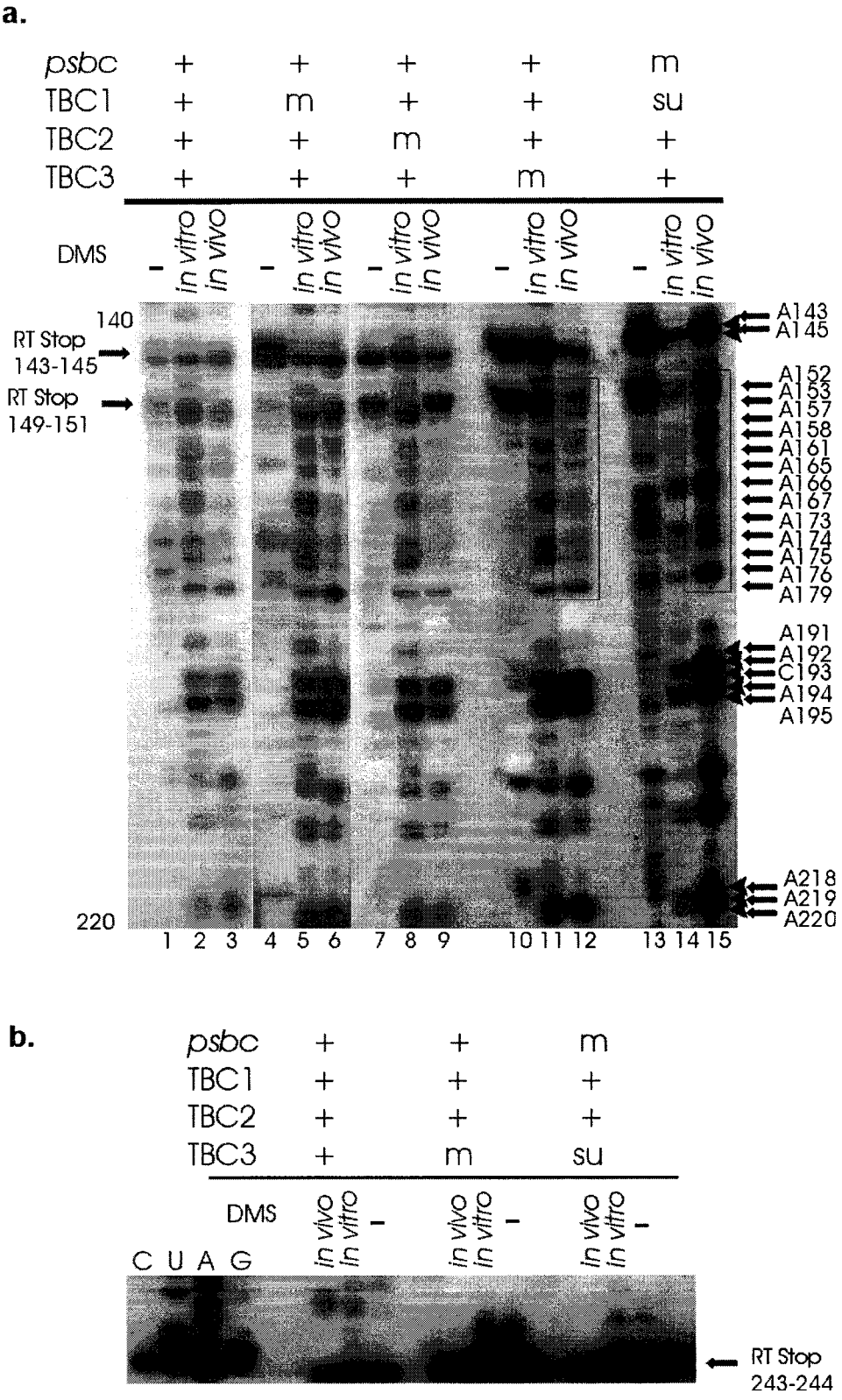


Figure 3.11. Legend on next page

Figure 3.11. Analysis of DMS methylation patterns of the *psbC* 5' UTR positions 143-220 and 243-244 in wild-type and mutant strains *in vitro* and *in vivo*. DMS treated RNA (*in vitro* and *in vivo*) and untreated total RNA was analyzed in the following strains; *Wild Type* (Lane 1-3), a *tbc1* mutant strain (Lane 4-6), a *tbc2* mutant strain (Lane 7-9), a strain with a wild-type *psbC* allele and the suppressor mutation in *tbc3* (Lane 10-12), a double mutant strain for *tbc1* and the suppressor mutation in the *psbC* 5' UTR (*psbC-F34sul*) (Lane 13-15). Primer extensions were performed with primer 'LP264" (Figure 3.4). Differences in *in vivo* DMS methylation pattern are indicated by arrow-heads .The region of lane 15 that reveals a hypermethylated sequence in the double mutant for *psbC-F34su1* and *tbc1* and in strain with *tbc3* mutation is indicated by the box. Bands that reveal strong premature reverse transcription (RT) stops are indicated as "RT Stop". Panel a shows results for positions 143-220. Panel b shows results that reveal the strong premature RT stop at positions 243-244 for *psbC-wild type*, *psbC-FuD34* and a double mutant for *psbC* 5' UTR (*psbC-FuD34*) and tbc3. This stop is enhanced on the 5' UTR with the *psbC-FuD34* mutation. For a summary of these results see Figure 3.14.

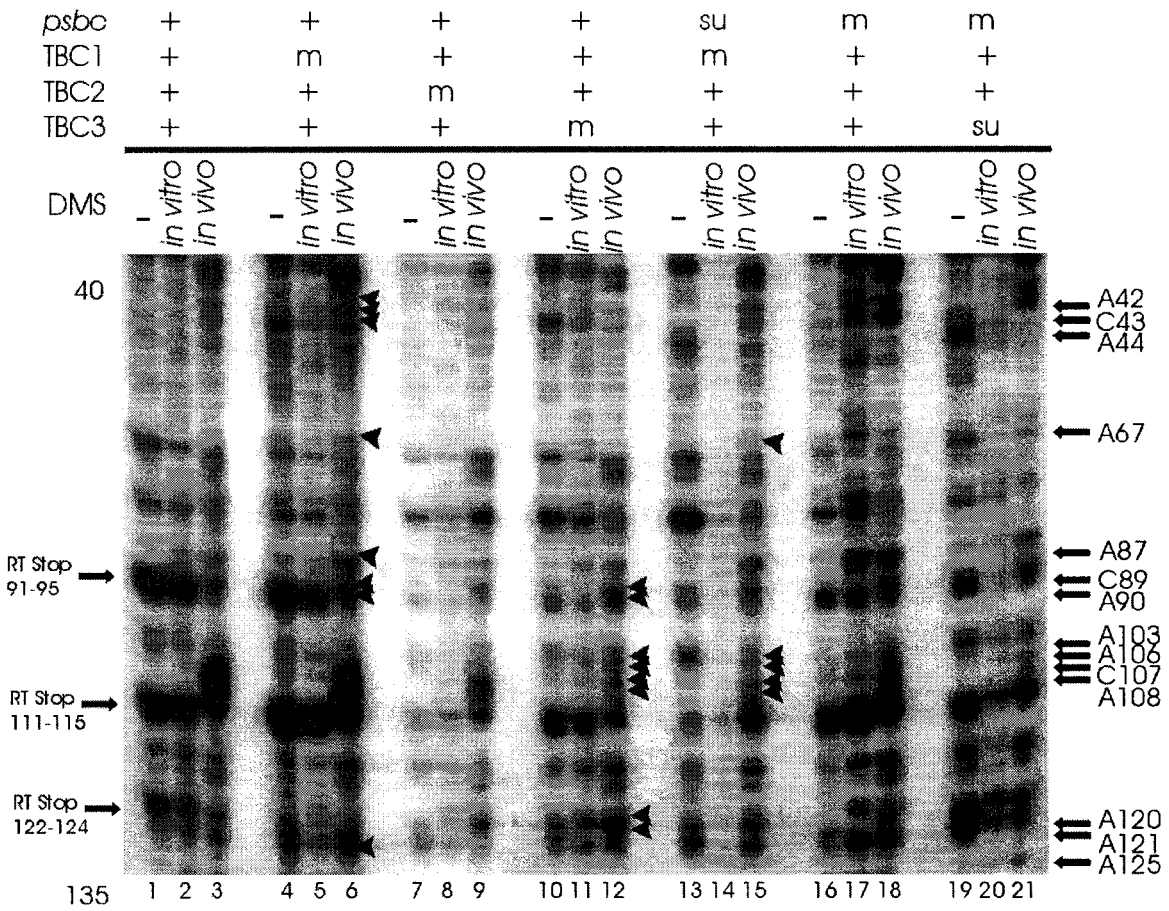


Figure 3.12. Analysis of DMS methylation patterns of the *psbC* 5' UTR positions 40-130 in wild-type and mutant strains *in vitro* and *in vivo*. DMS treated RNA (*in vitro* and *in vivo*) and untreated total RNA was analyzed in the following strains; *Wild Type* (Lane 1-3), a *tbc1* mutant strain (Lane 4-6), a *tbc2* mutant strain (Lane 7-9), a strain with a wild-type *psbC* allele and the suppressor mutation in *tbc3* (Lane 10-12), a double mutant strain for *tbc1* and the suppressor mutation in the *psbC* 5' UTR (*psbC-F34sul*) (Lane 13-15), a mutant for the *psbC* 5' UTR (with *psbC-FuD34*) (Lane 16-18) and a double mutant for the *psbC* 5' UTR (*psbC-FuD34*) and the suppressor mutation in *tbc3* (Lane 19-21). Primer extensions were performed with primer 'oligo191" (Figure 3.4). Differences in *in vivo* DMS methylation pattern are indicated by arrow-heads. Bands that reveal strong premature reverse transcription (RT) stops are indicated as "RT Stop". For a summary of these results see Figure 3.14.

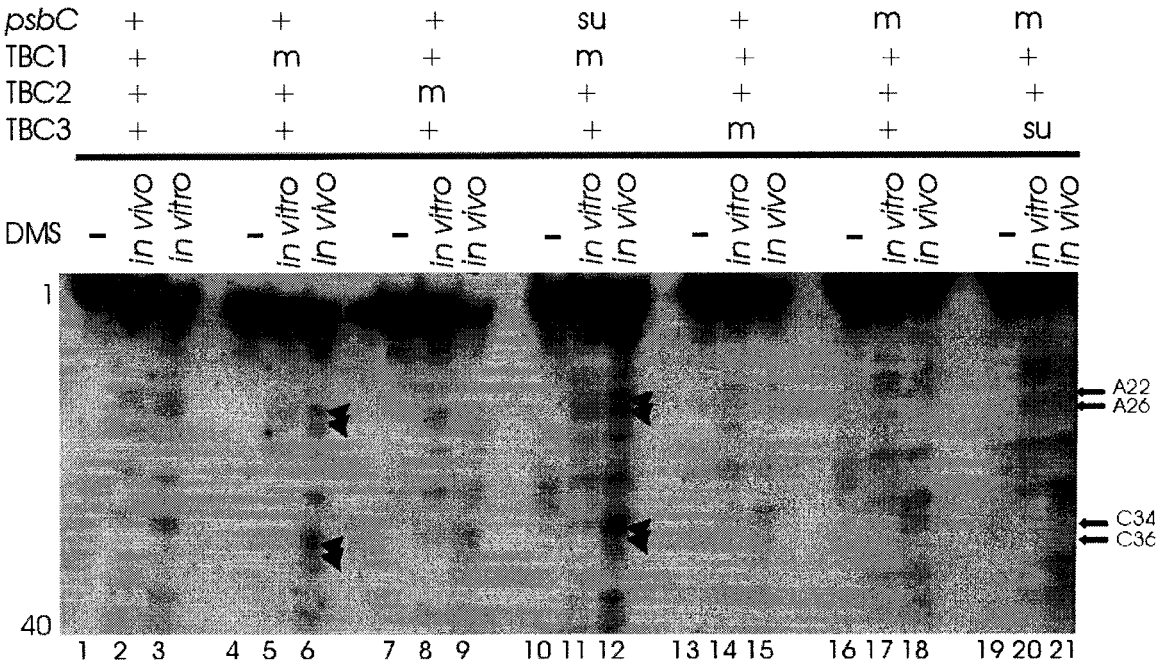
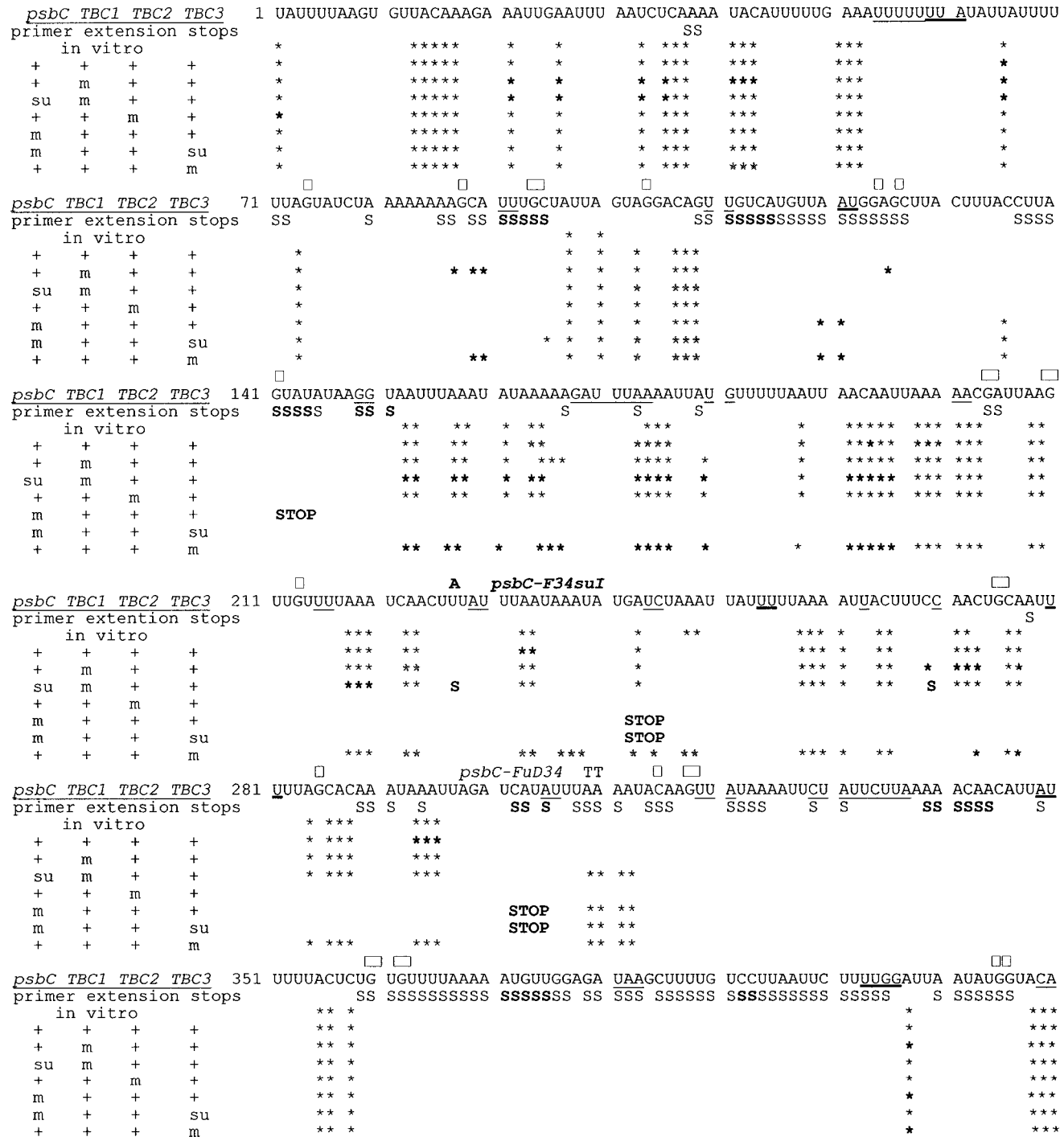


Figure 3.13. Analysis of DMS methylation patterns of the *psbC* 5' UTR positions 1-40 in wild-type and mutant strains *in vitro* and *in vivo*. DMS treated RNA (*in vitro* and *in vivo*) and untreated total RNA was analyzed in the following strains; *Wild Type* (Lane 1-3), a *tbc1* mutant strain (Lane 4-6), a *tbc2* mutant strain (Lane 7-9), a double mutant strain for *tbc1* and the suppressor mutation in the *psbC* 5' UTR, *psbC-F34sul*;(Lane 10-12), a strain with a wild-type *psbC* allele and the suppressor mutation in *tbc3* (Lane 13-15), a mutant for the *psbC* 5' UTR (with *psbC-FuD34*) (Lane 16-18) and a double mutant for the *psbC* 5' UTR (*psbC-FuD34*) and the suppressor mutation in *tbc3* (Lane 19-21). Primer extensions were performed with primer '5'ORF1" (Figure 3.4). Differences in *in vivo* DMS methylation pattern are indicated by arrow-heads. Bands that reveal strong premature reverse transcription (RT) stops are indicated as "RT Stop". For a summary of these results see Figure 3.14.

Figure 3.14. Legend on the next page.



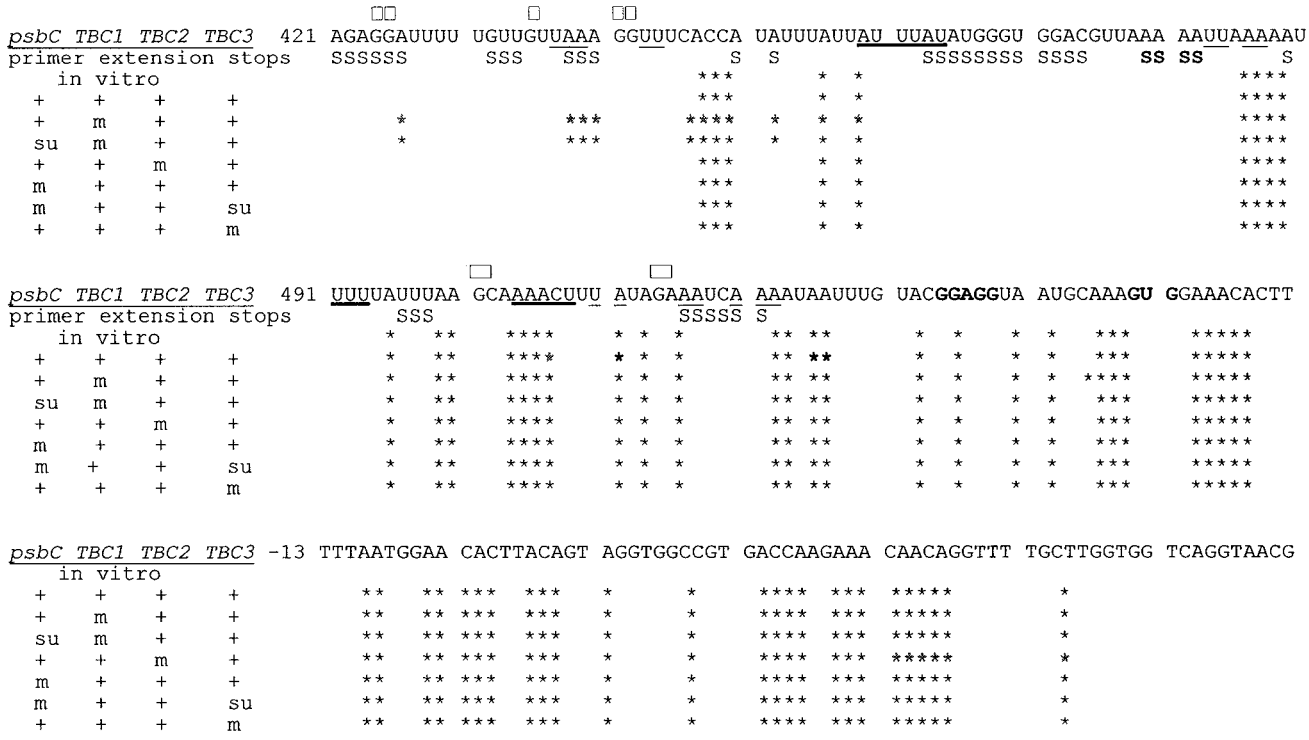


Figure 3.14. Summary of RNA structure mapping results. Results of RNA structure probing experiments are shown with respect to the sequence of the *psbC* 5' UTR. The genotypes of the strains analyzed are indicated by "wt" for wild-type, "m" for a mutation that abolishes translation (e.g. In *tbc1* or *tbc2*, or *psbC-FuD34*) and "su" for a suppressor mutation in either *psbC* (*psbC-F34su1*) or *tbc3*. Asterisks indicate bases methylated by DMS, with grey and black representing weak and strong methylation, respectively. Underlined sequence indicates RNase V1 cleavage sites. Thick underlining represents a stronger susceptibility to RNase V1. Arrows above the sequence indicate G residues that are RNase T1 substrates. Premature RT stops are indicated by "S", with stronger stops having the 'S' in bold text.

3.4. The *psbC-FuD34* and *psbC-F34sul* mutations alter the stability of SL3 and have long-range effects on 5' UTR structure

RNase T1 and RNase V1 digestions were performed on the 5' end-labeled *psbC* 5' UTRs with the *psbC-F34su1* and *psbC-FuD34* sequence to see if these mutations result in any structural changes (Figure 3.3). Differences were observed between the RNase digestion patterns of wild-type and mutant 5' UTRs. On the 5' UTR with the *psbC-F34sul* mutation, Gs at positions 415, 416, 318, 141, and 74 are more susceptible to RNase T1. On the *psbC-FuD34* mutant 5' UTR G318, within the SL3 stem and close to its base, is not susceptible to RNase T1 digestion, although this position is recognized on the wild-type 5' UTR. Conversely, G318 is an RNase T1 substrate on the *psbC-F34su1* 5' UTR and is digested more frequently than on the wild-type 5' UTR (Figure 3.3c). G318 is particularly interesting since *F34su1* mutation, T→A transversion at position 227 is predicted to convert a bulge close to G318 at the base of the stem to an internal loop (Rochaix *et al.*, 1989). (A bulge is where unpaired bases occur on only one strand of a stem while an internal loop occurs where juxtaposed bases on both strands are unpaired). The *psbC-FuD34* mutation corresponds to a deletion of C315 and insertion of two Ts between A309 and A310 (Rochaix et. al. 1989) that removes two bulges and stabilizes the base of the SL3 stem (Rochaix *et al.*, 1989). Differences in RNase susceptibility and DMS methylation pattern in the SL3 is shown in Figure 3.15. Various regions in *psbC* with *FuD34* mutant are less susceptible to RNase V1 digestion compared to the wild type *psbC* 5' UTR. These include nucleotides residues

around A230 and A305 indicating that these nucleotides may not be involved in base pairing in this *psbC*-translation defective mutant.

An extremely strong premature RT stop was detected at C302, A303 and A305 (Figure 3.10). These nucleotides are within the stem of SL3. This premature RT stop was detected on the wild-type and *psbC-F34sul* 5' UTRs although at a much lower intensity. This premature RT stop is so strong on the *psbC-FuD34* mutant 5' UTR that RT stops due to methylated bases 5' to it could not be determined. This site also corresponds with site in the SL3 stem that is not susceptible to RNase V1 in *psbC-FuD34* 5' UTR, which otherwise is cleaved in wild type UTR. Two other premature RT stops which were significantly stronger in *psbC-FuD34* compared to the wild type are at positions 243-244 (Figure 3.11b) and at position 141-145 (not shown). RT stops in *psbC-F34su1*, not seen in wild type UTR are at positions 268-271 and 227-228 (Figure 3.10). These results indicate a structural change in the SL3 that could alter its stability. There could also be a global change in the *psbC* 5' UTR in these mutants.

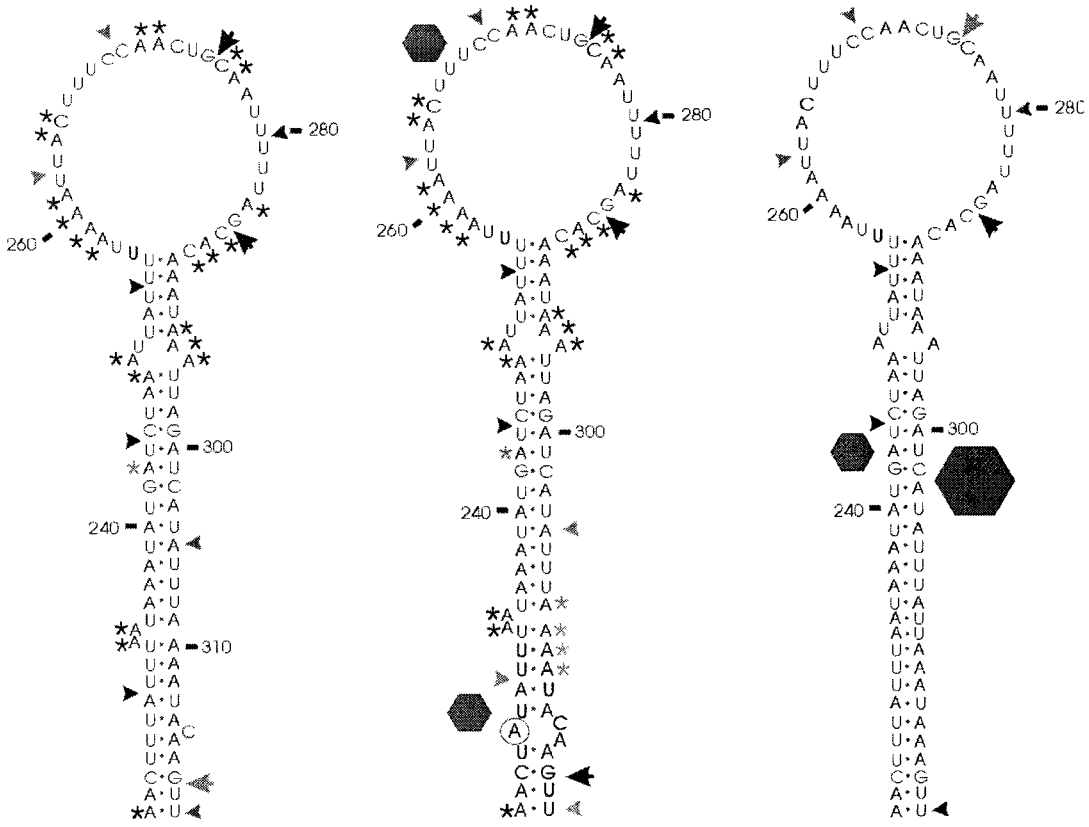


Figure 3.15. *In vitro* RNA structural analyses of SL3 with the wild-type sequence and the two mutations. Asterisks (*) indicate A and C residues that are methylated by DMS *in vitro*. Arrows indicate G residues that are substrates for RNase T1 and thus are not paired. Arrowheads indicated cleavage sites for RNase V1: double-stranded sequences. Most frequently methylated or cleaved positions are indicated by black asterisks, arrows and arrowheads. Less frequently methylated or cleaved sites are indicated in grey. Hexagons indicate locations of strong premature RT stops, with the size proportional to the strength of the block. Due to the strong premature RT stop on the *psbC-FuD34* 5' UTR DMS methylation sites could not be mapped and thus no asterisks are shown.

3.5. No evidence for a minor form of *psbC* mRNA with a 5' terminal extension was obtained in this study.

Primer extension can be used to detect mRNA 5' ends. End-labeled primers are extended from known region on the mRNA close to the 5' end with the enzyme reverse transcriptase. The length of the cDNA produced corresponds to the distance between the primer extension site and the 5' end of the mRNA. Two forms of the mRNAs of *psbA*, *psbB*, *psbD*, *petD* and *rbcL* genes in the chloroplasts have been reported, differing in length at the 5' end (reviewed by Rochaix, 1996; reviewed by Zerges, 2000). The shorter forms are more abundant and are thought to be the processed versions of the longer forms. Primer extensions on *psbC* 5' UTR showed primer extension products longer then the 5' end of *psbC* mRNA as shown in Figure 3.16. This was also observed earlier by F. Vingeault. The longer primer extension products are formed only with the primers that hybridize to regions upstream of the central SL3 and closer to the 5' end of the 5' UTR. The longer product is much less abundant than the mature *psbC* mRNA 5' end but the ends are not consistent between different primers.

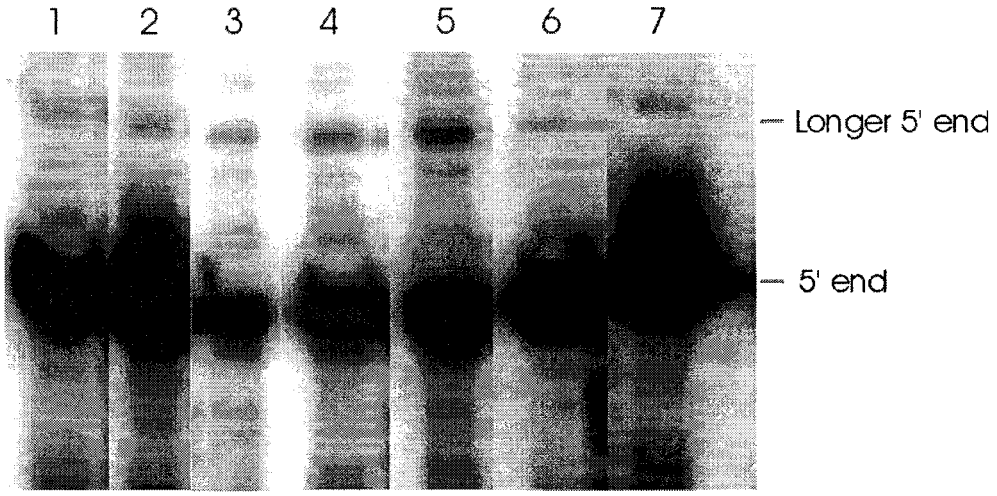


Figure 3.16. 5' end mapping of *psbC* mRNA. Total RNA from *Wild Type* (Lane 1), a *tbc1* mutant strain (Lane 2), a *tbc2* mutant strain (Lane 3), a double mutant strain for *tbc1* and the suppressor mutation in the *psbC* 5' UTR, *psbC-F34sul*; (Lane 4), a strain with a wild-type *psbC* allele and the suppressor mutation in *tbc3* (Lane 5), a mutant for the *psbC* 5' UTR (with *psbC-FuD34*) (Lane 6) and a double mutant for the *psbC* 5' UTR (*psbC-FuD34*) and the suppressor mutation in *tbc3* (Lane 7) were analyzed by primer extension with the primer 'oligo191' that hybridize close to the 5' end of *psbC* mRNA. Longer primer extension products were observed. The 5' end and minor larger 5' end are shown.

3.6 *psbC* 5' UTR Structure May Involves Long-Range Interactions

UV irradiation of nucleic acids results in stable bonds connecting or cross-linking two nucleotides. The susceptibility to UV cross-linking indicates that two nucleotides which may be widely separated in the primary sequences, are very closely apposed in the native RNA (reviewed by Branch *et al.*, 1989). Cross-linked RNA forms circular elements and has reduced electrophoretic mobility when denatured compared to linear RNA. 5' end-labeled *in vitro* transcribed *psbC* 5' UTR with wild type, *F34su1* and *FuD34* sequences were exposed to UV radiation and electrophoresed on denaturing polyacrylamide gel. 5' end-labeled *psbC* 5' UTR without UV exposure was taken as control. Two bands were observed for the UV irradiated samples (Figure 3.17) corresponding to the linear *psbC* 5' UTR and the cross-linked 5' UTR. Further experiments should address the cross-linked nucleotide residues by primer extension and RNase mapping. This will provide information on the nucleotides in *psbC* 5' UTR that are involved in long-range tertiary interactions.

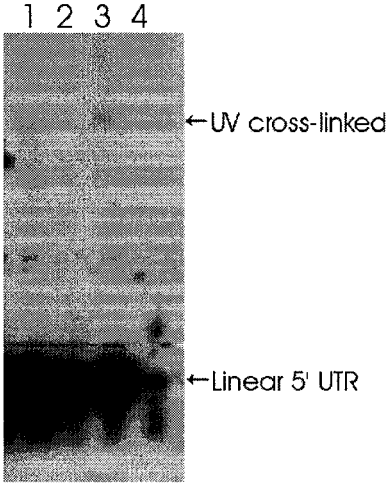


Figure 3.17. UV-induced cross-links in *psbC* 5' UTR suggest a tertiary structure. 5' end-labeled *psbC* 5' UTR with wild type (lane1), *F34su1* (lane 2) and *FuD34* (lane 3) sequences were exposed to UV radiation and a cross-linked product was detected. 5' UTR not exposed to UV (lane 4) was used as control. The 5' UTR and its cross-linked form are indicated by arrows.

Preliminary experiments characterized the circular dichroism (CD) spectra generated by the wild-type and mutant 5' UTRs to in order to reveal information regarding their structure. CD has been used to study secondary structure formation by a variety of RNAs (reviewed by Sosnick et al., 2000). In this study CD spectra were determined from 300 to 215 nm for *in vitro* transcribed *psbC* 5' UTR with wild type, *F34su1* and *FuD34* sequences (shown in Figure 3.18). The CD spectra of *psbC* 5' UTR shows all characteristics of double-stranded A-form RNA: large positive ellipticities at 260-270 nm, a cross-over at around 250 nm, and moderate negative ellipticities at 240 nm (Schmid *et al.*, 1996). CD spectra coupled with temperature ramping is used to study folding transitions in RNA. This becomes very complicated with large RNAs due to presence of multiple structures and the folding transitions and other effects could be superimposed. It would be ideal to dissect the 5' UTR and study the folding behaviour of cis-acting sequence elements separately and in association with each other.

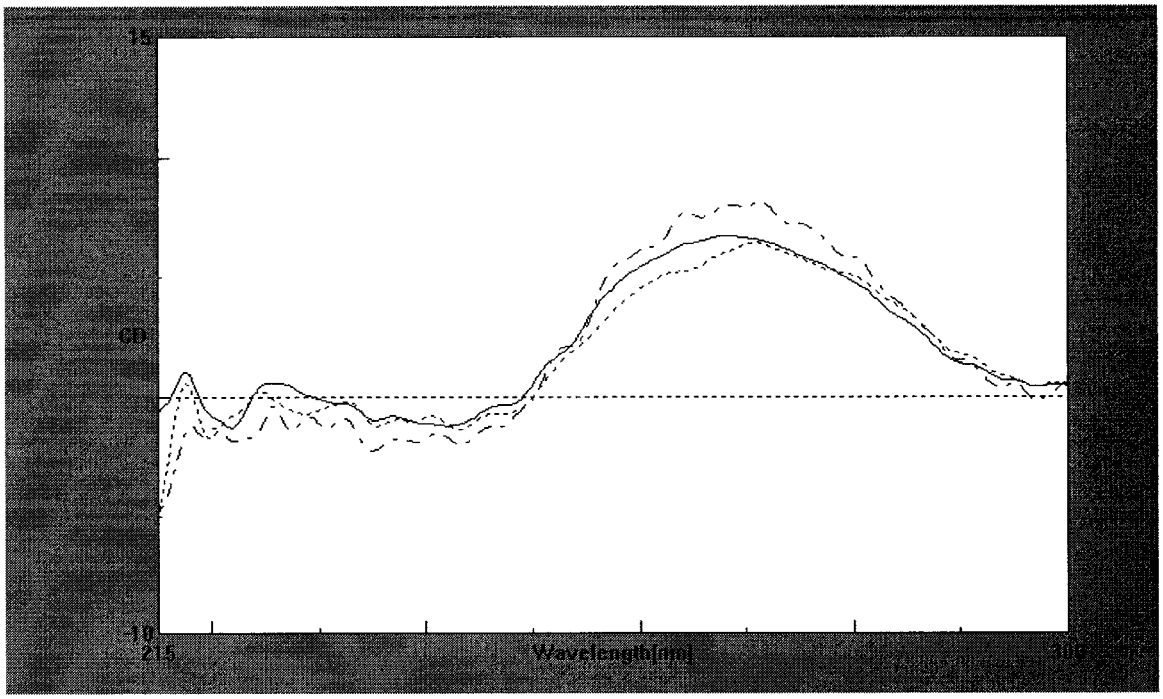


Figure 3.18. CD Spectra for *psbC* 5' UTR with Wild Type, *psbC-F34su1* and *psbC-FuD34* sequences. CD spectra for *psbC-Wild Type* (broken line), *psbC-F34su1* (dotted line) *and psbC-FuD34* (solid line) 5' UTRs, measured from 300 to 215 nm are shown. A double-stranded A-form RNA characteristics with a large positive ellipticities at 260-270 nm, a cross-over at around 250 nm, and moderate negative ellipticities at 240 nm are seen.

4. Discussion

4.1. The *psbC* 5' UTR is highly structured

Chemical and enzymatic probes were used to study RNA structures in the *psbC* 5' UTR. A DMS concentration was used to obtain, on average, one methylation per RNA molecule. Methylated As and Cs were detected by primer extension with 5' end-labeled primers that hybridize to the target RNA. Base-pairing protect bases from methylation by DMS (reviewed by Ehresmann *et al.*, 1987). Thus, the pattern of methylated A and C residues of an RNA *in vitro* reveals information regarding its structure. Methylated A and C residues block reverse transcription generating fragments that correspond to the distance of the methylated base from the 5' end of the primer. Many studies have used [^{32}P] end-labeled RNAs as substrates for DMS protection studies. Although their use would have avoided the high background or premature RT stops which hampered this study, primer extension was used for two reasons. First, the entire 547 base 5' UTR was used because genetic evidence revealed interactions between dispersed RNA sequences (Zerges *et al.*, 2003). [^{32}P] end-labeled subregions of the 5' UTR might not, therefore, form the relevant structure(s). Use of the entire ^{32}P-end-labeled 5' UTR would have precluded analysis of the central region because the cleavages at methylated bases there would generate fragments that would be too long and numerous to resolve by polyacrylamide gel electrophoresis. On the other hand, primer extension allows the use of multiple oligonucleotide primers to characterize the entire intact 5' UTR. Second, primer extension can characterize

DMS methylation patterns generated *in vivo*, where the used of [^{32}P] end-labeled RNAs is precluded. As described below, these *in vivo* DMS methylation patterns provided valuable information regarding the effects of *trans*-acting factors controlled by the three nuclear gene products (see Introduction).

RNase T1 and RNase V1 were used as enzymatic probes of *psbC* 5' UTR structure. Similarly, mild RNase treatment of [^{32}P] end-labeled RNA was performed to obtain, on average, one cleavage reaction per RNA molecule. The fragments generated by primer extension and RNase digestion of the 5' UTR were visualized by denaturing polyacrylamide gel electrophoresis and autoradiography.

The *psbC* 5' UTR is highly structured. The understanding of these structures revealed by this thesis will be extremely valuable in deciphering the molecular mechanisms of translation initiation in chloroplasts and the translational control exerted by the *trans*-acting *tbc1*, *tbc2* and *tbc3* gene products. Little is known about precise molecular mechanisms of translation initiation and its control in organelles and the roles of nuclear gene products therein. Moreover, since most studies have addressed translation control in the eukaryotic cytosol and *E. coli*, new mechanisms could be revealed by studies of organellar translation. Most previous studies have identified translational regulators in biochemical approaches involving *in vitro* translation reconstitution systems. The identification of translational regulators by genetic approaches, such as those used in this long-term project, could reveal new types of factors and molecular mechanisms. It should be emphasized that this study is one of the

most ambitious of its kind. Most of the large 547 5' UTR was characterized both *in vitro* and *in vivo* in several wild-type and mutant strains using multiple approaches. The results are extremely complex due to the complexity of the 5' UTR structure and the apparent concurrent existence of more than one structure formed by certain regions. (A few regions appear to be both single and double stranded, Figure 3.12). Despite these difficulties, the following conclusions could be made.

One of the first indications that the *psbC* 5' UTR is highly structured was its inability to hybridize to DNA or RNA probes larger than 25 bases (data not shown) or to many oligonucleotides of 17-25 bases (Figure 3.4). Moreover, this study found that reverse transcription terminates prematurely at many positions on *psbC* 5' UTR during primer extension reactions. Premature RT termination has been encountered on at least one other chloroplast mRNA, that of the *petD* gene (Higgs *et al.*, 1999). Reverse transcription can be arrested by modified or unusual bases and stable RNA structures (reviewed by Ehresmann *et al.*, 1987). These premature RT stops considerably hampered characterization of the DMS methylation patterns because bands due to these stops hide those due to stops at a methylated base at the same position. On the *psbC-FuD34* mutant 5' UTR, a potent block prevented RT from extending beyond it. That modified bases are responsible for these premature RT stops was ruled out by their detection on an *in vitro* transcribed *psbC* 5' UTR (data not shown). This RNA was synthesized in the absence any of the enzymes that are required to post-transcriptionally modify bases in RNA. The reverse transcription stops are, therefore, probably due to

stable RNA structures. As such, they provided information regarding the location of sequences that could be structured, as described below.

One region within the center of the *psbC* 5' UTR has the potential to form a large stem-loop structure, SL3 in Figure 3.1, which was predicted over a thirteen years ago (Rochaix *et al.*, 1989) and discussed in several subsequent studies (Zerges *et al.*, 2003; Zerges *et al.*, 1997; Zerges and Rochaix, 1994). It was shown to be one of two sequences of the 5' UTR that are absolutely required for translation (Zerges *et al.*, 1997; Zerges *et al.*, 2003). The SL3 apical loop is smaller than the loop proposed by Rochaix *et al.* (1989) because the RNase V1 cleavages at U253-U255 indicate these bases are paired to A289-A291 (Figure 3.1).

This thesis provides the first direct evidence that the large central stem-loop SL3 forms both *in vitro* and *in vivo* (see Results). For example, of the 52 A and C residues in SL3 the DMS methylation results for 34 A and C residues support the structure shown in Figure 3.1; bases predicted to be paired were not methylated and vice versa. (Paired A or C residues at the 3' end of a paired region are known to be accessible to DMS methylation (reviewed by Ehresmann *et al.,* 1987). Infrequent methylation of A243, which is predicted to be paired, partially detracts from the structure. Absence of methylation of three C residues in the loop (C269, C270 and C273,) detracts from the structure. However, C residues are a weaker DMS substrate than A residues. A tertiary structure involving these non-methylated C residues within the context of the SL3 loop is proposed below. A233 is predicted to be paired at the very base of the stem,

although it was methylated. Results of RNase T1 probing of the five G residues in SL3 support its structure, paired Gs were infrequently recognized or not at all. Two Gs in the loop were recognized by RNase T1. Finally, four of the five strong RNase V1 cleavage sites occur at sites that are predicted to be double-stranded. (A possible explanation of the strong RNase V1 cleavage near U280 and two weak cleavage sites in the SL3 loop is proposed below.) The remaining 13 A and C residues are sites of premature RT stops and their methylation status could not be accessed. Thus, this study provides strong evidence that SL3 has a 33 nt apical loop and a 26 bp stem with an internal loop and two bulges, one of 2 A's and the second of a C as shown in Figure 3.1.

Several additional stem-loop structures in the 5' UTR were predicted by *mfold* (Zuker, 1994). Those that are supported by the results of this study are shown in Figure 3.1. Others were excluded by the data and are not shown.

As described above RNase V1 cleavages were detected within the apical loop of SL3, including a particularly strong cleavage site adjacent to U280 and methylation by DMS of three C residues in the loop was not detected. Since, the weight of the evidence supports SL3, these conflicting results suggest that sequences within the loop are base-paired (Figure 3.1). This discrepancy can be reconciled by a pseudoknot tertiary structure formed by base-pairing of the SL3 apical loop to a sequence located elsewhere on the 5' UTR. Three sequences are complementary to the SL3 apical loop and they are shown at the bottom of Figure 4.1, and indicated as PKa, PKb, and PKc.

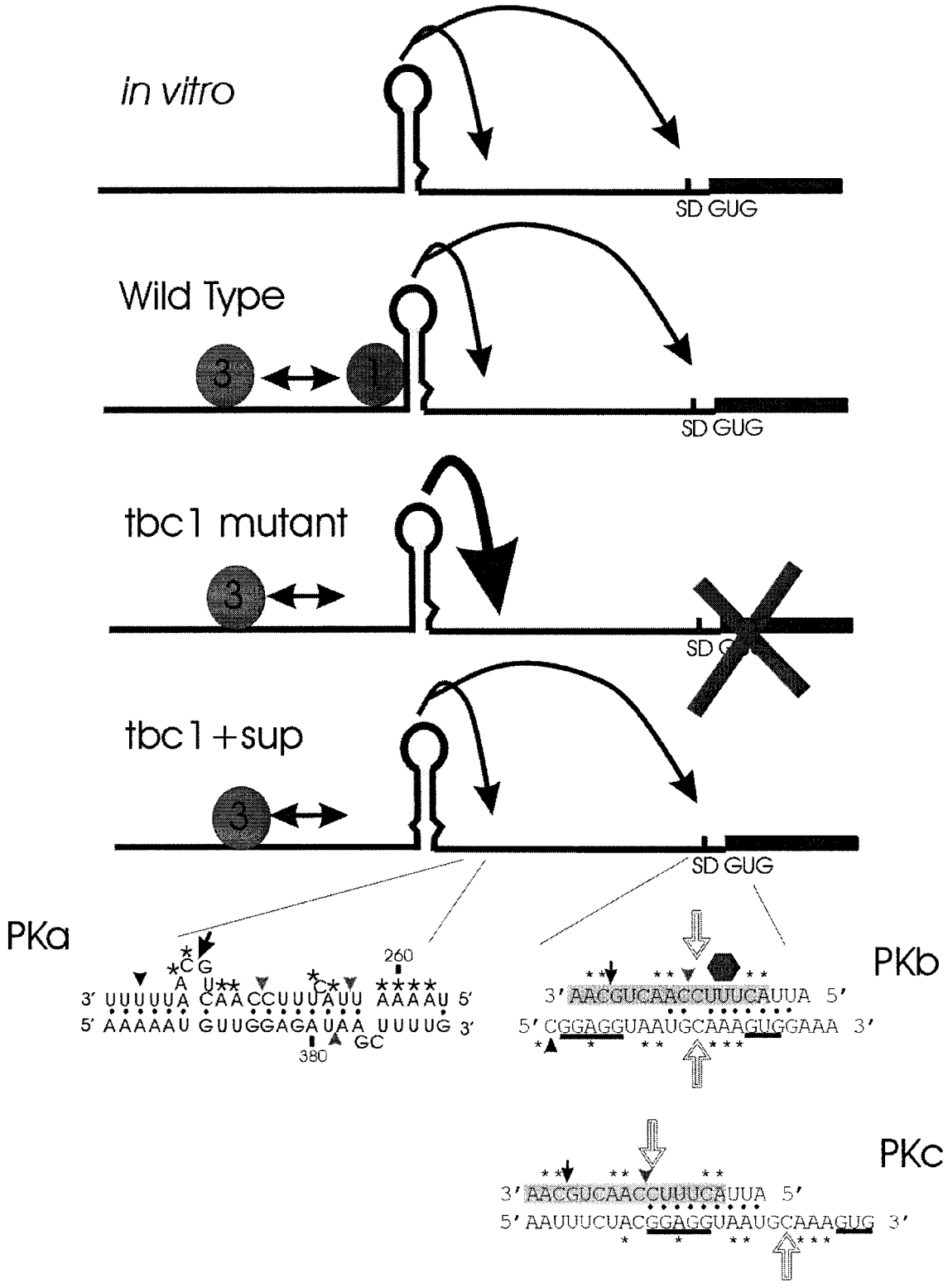


Figure 4.1. Legend in the next page.

Figure 3.17 A tentative model for alternate pseudoknot structures formed by sequences known to be involved in the *TBC1* role in translation. Basepairing of the loop to one of three sites located 3' to it (PKa, PKb or PKc) results in a pseudoknot tertiary structure. The *TBC1* and *TBC3* gene products are shown as red and blue balls, respectively. Asterisks (*) indicate A and C residues that are methylated by DMS *in vitro*. Arrows indicate G residues that are substrates for RNase T1 and thus are not paired. Arrowheads indicated cleavage sites for RNase V1: double-stranded sequences. Most frequently methylated or cleaved positions are indicated by black asterisks, arrows and arrowheads. Less frequently methylated or cleaved sites are indicated in gray. The hexagon indicates location of strong premature RT stop on the *psbC-F34sul* 5' UTR. The SD-like sequence and GUG initiation codon are underlined. In *Wild Type*, pseudoknot structures can form with one of the three different sites. In the *tbc1* mutant, PKb and PKc would not be involved, resulting in methylation of the C residues indicated by white arrows. In this model, *TBC1* is required for formation of a pseudoknot involving PKb or PKc sites in the translation initiation region. The *psbC-F34sul* mutation might restore translation in the presence of the *tbc1* mutation by increasing the flexibility of this stem due to the internal loop in the stem of SL3 generated. This flexibility would restore the ability of SL3 to form a pseudoknot with PKb or PKc in the absence of the wild-type *TBC1* gene product.

PKa is located from 47 to 60 bases 3' to SL3 base (377-393) and is partially within the *TBC1* target identified by a genetic approach (Zerges *et al.*, 2003). Results of the chemical and enzymatic probing experiments are consistent with 14 of the 21 potential basepairs between the SL3 apical loop and PKa (Figure 4.1). Moreover, bases in a bulged that would result from non-paired bases of the apical loop are methylate by DMS and G275 is a substrate of RNase T1. Each of the three RNase V1 cleavage sites in the SL3 apical loop (two are weak, one at U280 is strong) are paired in this pseudoknot structure. Thus, this pseudoknot structure is supported by the results of this study.

The two other sequences complementary to the apical loop of SL3, PKb and PKc, are within the translation initiation region. PKb spans the SD-like sequence (GGAGG). PKc spans SD-like sequence and the GUG initiation codon. The data from the enzymatic and chemical probing experiments do not strongly support the pseudoknot formed by PKb or PKc. Bases that would pair in these pseudoknots are unpaired, based on the DMS mapping results (Figure 3.12). However, it is intriguing that the two C residues methylated only in the *tbc1* mutant, C269 in the SL3 apical loop and C544 in the translation initiation region (indicated by the white arrows in Figure 4.1, would be brought into close proximity if the loop were to hybridize to PKb or PKc. Moreover, a previous study (Zerges *et al.*, 2003) found that the *TBC1* target includes SL3 and the translation initiation region (the SD-like sequence up to and including the GUG initiation codon). These observations led to the tentative model illustrated in Figure 4.1. SL3 might form the alternate pseudoknot structures with PKa and PKb or PKc.

Formation of a pseudoknot involving the translation initiation region would be facilitated by *TBC1* and be required for translation initiation to occur. In the *tbc1* mutant or *in vitro* only the pseudoknot involving PKa would form, thus explaining the *in vitro* evidence for it but not the other structures. The *psbC-F34sul* mutation should increase the flexibility of the SL3 stem by enlarging the internal loop near the base of the stem (see below) and could thereby restore the ability of SL3 to form the pseudoknot with the translation initiation region. The *psbC-FuD34* mutation could prevent the formation of a pseudoknot involving the translation initiation region by increasing the rigidity of the SL3 stem (due to the removal of the two bulges). Although the differences between the wild-type and mutant 5' UTRs in their accessibility to the chemical and enzymatic probes (described above) do not strongly support this model, a rigorous test of these potential pseudoknots and their possible roles in translational regulation would require assessments of the effects of site-directed mutagenesis of the PKa, PKb and PKc sites with and without compensatory mutations of the SL3 apical loop on translation initiation *in vivo*.

Long-range interactions are also supported by the preliminary results of UV induced RNA-RNA cross-linking. A UV cross-linked 5' UTR was observed (Figure 3.17). However, the locations of the cross-linked nucleotides and whether this is a true long-range interactions or not is yet to be ascertained by primer extension and RNase mapping of the cross-linked 5' UTR.

Rochaix *et al.* (1989) proposed that *psbC* translation requires *TBC1* function. A point mutation in SL3, *psbC-F34su1*, suppresses *tbc1* mutation by

destabilizing the base of the stem. They also proposed that *psbC-FuD34* abolishes translation by stabilizing the SL3 stem (due to its removal of the two bulges) and thereby prevents SL3 from melting. Indeed, the structure of the *psbC* 5' UTR is affected by *psbC-F34su1* and *psbC-FuD34*. In particular, the RNase T1 accessibility of G318 provides evidence supporting these hypotheses (Figure 3.15). On the wild-type UTR 5' RNase T1 weakly recognizes G318. This result and the infrequent cleavage between the 3' neighboring U residues by RNase V1 suggest the base of the wild-type SL3 stem can transiently melt, with an equilibrium existing between the base-paired and melted conformations of the SL3 stem base. Three differences between the wild-type and *psbC-F34su1* 5' UTRs were identified which demonstrate the SL3 stem is shifted towards the melted conformation by the point mutation *psbC-F34su1*. First on the *psbC-F34su1* 5' UTR a higher frequency of RNase T1 cleavage was observed adjacent to G318, indicating this base is more frequently unpaired. Second, the *psbC-F34su1* mutant 5' UTR was less frequently cleaved by RNase V1 between A229 and U230. Third, methylation of A309, A310, A311 and A312 was detected on the *psbC-F34su1* 5' UTR, indicating these positions can be unpaired. (Figure 3.15). A strong premature RT stop at position 237 was also detected only on the *psbC-F34su1*, at the position of the U to A point mutation. This premature RT stop and the one observed on wild-type and both mutant 5' UTRs at the opposing positions of the SL3 stem (A314, C315 and A316) prevented a determination of whether the bulge is converted to an internal loop as shown in Figure 3.15. Consistent with the hypothesis that *psbC-FuD34* blocks translation by stabilizing

the base of the SL3 stem (Rochaix et al. 1989), G318 is not susceptible to RNase T1 digestion. (Figures 3.15)

Other evidence indicate that structures other than stem-loops and pseudoknots could be formed by the *psbC* 5' UTR. Of the sequences that are base-paired, based on the chemical and enzymatic probing experiments, several cannot be predicted to form a stable stem-loop secondary structure: no partner sequence that is complementary and also paired sequence could be found. Two potent premature RT stops were observed on the opposing strands of the stem of the central stem-loop structure. These premature RT stops are much stronger on the 5' UTR with the *psbC-FuD34* mutation compared to the wild-type, although they were detected on both. These premature RT stops suggest that one or more RNA structures involve bases at these positions and that the *psbC-FuD34* mutation stabilizes these structures. On the *psbC-FuD34* mutant 5' UTR the premature RT stop in the 3' stem strand of the SL3 (positions 302-305) was so strong that the methylation status of A and C residues 5' to it could not be assessed by primer extension. A third premature RT stop, which is located upstream of SL3 (positions 141-145), was observed earlier by F. Vigneault and again in this study. This premature RT stop occurs just 3' to the loop of SL2 (Figures 3.14 and Figure 3.1). However, the A residues in the loop are not methylated by DMS and the single G in the loop is very weakly susceptible to RNase T1 digestion. These results indicate that either this stem-loop does not form or the bases in the loop are paired. It is tempting to speculate that this loop physically interacts with other sequences and thereby causes the RT stop at this

region. That this premature RT stop is strengthened by the *psbC-FuD34* mutation located approximately 165 nt 3' suggests that the loop interacts with these sequences. Although, *psbC-FuD34* could stabilize the central stem-loop and the interaction involving the upstream loop, the association between the two is difficult to reconcile with the following observation. When reverse transcriptase passes through an RNA structure, the RNA-RNA duplex is replaced by a DNA-RNA duplex and the stem-loop structure is abolished. Thus, one would expect that tertiary interactions involving the stem-loop structure would be eliminated. The 3' most premature RT stop in the central stem-loop (position 302-305) and the upstream stop (position 141-145) can be seen with the same primer extension, hence, they probably are not associated. Nevertheless this interaction cannot be excluded completely, since alternate interactions could be formed. Further investigation of this could be done by studying folding behaviour of the two elements present together on a short RNA fragment with temperature melting and circular dichroism and by site-directed mutagenesis of these sites, followed by primer extension on the mutant RNAs and analyses of their effects on translation in chloroplast transformants (Zerges and Rochaix, 1994; Zerges *et al.*, 1997).

The evidence for a structured *psbC* 5' UTR also comes from its CD spectrum from 300 to 215nm. This spectrum shows the characteristics of a double stranded RNA; large positive ellipticities at 260-270 nm, a cross-over around 250 nm, moderate negative ellipticities at 240 nm (Figure 3.18) (Sosnick *et al.*, 2000). The negative ellipticitiy band at 210 was not seen, probably due to

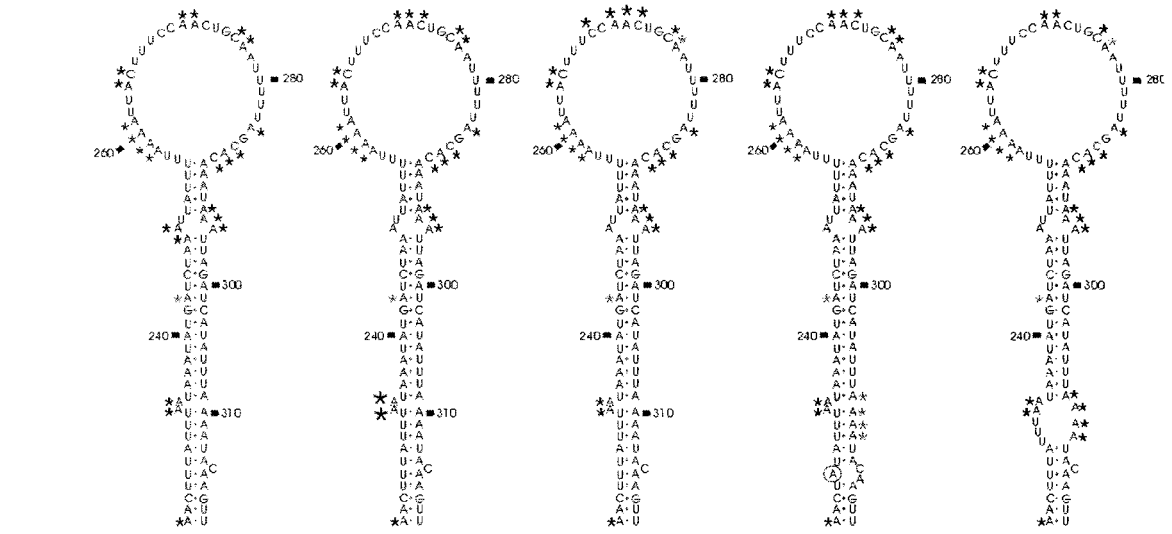
background caused by the buffer in this region of the spectrum. The spectrum for *psbC* 5' UTR with the wild type, *F34su1* and *FuD34* sequence showed similar characteristics and no dramatic differences between them were evident. CD is commonly used to study folding transitions in RNA (reviewed by Sosnick *et al.*, 2000). However, an RNA as large and structured as the *psbC* 5' UTR will show a complex folding behaviour and possibly multiple structures which generate superimposed effects on the spectrum and, therefore, be difficult to resolve. It would be ideal to dissect the 5' UTR and study folding behavior of each element separately and in association with each other. Future work should determine the spectra generated by different subregions of the wild-type and mutant 5' UTRs .

4.2. Translational Regulation of *psbC* 5' UTR by Nucleus-Encoded Factors

As described in the Introduction, a previous study determined the effects of a series of deletions within the *psbC* 5' UTR on its ability to drive expression of a reporter gene *in vivo* (Zerges *et al.*, 2003). One goal of this thesis was to characterize *in vitro* any RNA structures formed by the sequences that could interact with *TBC1*, *TBC2* and *TBC3*. (However, specific sequence targets of *TBC2* were not identified in the previous study.) The second goal was to determine *in vivo* whether any of the mutants affected in *psbC* translation have alterations of the 5' UTR structure or the composition of proteins bound to it. Because DMS can enter cells, it can be used to probe *in vivo* the structure of an RNA and the location of proteins bound to it (reviewed by Ehresmann *et al.*, 1987; Higgs *et al.*, 1999; Klaff *et al.*, 1997). In this thesis, this approach has

proved to be a powerful approach to determine the effects of the mutations that affect translation from the *psbC* 5' leader, both for the nuclear loci and the *psbC* 5' UTR. This work extends one of the most extensive genetic analyses of translational control to the biochemical level. As such, it extends our understanding of the molecular mechanisms of organelle-mRNA specific translational control.

The previous study found two principle target regions of the *TBC1* gene product within the *psbC* 5' UTR; one is SL3 and the 78 bases 3' to it (320-278) and the second is within the translation initiation region from the SD-like sequence to the GUG initiation codon. The *in vivo* methylation pattern in the *TBC1* mutant was extremely interesting in this regard. C269 within the loop of SL3 gains accessibility to DMS in the *TBC1* mutant but it is not accessible in *Wild Type* or any of the mutant strains (Figure 4.2). Neighboring C273 is also hypermethylated in the *tbc1* mutant. In the other *TBC1* target region, the translation initiation region, C554 was methylated in the *TBC1* mutant but not in any of the other strains.



psbC	wt	wt	wt	sup	wt
TBC1		wt	m	m	wt
TBC3	(in vitro)	wt	wt	wt	sup

Figure 4.2. *In vitro* RNA structural analyses of SL3 in wild-type and strains with mutation in *trans*-acting factors, *tbc1* and *tbc2*, and a double mutant with *tbc1* and *psbC-F34su1* mutations. Asterisks (*) indicate A and C residues that are methylated by DMS *in vivo*: single-stranded. Most frequently methylated positions are indicated by black asterisks. Less frequently methylated or cleaved sites are indicated in grey. Encircled is the T→A transversion in *psbC-F34su1* mutation at position 227. Genotypes are indicated at the bottom by 'wt' as wild type, 'm' as mutant defective in translation and 'sup' as suppressor mutation.

Thus, the genetic and biochemical probing experiments indicate that the *TBC1* target sequences are the SL3 and the translation initiation region. This raises the possibility that these regions physically interact in a RNA tertiary structure, as described above. It should be noted that most RNA structures that involve the translation initiation region are inhibitory, however structures that activate translation, such as the one proposed here, have been found in *E. coli* (Wulczyn and Kahmann, 1991).

A sequence complementary to the loop of SL3 was found at positions 374-380 (Figure 4.1). This sequence is within the region identified as containing part of the *TBC1* target sequence (Zerges *et al.*, 2003). Thus, it is interesting to speculate that a second pseudoknot structure can form by hybridization of the SL3 loop to this sequence (as described above and in Figure 4.1). The existence of both pseudoknots should be tested by site-directed mutagenesis. Mutagenesis of positions 374-380 should abolish *TBC1* dependence. Melting profile analyses based on circular dichroism using an RNA containing SL3 and 5' UTR sequences 3' to it might support or refute these structures.

The previous study (Zerges *et al.*, 2003) also found that sequences between positions 379 and 519 are involved in the *TBC1* interaction, but to a lesser extent than SL3 and the translation initiation regions to either side. In the *TBC1* mutant A and C residues between 426 and 459 are hypomethylated relative to *Wild Type* and the other mutant strains analyzed (Figure 3.8 and 3.14). Protection of these bases is consistent with bound protein(s), stabilization of an RNA structure, or both effects. Although sequences in this region have the

potential to form a weak stem-loop structure, results of chemical and enzymatic probing do not support it (Results). RNase V1 cleavages adjacent to the bases that are hypomethylated (437-440) in the *TBC1* mutant (Figure 3.2 and 3.14), however suggest these bases are in double-stranded RNA. Thus, they might be involved in a complex structure that we were not able to predict. Both effects are reversed by the suppressor mutation in SL3, *psbC-F34sul*, suggesting that they are directly mediated by *TBC1* function. More information regarding the structure of these sequences is required.

Alternatively, translational repressors or inactive activators could be bound to this region in the *TBC1* mutant. In this case the wild-type *TBC1* function would relieve this repression or function with a bound activator. This would be significant because the consensus has been that translation of chloroplast mRNAs involves positive regulatory factors, i.e. which stimulate translation, rather than repress it (reviewed by Zerges, 2000). An exception is the autoregulatory feedback repression of *petA* mRNA translation by the encoded polypeptide, cytochrome f (Choquet *et al.*, 1998). It is also possible that a translational activator is inactive and RNA-bound in the absence of the partner *TBC1* product. The premature RT stop within the 3' strand of the stem of SL3 is slightly stronger in *tbc1* mutant than on RNA from Wild Type. This is again the predicted *TBC1* target site (Zerges *et al.*, 2003).

The translational defects of the *TBC1* mutant is partially reversed by a single point mutation, in which the U227 is changed to an A residue (Rochaix *et al.*, 1989; Zerges *et al.*, 1994). Some of the altered DMS methylation patterns in

the *TBC1* mutant are reversed in the suppressed double mutant, i.e., with *F34su1* mutation in the 5' UTR, while others are not. Many altered methylation sites upstream of the central stem-loop overlap between *TBC1* and *TBC3* mutants. The effects described in detail above are reversed by the *psbC-F34sul* mutation, suggesting they are directly involved in the translational defect in the *TBC1* mutant and not secondary effects. As *TBC1* and *TBC3* interact genetically (Zerges *et al.*, 2003), the dispersed hypermethylated sites upstream of the central step-loop in *tbc1* mutant could be due to an effect of the *TBC1* mutation on structures or bound proteins that are dependent on wild-type *TBC3*.

In strains with dominant mutation in TBC3, a greater degree of melting of 3 A-U base-pairs was detected in the SL3 stem indicating melting at that region in the stem of SL3. This mutation in *tbc3* partially suppress *tbc1* mutation affects and the *FuD34* mutation in the stem of SL3 (Zerges *et al.*, 1997; Zerges *et al.*, 2003). Mutation in *rbc3* could restore translation via a similar mechanism as the *psbC-F34su1* mutation.

No conclusive results were obtained regarding *TBC2* target sequences in the *psbC* 5' UTR either with genetic analyses (Zerges *et al.*, 2003) or the *in vivo* DMS probing experiments described here. *TBC2* does promote *psbC* translation through the 5' UTR (Zerges and Rochaix, 1994; Zerges *et al.*, 2003). However, a stretch of nucleotides within the coding region, approximately 50 bases downstream of the initiation codon, is protected from methylation by DMS. This site could be bound by a translational repressor or a stalled ribosome. In the first case, *TBC2* would prevent binding of a translational repressor. The repressor

could be a 46 kDa RNA-binding protein (RBP) that was detected only in high-speed supernatant fractions of lysed chloroplasts from *tbc2* mutant strains (Zerges and Rochaix, 1994)

Several results support an interaction between *TBC1* and *TBC3*, which was suggested previously by the ability of the *TBC3* mutation to suppress a mutation of *TBC1*, but not of *TBC2* (Zerges *et al.*, 1997). First, these mutants share many common hypermethylated bases in the distal region of the 5' leader, between the 5' terminus and the central stem-loop structure (Figure 3.14). Second, this region was shown previously to be required for the interaction with *TBC3*, but not for the interaction with *TBC1* (Zerges *et al.*, 2003). Similarly, in the *TBC3* mutant, changes were detected in the methylation pattern of the SL3, which is not required for the interaction with *TBC3*. Moreover, both *TBC1* and *TBC3* also require sequences close to the initiation codon, suggesting that long-range interactions occur between the more distal targets and the initiation region (see above).

All these data are consistent with previously reported results of studies using genetic approaches (Zerges *et al.*, 1997; Zerges *et al.*, 2003). This study has extended our understanding of *psbC* translational control by identifying specific sequences and RNA structural elements that could be involved. TBC1 interacts with both the stem-loop in the center of the 5' leader and with the sequences immediately 5' to the initiation region. TBC3 interacts with the distal part of the 5' leader and a potential Shine-Dalgarno sequence adjacent to the initiation codon. Together, these *trans*-acting factors and their *cis*-acting targets

function to promote translation initiation. TBC2 appears to function in a distinct pathway because the mutation in it is not suppressed by the mutations in the *psbC* 5' UTR or the TBC3, which suppress the phenotype of the mutation in TBC1. Both pathways, the TBC1-TBC3 pathway and the TBC2 pathway, must function for translation to occur.

A model addressing the roles of TBC1, TBC3 and the *psbC* 5' UTR is presented in Figure 4.1. In this model, an RNA tertiary structure is required for translation. It is proposed to form by interactions between the RNA sequence elements that are required for translation and mediate the trans-acting roles of TBC1 and TBC3: i.e., SL3 and the proximal sequence elements at the initiation region. This structure could be one of the potential pseudoknots described above, although additional experiments are required to address this possibility. Furthermore, folding of the RNA into this tertiary structure is proposed to require the TBC1 product; in *tbc1* mutant this structure does not form and translation is defective. The *psbC-F34sul* mutation in SL3 might restore translation by opening the base of the SL3 stem, thereby, increasing its flexibility and, consequently, allowing formation of tertiary interaction(s). Conversely, the *psbC-FuD34* mutation would block translation by removing the bulged nucleotides within the stem, thereby, changing the properties of the stem such that this tertiary interaction does not occur. This appears to involve more than a stabilization of the SL3 stem. Similarly, in strains with dominant mutation in TBC3, a greater degree of melting of 3 A-U base-pairs was detected in the SL3 stem, suggesting that this mutation also restores translation via a similar mechanism as the *psbC-*

F34su1 mutation. Alternatively, the dominant *tbc3-rb1* mutation might increase an activity that can also promote formation of the tertiary structure in the absence of wild type *tbc1* function, with the resulting force on the SL3 stem casing the melting of the 3 A-U bas-pairs. In the wild type, pseudoknot structures might form with one of the tree different sites (Figure 4.1). In tbc1 mutant, PKb and PKc would not be involved, resulting in methylation of the C residues indicated by the white arrows. In this model, TBC1 is required for the formation of a pseudoknot involving PKb or PKc sites in the translation initiation region. The *psbC-F34su1* mutation might restore translation in the presence of *tbc1* mutation by increasing the flexibility of this stem due to the formation of an internal loop in the stem of SL3. This flexibility would restore the ability of SL3 to form a pseudoknot with PKb or PKc in the absence of the wild type *tbc1* gene product.

4.3. *psbC* Message Could Undergo 5' End-Processing

Stability, and possibly translation, of chloroplast mRNAs in *C. reinhardtii* may require 5' terminal mRNA processing (reviewed by Stern *et al.*, 1998). This model is based in part on the observations that mRNAs of certain chloroplast genes exist in two forms, one being longer at the 5' terminus and being less abundant than the shorter major form (reviewed by Rochaix, 1996 and Zerges, 2000). The longer minor form of *psbA* mRNA is not associated with poly-ribosomes and in the mutants defective in translation of this mRNA there is an increase in the level of longer form (Bruick *et al.*, 1998). However, this longer form was not detected in another study which used a different oligonucleotide in

primer extension reactions (Shapira *et al.*, 1997). In mutants defective in the stability of the *psbD* or *psbB* mRNAs, only the short major form is absent (Nickelsen *et al.*, 1999; Vaistij *et al.*, 2000). These results suggest that only the shorter form is translatable (reviewed by Rochaix, 1996).

As the mRNAs encoding three of the four core polypeptides of the photosystem II reaction center have the dual forms (*psbA, psbB*, *psbD*), a goal of this thesis address whether or not *psbC*, which encoded the fourth core subunit, also has a minor mRNA form with a 5' terminal extension. A longer form of *psbC* 5' UTR was observed by primer extension with primers that hybridize close to the 5' end. However, the longer primer extension product pattern was not consistent between different primers and no definite longer 5' end was detected. As the minor forms of the *psbB* and *psbD* mRNAs are extremely rare, the sensitivity of the primer extension experiments on *psbC* may not have been sufficient to detect such a minor form, if one does exist. Thus, although a longer minor form of *psbC* mRNA might exist no evidence was provided by this study.

The RNA structures and short sequence elements identified in this study will allow specific tests of their possible roles in translation initiation and its regulation by *TBC1*, *TBC2* and *TBC3*. The control of *psbC* translation is extremely complex. In addition to unraveling the molecular mechanisms of this regulation, it will be interesting to determine the physiological basis of such complex regulation, involving at least two distinct pathways and multiple interacting RNA sequence elements.

5. References

Alexander, A., Faber, N., and Klaff, P. (1998). "Characterization of protein-binding to the spinach chloroplast *psbA* mRNA 5' untranslated region." Nucleic Acids Research **26**(10): 2265-2272.

Auchincloss, A. H., Zerges, W., Perron, K., Girard-Bascou, J., and Rochaix, J-D. (2002). "Characterization of Tbc2, a nucleus-encoded factor specifically required for translation of chloroplast *psbC* mRNA in *Chlamydomonas reinhardtii*." Journal of Cell biology **157**: 953-62.

Bhattacharya, D., and Medlin, L. (1998). "Algal phylogeny and the origin of land plants." Plant Physiology **116**: 9-15.

Branch, A. D., Benenfeld, B.J., Paul, C. P., and Robertson, H. D. (1989). "Analysis of ultraviolet-induced RNA-RNA cross-links: A means for probing RNA structure-function relationships." Methods in Enzymology **180**: 418-442.

Bruik, R. K., and Mayfield, S. P. (1998). "Processing of the psbA 5' untranslated region in Chlamydomonas reinhardtii depends upon factors mediating ribosome association." Journal of Cell biology **143**: 1145-53.

Chen, X., Kindle, K. L. and Stern, D. B. (1995). "The initiation codon determines the efficiency but not the site of translation initiation in *Chlamydomonas* chloroplasts." Plant Cell(7).

Chen, X., Simpson, C. L., Kindle, K. L., and Stern, D. B. (1993). "A dominant mutation in the *Chlamydomonas reinhardtii* nuclear gene SIM30

suppresses translational defects caused by initiation codon mutations in chloroplast genes." Genetics **145**: 935-943.

Choquet, Y., Stern, B. D., Wostrikoff, K., Kuras, R., Girird-Bascou, J., and Wollman, F. A. (1998). "Translation of cytochrome f is autoregulated through the 5' untranslated region of petA mRNA in *Chlamydomonas* chloroplasts." Proceedings of National Academy for Science USA **95**: 4380-85.

Conn, G. L., Gutell, R. R., and Draper, D. E. (1998). "A functional ribosomal RNA tertiary structure involves a base triple interaction." Biochemistry **37**(34): 11980-88.

Costa, M., and Michel, F. (1995). "Frequent use of the same tertiary motif by self-folding RNAs." EMBO J **14**(6): 1276-85.

Danon, A., and Mayfield, S. P. (1991). "Light regulated translational activators: identification of chloroplast gene specific mRNA binding proteins." EMBO J **10**: 3993-4001.

Doherty, E. A., and Doudna, J. A. (1997). "The P4-P6 domain directs higher folding of the *Tetrahymena* Ribozyme core." Biochemistry **36**: 3159-3169.

Doudna, J. A., and Doherty, E. A. (1997). "Emerging themes in RNA folding." Folding and Design **2**: R65-R70.

Downs, W. D., and Cech, T. R. (1990). "An ultraviolet-inducible adenosine-adenosine cross-link reflects the catalytic structure of the *Tetrahymena* ribozyme." Biochemistry **29**: 5605-5613.

Draper, D. E. (1996). "Strategies for RNA folding." TIBS **21**: 145-149.

Draper, D. E., and Misra, V. K. (1995). "Protein-RNA recognition." Annual Review of Biochemistry **64**: 593-620.

Draper, D. E., and Misra, V. K. (1998). "RNA shows its metal." Nat. Struct. Biol. **5**(11): 927-930.

Drapier, D., Girard-Bascow, J., and Wollman F. A. (1992). "Evidence for nuclear control of the expression of the *atpA* and *atpB* chloroplast genes in *Chlamydomonas*." Plant Cell **4**: 283-95.

Ehresmann, C., Baudin, F., Mougel, M., Romby, P., Ebel J-P., and Ehersmann, B. (1987). "Probing the structure of RNAs in solution." Nucleic Acids Research **15**: 9109-9128.

Eibl, C., Zou, Z., Bech, A., Kim, M., Mullet, J. and Koop, H. U. (1999). "*In vitro* 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 Journal **19**: 333-45.

Fargo, D. C., Boynton, J.E., and Gillham, N. W. (1999). "Mutations altering the predicted secondary structure of a chloroplast 5' untranslated region affect its physical and biochemical properties as well as its ability to promote translation of reporter mRNA both in *Chlamydomonas reinhardtii* chloroplast and in *Escherichia coli*." Molecular and Cellular Biology **19**(10): 6980-6990.

Fargo, D. C., Hu, E., Boynton, J. E., and Gillham, N. W. (2000). "Mutations that alter the higher-order structure of its 5' untranslated region affect the

stability of chloroplast *rps7* mRNA." Molecular and General Genetics **264**: 291-299.

Fisk, D. G., Walker, M. B. and Barkan, A. (1999). "Molecular cloning of the maize gene *crp1* reveals similarity between regulators of mitochondria and chloroplast gene expression." EMBO J **18**(9): 2621-30.

Fox, T. D. (1996). "Genetics of mitochondrial translation. In translation control." Hershey, J . W. B., Mathew, M. B., and Sonenberg, N., eds., (New York: Cold Spring Harbor Press), pp. 733-758.

Franzetti, B., Carol, P., and Mache, R. (1992). "Characterization and RNA-binding properties of a chloroplast S1-like ribosomal protein." Journal of Biological Chemistry **267**: 19075-81.

Gillham, N. W., Boynton, J. E., and Hauser, C. R. (1994). "Translational regulation of gene expression in chloroplast and mitochondria." Annual Review of Genetics **28**: 71-93.

Girard-Bascou, J., Pierre, Y., and Drapier, D. (1992). "A nuclear mutation affects the synthesis of the chloroplast *psbA* gene product in *Chlamydomonas reinhardtii*." Current Genetics **22**: 47-52.

Gold, J. C., and Spermulli, L. L. (1985). "*Euglena gracilis* chloroplast initiation factor 2. Identification and initiation characterization." Journal of Biological Chemistry **260**: 14897-900.

Goldschmidt-Clermont, M. (1998). "Coordination of nuclear and chloroplast gene expression in Plant cells." Int. Rev. Cyto. **177**(115-180).

Gray, M. W. (1992). "The endosymbiont hypothesis revisited." Int. Rev. Cyto. **141**: 233-357.

Harris, E. H. (1989). "The Chlamydomonas Sourcebook: A comprehensive Guide to Biology and Laboratory Use (Sand Diego, CA: Academic Press)."

Harris, E. H. (2001). "Chlamydomonas as model organism." Annu. Rev. Plant Physiol. Plant Mol. Biol. **52**: 363-406.

Harris, E. H., Boynton, J. E., and Gillham, N. W. (1994). "Chloroplast ribosomes and protein synthesis." Microbiol. Rev. **58**: 700-754.

Harris, H. E. (1998). "Introduction to Chlamydomonas." In The Molecular Biology of Cloroplasts and Mitochondria in Chlamydomonas. Rochaix, J-D., Goldschmidt-Clermont, M., and Merchant, S., eds (Netherlands: Kluwer Academic Publisher), pp 1-11.

Hauser, C. A., Gillham, M. W., and Boynton, J. E. (1998). "Regulation of chloroplast translation." In The Molecular Biology of Chloroplasts and Mitochondria in Chlamydomonas. Rochaix, J-D., Goldschmidt-Clermont, M., and Merchant, S., eds (Netherlands: Kluwer Academic Publisher), pp197-217.

Hauser, C. A., Gillham, N.W., and Boynton, J. E. (1996). "Translational regulation of chloroplast genes: Proteins binding to the 5'-untreanslated regions of chloroplast mRNAs in *Chlamydomonas reinhardtii*." The Journal of Biological Chemistry **271**(3): 1486-1497.

Heifetz, P. B. (2000). "Genetic engineering of the chloroplast." Biochemie **82**: 655-666.

Higgs, D. C., Shapiro, R. S., Kindle, K. K., and Stern, D. B. (1999). "Small cis-acting sequences that specify secondary structures in a chloroplast mRNA are essential for RNA stability and translation." Molecular and Cellular Biology **19**(12): 8479-8491.

Hilbers, C. W., Robillard, G. T., Shulman, R. G., Black, R. D., Webb, P. K., Fresco, R. and Reisner, D. (1976). "Thermal unfolding of yeast glycine transfer RNA." Biochemistry **15**(9): 1874-1882.

Hirose, T., and Suguira, M. (1996). "*Cis*-acting elements and trans-acting factors for accurate translation of chloroplast *psbA* mRNA: development of an in vitro translation system for tobacco chloroplast." EMBO J **15**: 1687-95.

Hirose . T., Kusumegi, T., and 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 Letters **430** (3): 257-260.

Hotchkiss, T. L., and Hollingsworth, M. J. (1999). "ATP synthase 5' untranslated regions are specifically bound by chloroplast polypeptides." Current Genetics **35**: 512-520.

Isel, C., Wsthof, E., Massire, C., Stuart, F. J., Grice, L., Ehresmann, B., Ehresmann, C., and Marquet, R. (1999). "Structural basis for the specificity of the initiation of HIV-1 reverse transcription." The EMBO journal **18**(4): 1038-1048.

Kindle, K. L. (1998). "Nuclear transformation: technology and applications." In The Molecular Biology of Chloroplasts and Mitochondria in

Chlamydomonas. Rochaix, J-D., Goldschmidt-Clermont, M., and Merchant, S., eds (Netherlands: Kluwer Academic Publisher), pp 41-61.

Klaff, P., Mundt, S.M., and Steger, G. (1997). "Complex formation of the spinach chloroplast psbA mRNA 5' untranslated region with proteins is dependent on the RNA structure." RNA **3**: 1468-1479.

Klaff, P., and Gruissem W. (1991). "Changes in Chloroplast mRNA Stability during Leaf Development." Plant Cell **3** (5): 517-529.

Kraus, B. L., and Spermulli, L. L. (1986). "Chloroplast initiation factor 3 from *Euglene gracilis*. Identification and initial characterization." Journal of Biological Chemistry **261**:4781-84.

Krol, A., and Carbon, P. (1989). "A guide for probing native small nuclear RNA and riboprotein structures." Methods in Enzymology **180**: 212-227.

Kuchka, M., Goldschmidt-Clermont, M., Van Dillewijn, J., and Rochaix J-D. (1989). "Mutations at the *Chlamydomonas* NAC2 locus specifically affect the stability of the chloroplast *psbD* transcript encoding polypeptide D2 of the PSII." Cell **58**(869-76).

Kuchka, M., Mayfield, s., and Rochaix J-D. (1988). "Nuclear mutations specifically affect the synthesis and /or degradation of the chloroplast-encoded D2 polypeptide of photosystem II in *Chlamydomonas reinhardtii*." EMBO J **7**: 319-24.

Laing, L. F., Gluick, T. C., and Draper, D. E. (1994). "Stabilization of RNA structure by Mg ions: Specific and non-specific effects." Journal of Molecular Biology **237**: 577-587.

Lisitsky, I., Klaff, P., and Scuster, G. (1996). "Addition of destabilizing poly(A)-rich sequences to endonuclease cleavage sites during the degradation of chloroplast mRNA." Proceedings of National Academy for Science USA **93**: 13398-403.

Lyngso, R. B., and Pedersen, C. N. S. (2000). "RNA pseudoknot prediction in energy-based models." Journal of Computational Biology **7**(3/4): 409-427.

Lyons, A. J., Lytle, J. R., Gomez, J., and Robertson, H. D. (2001). "Hepatitis C virus internal ribosome entry site RNA contains a tertiary structural element in a functional domain of stem-loop II." Nucleic Acids Research **29**(12): 2535-2541.

Martin, W., Rujan, T., Richly, E., Hansen, A., Cornelsen, S., Lins, T., Leister, D., Stoebe, B., Hasegawa, M., and Penny, D. (2002). "Evolutionary analysis of *Arabidopsis*, cyanobacterial, and chloroplast genomes reveals plastid phylogeny and thousands of cyanobacterial genes in the nucleus." Proceedings of National Academy for Science USA **99**(19): 12246-251.

Mathews, D. H., Sabina, J., Zuker, M., and Turner, D. H. (1999). "Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure." Journal of Molecular Biology **288**: 911-40.

Maul, J. E., Lilly, J. W., Cui, L., dePamphilis, C. W., Miller, W., Harris, E. H., and Stern, D. B. (2002). "The *Chlamydomonas reinhardtii* plastid chromosome: Islands of genes in a sea of repeats." Plant Cell **14**(2659-2679).

Mayfield, S. P., Cohen, A., Danon, A., and Yohn, C. B. (1994). "Translation of *psbA* mRNA of *Chlamydomonas reinhardtii* requires a structured RNA

element contained within 5' untranslated region." Journal of Cell biology **127**: 1537-45.

McCarthy, J. E., and Brimacombe, B. (1994). "Prokaryotic translation. The interactive pathway leading to initiation." Trends in Genetics **10**(402-407).

Met, L. J., and Rochaix, J-D. (1998). "Perspectives." In The Molecular Biology of Chloroplasts and Mitochondria in *Chlamydomonas*. Rochaix, J-D., Goldschmidt-Clermont, M., and Merchant, S., eds (Netherlands: Kluwer Academic Publisher), pp 685-703.

Misra, V. K., and Draper, D. E. (1998). "On the role of magnesium ions in RNA stability." Biopolymers **48**(2-3): 113-35.

Moazed, D., Stern, S., and Noller, H. F. (1986). "Rapid chemical probing of conformation in 16 S ribosomal RNA and 30 S ribosomal subunits using primer extension." Journal of Molecular Biology **187**: 399-416.

Murphy, F. L., Wang, Y-H., Griffith, J-D., and Cech, T. R. (1994). "Coaxially stacked RNA helices in the catalytic center of the *Tetrahymena* ribozyme." Science **265**: 1709-1712.

Nickelsen, J., Fleischmann, M., Boudreau, E., Rahire, M., and Rochaix, J-D. (1999). "Identification of *cis*-acting RNA leader elements required for chloroplast *psbD* gene expression in Chlamydomonas." The Plant Cell **11**: 957-970.

Nickelsen, J., fleisheman, M., Bourdeau, E., Rahir, M., and Rochaix J. D. (1999). "Identification of *cis*-acting RNA leader elements required for chloroplast *psbD* gene expression in *Chlamydomonas*." Plant Cell **11**: 957-970.

Palmer, J. D. (1990). "Contrasting modes and tempos of genome evolution in land plant organelles." Trends in Genetics **6**(4): 115-120.

Pestova, T. V., Kolupaeva, V. G., Lomakin, I. B., Pilipenko, E. V., Shtsky, I. N., Agol. V. I., and Hellen, C. U. T. (2001). "Molecular mechanisms of translation initiation in eukaryotes." Proceedings of National Academy for Science USA **98**: 7029-36.

Rochaix, J.-D. (1996). "Post-translational regulation of chloroplast gene expression in *Chlamydomonas reinhardtii*." Plant Molecular Biology **32**: 327-341.

Rochaix, J.-D., Kuchka, m., Mayfield, S., Rahire, S., Girard-Bascou, J., and Bennoun, P. (1989). "Nuclear and chloroplast mutations affect the synthesis or stability of the chloroplast *psbC* gene product in *Chlamydomonas reinhardtii*." The EMBO journal **8**(4): 1013-1021.

Sager, R., and Ramanis, Z. (1968). "The pattern of segregation of cytoplasmic genes in *Chlamydomonas*." Proceedings of National Academy for Science USA **61**(1): 324-331.

Sakamoto, W., Kindle, K. I., and Stern, D. B. (1993). "*In vivo* analysis of *Chlamydomonas* chloroplast *petD* gene expression using stable transformation of beta-glucuronidase translational fusion." Proceedings of National Academy for Science USA **90**(2): 497-501.

Sakamoto, W., Sturm, N. R., Kindle, K. I., and Stern, D. B. (1994). "*petD* mRNA maturation in *Chlamydomonas reinhardtii* chloroplasts: role of 5' endonucleolytic processing." Molecular and Cellular Biology **14**: 6180-86.

Sambrook, J., and Russell, D. W. (2001). "Molecular Cloning: A Laboratory Manual 3rd ed. (New York: Cold Spring Harbor Laboratory)."

Schmid, B., Read, L. K., Stuart, K., and Goringer, H. U. (1996). "Experimental verification of the secondary structure of guide RNA-pre-mRNA chimaeric molecules in *Trypanosoma brucei*." European Journal of Biochemistry **240**: 721-731.

Senecoff, J. F., and Meagher, R. B. (1993). "*In vivo* analysis of plant 18 S ribosomal RNA structure." Methods in Enzymology **224**: 357-373.

Shapira, M., Lerse, A., Heifetz, P.B., Irihimovitz, V., Osmond, C.B., Gillham, N. W., and Boyant, J. E. (1997). "Differential regulation of chloroplast gene expression in *Chlamydomonas reinhardtii* during photoacclimation: light stress transiently suppresses synthesis of the Rubisco LSU protein while enhancing synthesis of the PSII D1 protein." Plant Molecular Biology **33**: 1001-1011.

Silflow, C. D. (1998). "Organization of the nuclear genome." In The Molecular Biology of Chloroplasts and Mitochondria in Chlamydomonas. Rochaix, J-D., Goldschmidt-Clermont, M., and Merchant, S., eds (Netherlands: Kluwer Academic Publisher), pp 25-40.

Sodeinde, O. A., and Kindle, K. L. (1993). "Homologous recombination in nuclear genome of *Chlamydomonas reinhardtii*." Proceedings of National Academy for Science USA **90**: 9199-9203.

Sosnick, T. R., Fang, Z., and Shelton, V. M. (2000). "Application of circular dichroism to study RNA folding transitions." Methods in Enzymology **317**: 393-409.

Stampecchia, O., Girard-Bascou, J., Zanasco, J-L., Zerges, W., Bennoun, P., and Rochaix, J-D. (1997). "A nuclear-encoded function essential for translation of the chloroplast *psaB* mRNA in Chlamydomonas." The Plant Cell **9**: 773-782.

Stein, A., and Crothers, D. M. (1976). "Conformational charges of transfer RNA. The role of Magnesium(II)." Biochemistry **15**(1): 160-168.

Stern, D., Higgs, D., and Yang, J. (1997). "Transcription and translation in chloroplasts." Trends in Plant Science **2**: 308-315.

Stern, D. D. a. D., R. G. (1998). "Chloroplast RNA synthesis and processing." In The Molecular Biology of Chloroplasts and Mitochondria in Chlamydomonas. Rochaix, J-D., Goldschmidt-Clermont, M., and Merchant, S., eds (Netherlands: Kluwer Academic Publisher), pp 165-183.

Stern, S., Moased, D., and Noller. H. F. (1988). "Structural analysis of RNA using chemical and enzymatic probing monitored by primer extension." Methods in Enzymology **164**: 481-489.

Strobel, S. A., and Doudna, J. A. (1997). "RNA seeing double: close-packing of helices in RNA tertiary structure." TIBS **22**: 262-266.

Svab, Z., and Maliga, P. (1993). "High frequency plastid transformation in tobacco by selection for a chimeric *aadA* gene." Proceedings of National Academy for Science USA **90**: 913-917.

Szewczak, A. A., Podell, E. R., Bevilacqua, P. C., and Cech, T. R. (1998). "Thermodynamic Stability of the P4-P6 domain RNA tertiary structure measured by temperature gradient gel electrophoresis." Biochemistry **37**: 11162-11170.

Vaistij, F. E., Gold-Clarmont, M., Wostrikoff, K., and Rochaix, J-D. (2000). "Stability determinants in the chloroplast *psbB/T/H* mRNAs of *Chlamydomonas reinhardtii*." The Plant Journal **21**(5): 469-482.

Wu, H. Y., and Kuchka, M. R. (1995). "A nuclear suppressor overcomes defects in the synthesis of the chloroplast *psbD* gene product caused by mutations in two distinct nuclear genes of *Chlamydomonas*." Current Genetics **27**: 263-269.

Wu, M., and Tinoco, Jr. J. (1998). "RNA folding causes secondary structure rearrangement." Proceedings of National Academy for Science USA **95**: 11555-11560.

Wulczyn, F. G., and Kahmann, T. (1991). "Translational stimulation: RNA sequence and structure requirements for binding of Com protein." Cell **65**: 259-269.

Yael, S. K., and Danon, A. (2002). "The 3' untranslated region of chloroplast *psbA* mRNA stabilizes binding of regulatory proteins to the leader of the message." Journal of Biological Chemistry **277**(21): 18665-669.

Yamaguchi, K., and Subramanian, A. R. (2000a). "The plastid ribosomal proteins. Identification of all the proteins in the 50 S subunit of an organelle

ribosome (chloroplast)." Journal of Biological Chemistry **275**(37): 28466-82.

Yamaguchi, K., von Knoblauch, K., and Subramanian, A. R. (2000b). "The plastid ribosomal proteins. Identification of all the proteins in the 30 S subunit of an organelle ribosome (chloroplast)." Journal of Biological Chemistry **275**(37): 28455-65.

Yohn, C. B., Cohen, A., Danon, A., and Mayfield, S. P. (1998). "A poly(A) binding protein functions in the chloroplast as a message-specific translation factor." Proceedings of National Academy for Science USA **95**: 2238-43.

Zerges, W. (2000). "Translation in chloroplasts." Biochimie **82**: 583-601.

Zerges, W., and Rochaix, J-D. (1994). "The 5' leader of a chloroplast mRNA mediates the translational requirements for two nucleus-encoded functions in *Chlamydomonas reinhardtii*." Molecular and Cellular Biology **14**(8): 5268-5277.

Zerges, W., Auchinclos, A. H., and Rochaix, J-D. (2003). "Multiple Translational Control Sequences in the 5' Leader of the Chloroplast *psbC* mRNA Interact With Nuclear Gene Products in *Chlamydomonas reinhardtii*." Genetics **163**: 895-904.

Zerges, W., Girard-Bascou, J., and Rochaix, J-D. (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*." Molecular and Cellular Biology **17**(6): 3440-3448.

Zuker, M. (1994). "Prediction of RNA secondary structure by energy minimization." Methods in Enzymology **25**: 267-94.