Structural Analyses of the 5' untranslated region of the *psbC* Chloroplast mRNA in *Chlamydomonas reinhardtii*

Frédéric Vigneault

A Thesis

In

The Department

Of

Biology

Presented in Partial Fulfilment of the Requirements

for the Degree of Master of Science in Biology at

Concordia University

Montreal, Quebec, Canada

August 2002

ABSTRACT

Structural Analyses of the 5' untranslated region of the *psbC* Chloroplast mRNA in *Chlamydomonas reinhardtii*

Frédéric Vigneault

Chloroplasts carry out photosynthesis in plant cells and algae, and possess their own genomes. Nuclear encoded gene products synthesized in the cytosol are transported into the chloroplast to control the gene expression. In the green algae *Chlamydomonas reinhardtii*, translation of the *psbC* mRNA, which encodes subunit P6 of the photosystem II, has been shown to be regulated by the *TBC1*, *TBC2* and *TBC3* nuclear products via its 547 nt long 5' untranslated region (5' UTR). The middle region (222-320) is absolutely required for translation and is predicted to form a stable stem-loop structure. The chloroplast mutant *psbC-FuD34* modifies the sequence of this stem-loop and abolishes the translation. It is predicted to increase the stem-loop stability. Another chloroplast mutation, *psbC-F34sul*, partially suppresses the requirement for the *TBC1* factor and is predicted to decrease the 222-320 stem-loop stability. This study did not reveal a precursor 5' end that needs to be processed in *psbC* transcripts, contrary to other mRNAs encoding PSII subunits. Dimethyl sulfate was used *in vivo* and *in vitro* to analyse 5' UTR structures. It did not clearly establish the presence of the 222-320 stem-loop structure, but revealed some structural differences between the wild-type and *psbC-FuD34*. The *in vivo* methylation pattern generally agreed with computer predicted secondary structure. Temperature gradient gel electrophoresis was used to compare the melting profiles of wild type and several mutant 5' UTRs. Important differences in the unfolding pattern exist between the wild type, *psbC-FuD34* and deletion mutants near the middle stem-loop. Moreover, more than one conformation can exist at the same conditions. This study suggests that structural interactions with the 222-320 stem-loop play an important role in the translation of the *psbC* mRNA.

ACKNOWLEDGEMENTS

I was looking for something that will be different from ecology when I have received an email from a certain Bill Zerges that was looking for students with good potential to work on factors affecting the translation in the chloroplast of a small green alga...

I first of all thank my supervisor Dr. William Zerges for giving me the opportunity to discover cellular and molecular biology and for all his patience. I also want to thank my committee members Dr. Claire Cupples and Dr. Muriel Herrington. Special thanks to my lab-mates Shengwu, Dana, Munir and Hamid (in order of appearance) for three good years. I thank all my friends and especially those of you who actively participated in the "entertainment part" of a Graduate degree, and finally to Éliane for her support during those past seven years.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

1.1 *Chlamydomonas reinhardtii*

According to systematic and evolutionary studies, the green algae lineage is more closely related to plants than any other algal group (reviewed by Mets and Rochaix, 1998). *Chlamydomonas reinhardtii* is a unicellular green alga whose photosynthetic apparatus closely resembles that of land plants and is used as a plant cell model (reviewed by Weeks, 1992). The cells are oval shaped and are about μm long with two flagella (of about 10 μm long also). Nearly half of the cell volume is occupied by a single cup shaped chloroplast. The nuclear genome has seventeen linkage groups and is estimated to be 100 Mbp. Sequencing of the nuclear genome is underway and due to be completed by the end of 2002. The *C. reinhardtii* chloroplast genome has recently been entirely sequenced. It is circular, consists of 203 337 bp, contains 34 genes involved in photosynthesis, 31 genes involved in chloroplast transcription and translation, 29 tRNA genes, 1 protease gene and 9 genes of unknown functions (Maul *et al*, in press; Rochaix, 2001b). When supplied with a reduced carbon source, such as acetate, *C. reinhardtii* is able to grow heterotrophically without any photosynthetic function (reviewed by Harris, 1998; Lefebvre and Silflow, 1999). Therefore, unlike vascular plants, *C. reinhardtii* mutants can be obtained which are deficient in photosynthesis, because they can be maintained on media containing acetate. Moreover, *C. reinhardtii* haploid cells have two different life cycles. Usually, the vegetative cells undergo mitosis. However, upon nitrogen starvation cells differentiate into gametes, and opposite mating-type (mt+ and mt-) will combine in a zygote, which will then undergo meiosis producing a tetrad of four

haploid spores (reviewed by Harris, 1998; Rochaix, 2001b). The nuclear genome is inherited in a Mendelian fashion while the chloroplast genome is transmitted uniparentally from the mt+ partner (while the mitochondria genome is inherited from the mt-) (reviewed by Harris, 1998; Rochaix, 2001b). Chloroplast mutants can easily be obtained by homologous recombination, which causes the replacement of the endogenous DNA with the introduced one, or following chemical mutagenesis. Nuclear mutants are obtained by screening and selection following mutagenesis, and genes can be rescued by complementation (reviewed by Mets and Rochaix, 1998; Simpson and Stern, 2001). Photosynthetic mutants, particularly those defective in photosystem II (PSII), can be identified by their high fluorescence profile (reviewed by Bennoun and Beal, 1998; Joliot *et al*, 1998). *C. reinhardtii* can be manipulated and characterized using genetic, molecular, biochemical and microbial techniques and therefore represent a system of choice to study the assembly of the photosynthetic apparatus and the expression of the genes encoding its components.

1.2 Nucleus-encoded factors coordinate expression of chloroplast genes

The chloroplast is an essential organelle for plants and algae. It performs photosynthesis, which is required not only for the photoautotrophic growth of the plant, but also as the primary source of energy and oxygen for many organisms. These organelles also produce amino acids and lipids, and carry out essential steps in the assimilation of C, N, S, and P. Chloroplasts are thought to have evolved from a single ancestor related to cyanobacteria by an association between this photosynthetic ancestor within his host cell, by an event called endosymbiosis that occurred one or two billion

years ago (reviewed by Sugiura *et al*, 1998). Chloroplasts evolution occurred following successive secondary symbiosis events, which involved the massive transfer of novel genes of bacterial origin to the nuclear genome of the host cell (reviewed by Barkan and Goldschmidt-Clermont, 2000; Cavalier-Smith, 2000). The expression of these genomes and the biochemical pathways they encode have become tightly integrated within the cell (Zerges *et al*, 1997). This theory is strongly supported by the similarity of the gene sequences and arrangements between the actual chloroplast of plants and cyanobacteria (reviewed by Sugiura *et al*, 1998). Complex signalling pathways have evolved to coordinate the chloroplast genome expression in response to environmental factors such as light intensity (reviewed by Zerges, 2000). Current day chloroplasts contain semi-autonomous circular genomes and the complete transcriptional and translational machinery, which are distinct from those of the nucleus-cytosol (reviewed by Zerges, 2000). Indeed the biosynthesis and assembly of the photosynthetic apparatus rely essentially upon nuclear regulatory factors, synthesized by 80S ribosomes in the cytosol and subsequently imported into chloroplasts, where they control plastid gene expression. Many of these factors act in the expression of individual chloroplast genes by regulating mRNA stability or translation. This coordinated interaction occurs mainly at the post-transcriptional level (Figure 1) (Nickelsen, 2000; reviewed by Rochaix, 1996; Stern *et al*, 1997; reviewed by Zerges, 2000).

Both chloroplast and nuclear mutants deficient in photosynthetic activity have been analyzed in *C. reinhardtii* (Rochaix *et al*, 1989; Zerges and Rochaix, 1994). Nuclear mutations affecting photosynthesis have been located within genes that encode structural components of the photosynthetic apparatus or factors affecting the expression of

chloroplast genes (Rochaix, 1996; Rochaix *et al*, 1989). Moreover, analysis of mutants affected in the translation of specific chloroplast mRNAs have revealed a large number of nuclear loci involved in chloroplast gene expression and that the interactions between the nuclear and chloroplast genetic systems are rather complex (reviewed by Barkan and Goldschmidt-Clermont, 2000; Mets and Rochaix, 1998). In addition, the expression of one chloroplast gene can be regulated by several nuclear factors and at multiple levels (reviewed by Rochaix, 1996; Sugiura *et al*, 1998). A striking example is the expression of the chloroplast *psaA* mRNA of *C. reinhardtii* (encode the 83 kDa PsaA subunit of the photosystem I reaction center), which requires the products of at least 14 nuclear loci (reviewed by Rochaix, 2001a). Genetic and molecular analysis of nuclear mutants of several green algae, and from higher plants, indicates that the majority of these nuclear factors interact with the 5' and 3' untranslated regions (UTRs) of chloroplast mRNAs, which correspond to the sequences of the mRNA between the coding sequence and the 5' and 3' termini respectively. These factors are mRNA specific, but are not necessarily sequence specific and could therefore require other specific *cis*-elements and RNA structures that can form. For example, three nuclear loci (*TBC1*, *TBC2* and *TBC3*) have been shown to interact with the *psbC* 5' UTR mRNA in *C. reinhardtii* (encode the subunit P6 of the PSII) (Zerges *et al*, 1997; Zerges and Rochaix, 1994). A 47 kDa protein interacts with the *psbA* 5' UTR (encode subunit D1 of PSII) (Danon and Mayfield, 1991), and a 40 kDa protein binds *psbD* 5' UTR (encode subunit D2 of PSII) (Ossenbuhl and Nickelsen, 2000). These *trans*-acting regulators of translation have been identified with genetic approaches, and analysis of their biochemical functions is just beginning (Barkan *et al*, 1994; Danon, 1997; Higgs *et al*, 1999; Hirose *et al*, 1999; Hirose

and Sugiura, 1996; Hsu-Ching and Stern, 1991; Levy *et al*, 1999; Rochaix *et al*, 1989; Sugiura *et al*, 1998; Zerges and Rochaix, 1998). Until now, most factors involved in translation have been identified with biochemical approaches using *in vivo* translation systems. An *in vitro* translation system from isolated tobacco chloroplast has been developed that facilitate the detailed analysis of factors and mechanisms involved in translation that were not revealed by biochemical experiments (Hirose and Sugiura, 1996). This system is dependent of the added mRNA and seems to be reliable compared to *in vivo* data.

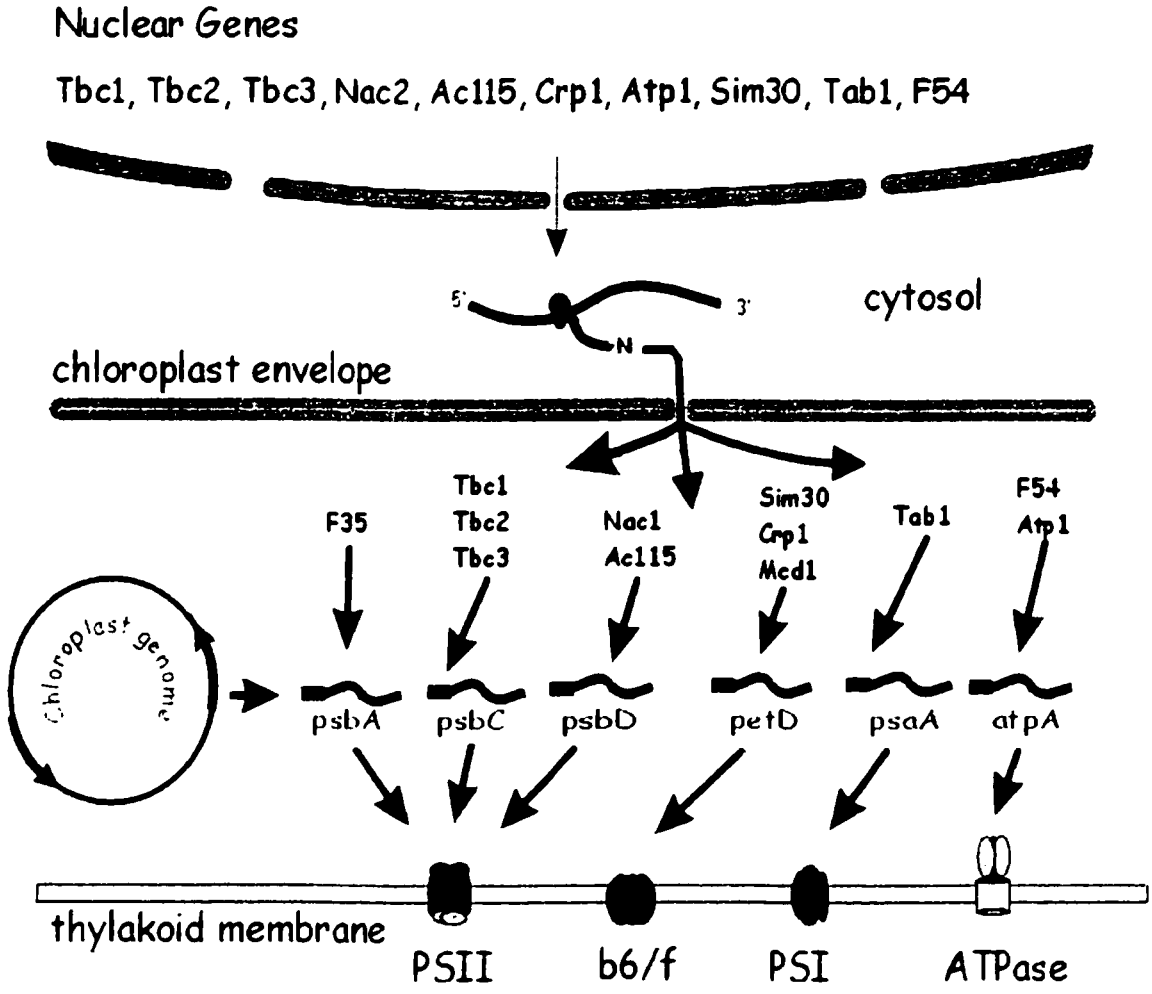

Figure 1. Nuclear encoded gene factors control the translation of chloroplast mRNAs. These factors are synthesized by the cytosolic ribosomes and imported in the chloroplast were they interact specifically with their target mRNA, usually within its 5' untranslated region (box at 5' end of each chloroplast mRNA). Chloroplast mRNAs are transcribed from the chloroplast genome by its own plastid stromal machinery and each of them encodes one subunit of the photosynthetic apparatus (Zerges, 2000).

1.3 Photosystem II assembly

The photosynthetic apparatus is composed of several multisubunit complexes – photosystem I, photosystem II (PSII), cytochrome b_6f, and ATP synthase. These complexes are embedded in the thylakoid membrane and consist of both nucleus and chloroplast genome-encoded proteins (reviewed by Barkan and Goldschmidt-Clermont, 2000; Zerges *et al*, 1997). Of these complexes, PSII is of particular interest, since it is the unique enzyme complex that can oxidize water using light as the energy source and therefore, represents the first reaction in the photosynthetic electrons transport chain and the source of most bimolecular oxygen for respiration. It is also the site of action of potent photosynthesis herbicides, such as triazine, and it is subject to photoinhibition due to damage caused by high light intensity (Rutherford, 1989; Zerges *et al*, 1997). The PSII core reaction center is comprised of more than twenty-five polypeptides and is surrounded by a light-harvesting complex, which are embedded in the thylakoid membrane (reviewed by Redding and Peltier, 1998). The oxygen-evolving complex (OEC) is located in the lumen, and is associated with PSII, peripherally associated with the thylakoid membrane.

At least 16 subunits, including the central most hydrophobic proteins of the reaction centre, are encoded by chloroplast genes and synthesized in the organelle while the rest are encoded by nuclear genes and translated in the cytoplasm, and then imported into the chloroplast to the thylakoid (reviewed by Erickson, 1998; Innis *et al*, 1988; Kapazoglou *et al*, 1995; van Wijk *et al*, 1997). The PSII reaction center includes subunits D1, D2, P5 and P6 (the two chlorophyll binding proteins CP47 and CP43 respectively in higher plants) encoded by chloroplast genes *psbA*, *psbD*, *psbB* and *psbC* respectively, and the

cytochrome b_{559} (cyt b_{559} is encoded by *psbE* and *psbF*). (de Vitry *et al*, 1989; reviewed by Erickson, 1998; Rochaix *et al*, 1989). Analysis of the PSII assembly in *C. reinhardtii* revealed that the subunits are synthesized and inserted into the assembling complex in a specific order (de Vitry *et al*, 1989; reviewed by Erickson, 1998; reviewed by Zerges, 2002). The subunits D1, D2 and P5 are synthesized in a concerted manner, while P6 synthesis is independent. It was proposed that D2 might serve as a receptor for D1. In the absence of D2, D1 synthesis is severely compromised. In a similar way, absence of D1 impairs the synthesis of P5 (de Vitry *et al*, 1989; van Wijk *et al*, 1997). Moreover *psbE* null mutants affecting cytochrome b_{559} synthesis also prevent accumulation of D1, D2 and CP47 subunits and greatly reduce that of CP43, which means that cyt b_{559} is probably involved in PSII assembly at early stage (Morais *et al*, 1998).

1.4 Translation initiation in chloroplast

In every organism, protein products are synthesized by ribosomes that read the genetic information contained in mRNAs by a process called translation. mRNAs are not only simple copies of genes, but also play an active role in translational regulation, initiation and termination. In prokaryotes, translation initiation occurs in the cytosol (Figure 2a). The S1 protein binds the 30S small ribosomal subunit to activate it. This is done in the presence of the initiation factors IF-3, which promotes the dissociation of the 50S ribosomal subunits, while IF-1 prevents binding of any tRNA at the aminoacyl acceptor-site and also promotes mRNA binding (reviewed by Kozak, 1999). The 30S subunit binds to the Shine-Dalgarno (SD) sequence consensus (GGAGG). In *Escherichia coli*, almost all mRNAs have the SD sequence located at 7±2 nt 5' to the AUG or GUG

translation initiation codon (McCarthy and Brimacombe, 1994). The SD sequence is complementary to the 3'end anti-SD sequence of the 16S rRNA of the 30S subunit and the two interact by base-pairing. The AUG initiation codon is present in 91% of the *E. coli* mRNAs, while GUG and UUG are also used for 8% and 1% of the mRNAs respectively (reviewed by Makrides, 1996). The 30S small subunit of the ribosome first interacts with the initiation codon. Bound IF-2 aids in the selection of the initiator $tRNA_f^{Met}$ (to the peptidyl-site). Binding of GTP (guanosine 5'-triphosphate) forms the ternary complex. Then the 50S subunit assembles with the 30S subunit to form the 70S ribosome of the initiation complex (reviewed by Kozak, 1999). Studies have shown that it really is the 30S subunit that recognizes the start codon and not one of the initiation factors (Wang *et al*, 1989). *Trans*-acting factors and *cis*-elements usually modulate the translation initiation by preventing the binding of the small ribosomal subunit to the translation initiation region, either through inhibitory RNA structures, bound regulatory proteins, or both. Stem-loop structures within the 5' UTR have been shown to conceal the SD sequence or the initiation codon hindering their accessibility to the 30S subunit (de Smit and van Duin, 1990).

Translation initiation in eukaryotes involves a similar process (Figure 2b). Instead of a SD sequence, the 40S small ribosomal subunit is recruited by initiation factors bound to the 5'm^7G cap structure. The small subunit then scans in a 3' direction to the first AUG in the proper sequence context for initiation. No GUG initiation codon has been reported to date in eukaryotes (Bonham-Smith and Bourque, 1989). There the $tRNA_i^{Met}$ binds to peptidyl-site of the 40S subunit, followed by the binding of the 60S large ribosomal subunit to complete the initiation complex. As in prokaryotes, at least 10 eukaryotic

initiation factors (eIF) are required for initiation (reviewed by Kozak, 1999). For example, eIF-4E is the cap binding protein that recruits the 40S subunit. During the scanning process, the mRNA is kept unstructured by eIF-4A and eIF-4B which have RNA helicase unwinding activity (reviewed by Pain, 1996). The ternary complex formation is catalyzed by eIF-2 which binds GTP and $tRNA_i^{Met}$, while the eIF-3 directs the binding of the 60S subunit to form the 80S ribosome complex (Rozen *et al*, 1990). Poly(A)-binding proteins (PABP) both bind to the poly-A sequence at the 3' end and interact with eIF-4G on the 5' m^7G cap. This suggests an interaction between the 5' and 3'ends during the initiation process, at least in yeast and plants (Craig *et al*, 1998). Shortening of the poly(A) tail and successive decapping was shown to lead to 5'->3' degradation in several yeast transcripts (Muhlrad *et al*, 1995).

Chloroplast translation systems are closely related to those of prokaryotes, reflecting their common evolutionary origins. Nevertheless, there are also many differences, which could have some similarities with the eukaryotic cytosolic systems (Figure 2d). Chloroplast ribosomes are similar to their eubacterial counterparts both in size and sequences (reviewed by Barkan and Goldschmidt-Clermont, 2000). The chloroplast homologues of the 70S ribosome include the 30S and 50S subunits with their respective rRNA 16S, 5S and 23S. Identification of all the protein component of the 30S and 50S chloroplast ribosomal subunits has been achieved (Yamaguchi and Subramanian, 2000; Yamaguchi *et al*, 2000). Twenty-one orthologues to *E. coli* 30S ribosomal subunit and thirty-one to the 50S ribosomal subunit were found. In addition, four proteins were found to be specific to 30S chloroplast ribosome subunits and two to the 50S. Two initiation factors homologous to IF-2 and IF-3 have been characterized in *Euglena gracilis* (Gold

and Spremulli, 1985; Kraus and Spremulli, 1986). A chloroplast gene coding for an IF-1 homologue has been found in spinach (reviewed by Stern *et al*, 1997). A spinach homologue of S1 has been found associated with the 30S subunit and requires a *cis*-element in *psbA* (Franzetti *et al*, 1992). The chloroplast translational machinery also includes tRNAs ($tRNA_f^{Met}$ is also used as the initiator tRNA), aminoacyl tRNA synthases, elongation factors, ribosome release factors, and several others which are similar to their counterparts in eubacteria (Sugiura *et al*, 1998; Wang *et al*, 1989).

Chloroplast 16S rRNA anti-SD sequences are highly conserved (Bonham-Smith and Bourque, 1989). But the presence of a SD-like sequence in chloroplast mRNAs does not seem to be as prevalent as it is in prokaryotes (reviewed by Zerges, 2000). SD-like sequence within 100 nt upstream of the initiation codon have been found in only a third of the chloroplast transcripts studied to date (Bonham-Smith and Bourque, 1989). Among the seventy-nine chloroplast genes in tobacco, thirty-eight have a SD-like sequence within the first 18 to 6 nt from the initiation codon (reviewed by Sugiura *et al*, 1998). Chloroplast transformation in *C. reinhardtii* has allowed the use of site directed mutagenesis to study the SD-like sequences *in vivo*. Deletion of SD-like sequences from *petD* (GGA, -12 to -10), *atpB* (GGAGG -115 to -111), *atpE* (GAAG, -22 to -19), *rps4* (GAAG, -149 to -146) and *rps7* (GGA, -116 to -114) mRNAs did not affect the translation, neither did insertion of the SD consensus sequence (GGAGG, -9 to -5) enhance the expression (Fargo *et al*, 1998; Sakamoto *et al*, 1994). Replacement of the SD-like sequence in *psbD* (GGAG, -14 to -11) by AAAG had no effect while CCUC reduced the accumulation of D2 to 25% of the wild type level (Nickelsen *et al*, 1999). Similar results were obtained for *psbC* (GGAGG, -14 to -10) (Zerges *et al*, submitted).

Mutation of the *psbA* SD-like sequence (GGAG, -31 to -28) is ambiguous since it abolished translation, but the mRNA level was also strongly reduced (Mayfield *et al*, 1994). It is interesting to note that the tobacco *psbA* does not require a SD-like sequence, although deletion of sites complementary to the 3' terminus of 16S rRNA significantly reduced translation (Hirose and Sugiura, 1996). Another interesting fact is that neither the deletion nor the replacement of the SD-like sequence in *atpB*, *atpE*, *rps4* and *rps7* affected their translation in *E. coli* (Fargo *et al*, 1998).

The initiation codon of chloroplast mRNAs is usually AUG and in some cases GUG. Among the seventy-nine chloroplast genes encoding proteins in tobacco, only two mRNAs, *rps19* and *psbC*, have GUG as the initiation codon (Sugiura *et al*, 1998). In the *C. reinhardtii* chloroplast, the *psbC* mRNA also uses GUG as initiation codon (Rochaix, 1996). Mutation analysis allowed researchers to study the importance of the initiation codon. For instance, changing the initiation codon of the *psbA* mRNA of tobacco from AUG to ACG abolished its translation, while the replacement by a GUG reduced translation to 10% of the wild type level (Hirose and Sugiura, 1996). Changing the *petD* initiation codon from AUG to AUU or AUC decreased translation by five fold (Chen *et al*, 1993). In *petA* (encode cytochrome *f* of the cyt b_6f) changing the AUG to ACC, ACU, ACG, or AUU severely affected synthesis (from 1% to 20% of the wild type level), and changing to UUC prevented translation (reviewed by Sugiura *et al*, 1998). These studies also revealed that the initiation codon controls the efficiency of the translation, but not the site of initiation (Chen *et al*, 1995).

Chloroplast mRNAs are not capped *in vivo* with 5'm^7G (reviewed by Danon, 1997). Poly(A)-rich sequences have been found in a few chloroplast mRNAs, but they

do not represent the long poly(A) tail found in cytoplasmic mRNAs (Lisitsky *et al*, 1996). Chloroplast poly-adenylation seems to target the mRNA for rapid degradation as it does in eubacteria (Rott *et al*, 1998; reviewed by Sugiura *et al*, 1998). A 47 kDa protein, homologue to a PABP, has been reported to be essential for translation of the *psbA* mRNA in *C. reinhardtii* and mediations of its translational activation by light regulation through binding to the 5' UTR (Yohn *et al*, 1998). Meanwhile the 3' and 5' ends of the *psbA* mRNA of tobacco might interact together, since the replacement of the *psbA* 3'UTR by that of the *rpl32* mRNA (chloroplast 50S ribosomal protein L32) decreased the synthesis of a reporter gene (Eibl *et al*, 1999). Scanning has been proposed to occur over 30 nt on the 5' UTR of the *psbA* mRNA of barley, but evidence against long range scanning has been reported on the 5' UTR of the *petD* mRNA (encode the subunit IV of the cytochrome b_6f complex) (Drager *et al*, 1999; reviewed by Zerges, 2000).

Many chloroplast mRNAs are thought to be processed to give their mature 5' ends (reviewed by Monde *et al*, 2000; Rochaix, 1996). Processing allows both the accumulation of stable transcripts and their translation to occur. While it is sometimes carried out by splicing, a combination of 5'->3' exoribonucleases that remove nucleotides from the 5' end and endoribonucleases that cleave internally, seems to be the more common pathway and may be the exclusive pathway for chloroplast mRNAs processing and degradation in *C. reinhardtii* (Drager *et al*, 1998; Rott *et al*, 1998; Stern and Gruissem, 1987). Only *psaA*, *psbA* and 23S rRNA genes contain introns, and the last two are able to self-splice (Drager *et al*, 1999; reviewed by Monde *et al*, 2000). A *cis*-element within the first 24 nt of the *petD* 5' UTR (362 nt) forming a stem-loop structure has been shown to be required for mRNA processing and stability, and interaction with

the *MCD1* nuclear gene product protecting the mRNA from 5'->3' degradation, thereby stabilizing it (Esposito *et al*, 2001; Higgs *et al*, 1999). Some of the chloroplast mRNAs encoding PSII subunit polypeptides have been found to accumulate two forms of their transcripts, which vary at their 5' termini. One form has additional sequences and thus, is longer. For example, the *psbA* mRNA has two forms, with termini corresponding to nucleotides located -90 and -36 bases from the AUG initiation codon (Bruick and Mayfield, 1998). *psbB* has mRNAs extending to positions -147 or -35 (Vaistij *et al*, 2000), and *psbD* mRNAs extend to positions -74 and -47 (Nickelsen *et al*, 1999). Moreover, in all cases the shorter form is much more abundant and corresponds to the mature and translatable mRNA. Processing of the *psbA* mRNA occurs at a stem-loop structure that binds a protein complex containing the 47 kDa PABP homologue, and is suggested to be essential for ribosome association (Bruick and Mayfield, 1998). A nuclear mutant of the *Mbb1* gene is affected in *psbB* mRNA stability and processing, and accumulates only the longer form, while wild type accumulate both the short and the long forms (Vaistij *et al*, 2000). Processing of the 5' UTR has been proposed to be essential for *psbD* stability and requires the nuclear *Nac2* function, because only the shorter transcript is found associated with polyribosomes (Nickelsen *et al*, 1999). Since PSII assembly involves the concerted expression of *psbA*, *psbB* and *psbD* genes, it is therefore possible that 5' end processing is coordinated to allow the proper translation of these mRNAs in the temporal order required for PSII assembly (reviewed by Erickson, 1998).

Contrary to the majority of prokaryotic and cytosolic mRNAs, chloroplast mRNAs contain long AU-rich 5' UTRs with almost no sequence conservation (reviewed by Hauser *et al*, 1998). Moreover, computer analysis based on minimum energy models,

predict these 5' UTRs to be highly folded at the secondary structure level, containing stable stem-loop structures (Fargo *et al*, 1999; Zuker, 1994). Besides, in contrast to their prokaryotic and cytosolic counterparts, secondary structures and regulatory factors in chloroplast mRNAs appear to involve positive *cis*-acting elements (reviewed by Zerges, 2000). This could be similar to the 'internal ribosome entry system' (IRES), used by the mRNAs of picorna viruses and by few mRNAs in the eukaryote nuclear-cytosolic compartments (Figure 2c). This system is independent of any 5'm^7G cap, Shine-Dalgarno or poly(A) structure, but involves a stem-loop structure and other regulatory factors in the 5' UTR to recruit the ribosome and promote translation (reviewed by Zerges, 2000).

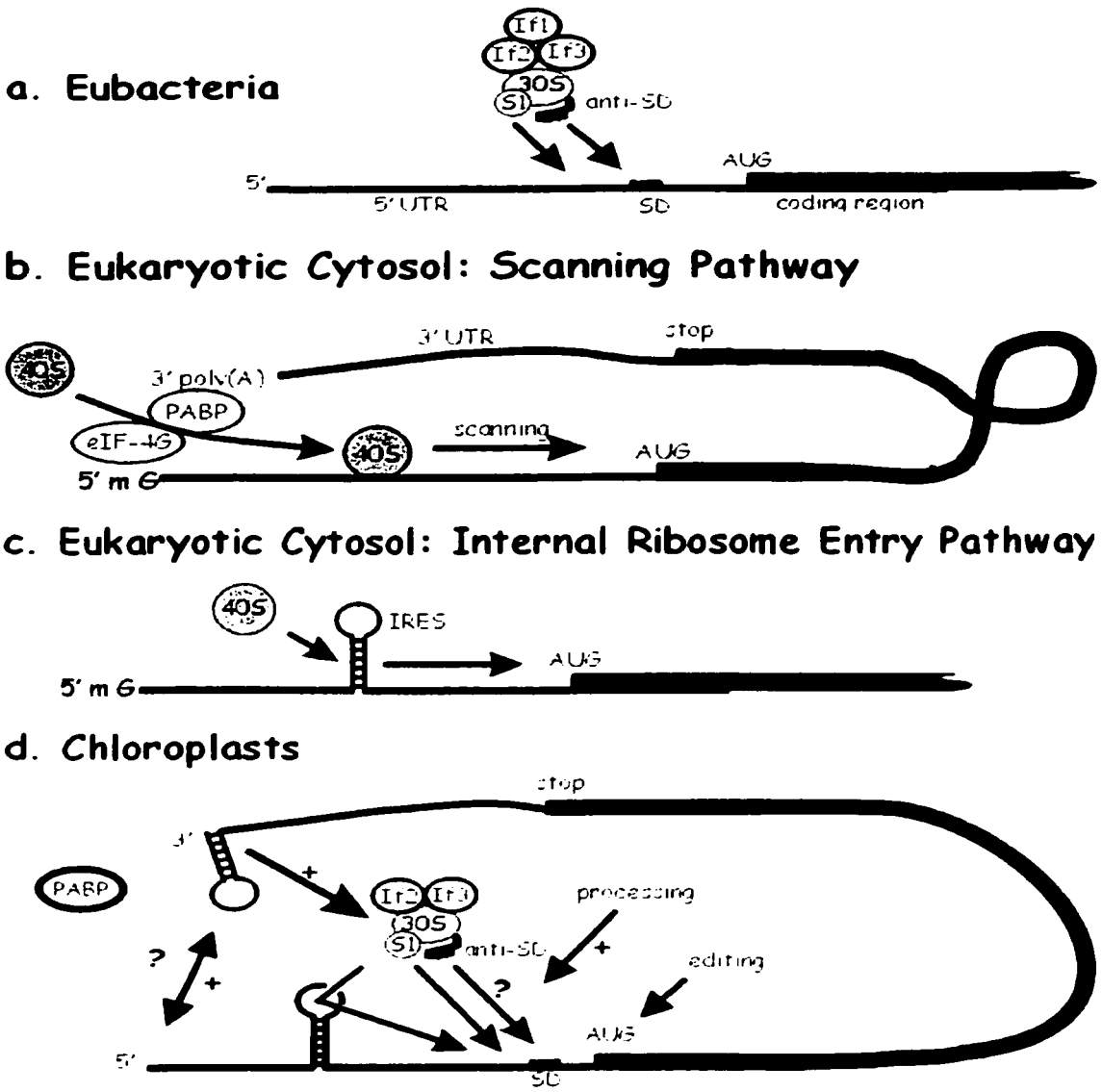

Figure 2. General mechanisms of translation initiation. **a.** In prokaryotes, initiation factors interact with the 30S ribosomal small subunit to recognize and bind to the RNA via the 16S rRNA and the 5' UTR Shine-Dalgarno sequence near the AUG. **b.** In the cytosol of eukaryotes, various eukaryotic initiation factors interact with features at the 5' and 3' ends to recruit the 40S ribosomal small subunit and to allow its scanning to the AUG. **c.** Internal ribosomal entry site represent a minor pathway used by some cytosolic mRNA and picorna viruses, in which the 40S is recruited via a structural motif of the 5' UTR. **d.** Similarities to the three pathways can be found in chloroplasts. Homologues of the prokaryotic machinery have been identified, such as the 30S, the 16S rRNA and its conserved anti-SD, and two initiation factors. Most chloroplast mRNAs do not seem to use a SD element. Structural elements in the 5'UTR of some of them are required. Even if they do not have any 5' m^7G cap or any poly(A) tail, interaction between the 3' and 5' UTRs have been found to be important in some cases, and a PABP homologue could interact with the 5' UTR of a mRNA (Zerges, 2000).

1.5 Importance of RNA structure

The structure adopted by RNA is often of primary importance. In the case of *Tetrahymena thermophila*, the self-splicing group I intron require the proper configuration of two of its subdomain structures (P4 and P6 domain) containing its catalytic activity (Szewczak *et al*, 1998). As seen above, translational regulation is often mediated by sequences in the 5' UTR. Chloroplast mRNA 5' UTRs show little sequence conservation, except for their AU-richness. However, many are predicted to be highly structured, suggesting that secondary and tertiary structures play a major role in translation regulation. Computer prediction, UV-melting profile, chemical and enzymatic probing, and mutagenesis have been used in the past to determine structural *cis*-elements affecting the translation of chloroplast mRNAs (Fargo *et al*, 1999; Fargo *et al*, 2000; Fargo *et al*, 1998; Higgs *et al*, 1999; Koo and Spremulli, 1994).

RNA thermodynamic parameters have been determined for a variety of structural features, including Watson-Crick base-pairing interactions (i.e. to form helices) and non-Watson-Crick base-pairing interactions between non-paired bases in loops and bulges, and other mismatches and stacking interactions (Mathews *et al*, 1999; reviewed by Tinoco and Bustamante, 1999). The RNA structure is theoretically hierarchical due to the fact that the primary sequence determines the secondary structure, and the secondary structure determines the tertiary structure (reviewed by Tinoco and Bustamante, 1999). The folding behaviour generally agrees with the search for the lowest free energy structure (reviewed by Tinoco and Bustamante, 1999). Even so, secondary and tertiary structures may transiently form sub-optimal structures before reaching the final energy state (reviewed by Tinoco and Bustamante, 1999). Trapped intermediate states could

possibly form. The unfolding behaviour becomes highly complex as the sizes of the RNA increases, and in the same way, resolving its structure becomes a complex problem. As the temperature of a solution containing a structured RNA is raised, tertiary structures unfold at lower temperatures and are generally sensitive to the presence of divalent ions, such as K^+ and Mg^{2+}, while secondary structures melt at higher temperatures and are stabilized by hydrogen bonds and stacking interactions (Draper and Gluick, 1995; Draper *et al*, 1995; Gluick and Draper, 1994; Gluick *et al*, 1997). Complex structures, like the P4-P6 domain of *T. thermophila* group I intron, will have their secondary structure almost unperturbed at temperature as high as 60°C, and that does not seem to be exceptional (Szewczak *et al*, 1998). Computer predictions that use energy minimization to construct secondary structure models of a given RNA sequence become useful to generate potential secondary structures. The *mfold* software is currently used for this purpose (http://128.151.176.70/RNAstructure.html) (Mathews *et al*, 1999; Zuker, 1994). Consistent results (64-86%) are usually found between computer prediction and experimental data (Mathews *et al*, 1999; Tinoco and Bustamante, 1999; Zuker, 1994).

On a native polyacrylamide gel, RNA mobility is determined by a combination of both folded and unfolded structure mobilities (Szewczak *et al*, 1998). Temperature gradient gel electrophoresis (TGGE) is a relatively simple technique that can be used to observe the structural folding pattern of any RNA, as revealed by its electrophoretic mobility in a horizontal temperature gradient within the gel. A gradient of temperature perpendicular to the direction of electrophoresis is applied over a native polyacrylamide gel with a low crosslinking ratio of bis-acrylamide to acrylamide (Birmes *et al*, 1990; Wartell *et al*, 1990). TGGE requires only trace amounts of radiolabeled material, and

allows direct visualization of the mobility and structure of RNA with respect to temperature and, unlike other techniques, the existence of multiple isoforms can be detected. TGGE is sensitive enough to detect small differences in free energy (Bevilacqua and Bevilacqua, 1998). However, the temperature range is limited by the tendency of the glass plates used to cast the gel to break at high temperature (above 60-65°C) and condensation and/or drying of the gels can also occurs at these temperature (Birmes *et al*, 1990; Szewczak *et al*, 1998). It also limits the type of solvent and buffer that can be used. The presence of metal ions (and Mg^{2+}) is known to cause RNA degradation at high temperature (Szewczak *et al*, 1998). In most cases, an increase in temperature is also known to induce a decrease in pH due to proton dissociation. A previous study showed that 0,09M Tris-borate-EDTA buffer pH 8.3, decreases by 0.68 units over a 50°C temperature range (from 20° to 70°C) (Birmes *et al*, 1990). The buffer stability is more critical when studying protein conformation, but a major change of the pH can also affect nucleic acids (Sambrook and Russell, 2000; Wartell *et al*, 1990).

The melting temperature (Tm) or hyperchromicity represents the temperature of a thermal transition between to folded and unfolded states and can be determined by observing the variation in RNA mobility on a TGGE. The transition between two states will have a different mobility and the Tm is situated in the middle of this transition. It is usually done by plotting a straight line through the lower baseline and a straight line through the upper base line of a each transition, and then determining the equidistant point from the two (Szewczak *et al*, 1998). Whenever possible, a UV-melting experiment (variation of the UV light absorbance at 260 nm at increasing temperature) is used to corroborate the Tm determined by TGGE, but they usually correlate (Bevilacqua

and Bevilacqua, 1998; Brion *et al*, 1999). Since the general electrophoretic mobility of the gel increases with temperature regardless of changes in structure, the RNA and the dye will run faster at higher temperature, creating a negative slope of mobility vs. temperature. Therefore, mobility changes resulting from structural transitions of RNA can be determined relative to a standard RNA with little or no structure. TGGE can be a useful tool for the analysis of RNA tertiary structure melting. However, it cannot alone reveal the exact nature of the structures that form in RNA. Even the Tm does not necessarily correspond to an individual transition, as well as the unfolded state does not necessarily indicate a perfectly single-stranded molecule (Laing *et al*, 1994). This must be defined by other methods, such as chemical and enzymatic structure probing, and ideally NMR and X-ray diffraction would be required, but they are far more expensive and require large amount of RNA, and until now, only tRNAs and rRNAs have yielded crystals able to diffract sufficiently well for high resolution structural determinations (Brodersen *et al*, 2002).

The principal objection to the use of specific ribonucleases in enzymatic probing is their large size, compared to the RNA, which can affect their accessibility to specific nucleotides, and can cause secondary and tertiary structure modification due to nucleotides nearest-neighbour effects and steric hindrance (reviewed by Knapp, 1989). Moreover, RNases can only be used for structure probing *in vitro*. Primer extension associated with dimethyl sulfate (DMS) modification represents a good way to probe chemically RNA structures. DMS specifically methylates adenine (at N-1 position), cytosine (N-3) and guanine (N-7) residues that are not involved in Watson-Crick interactions, such as base-pairing and protein binding (reviewed by Ehresmann *et al*,

1987). Only short incubation times are required and any temperature and various buffers composition can be used. In addition, it does not affect the transcript structure formation. DMS can be used *in vitro* and *in vivo*, because it can cross membranes. RNA can then be prepared and the methylation pattern determined by primer extension using a ^{32}P-labeled oligonucleotide hybridized specifically to the mRNA of interest. The enzyme reverse transcriptase will stop extending a cDNA copy from this oligonucleotide primer at the methylated A and C residues. Probing for methylated G residues require tedious aniline treatment and is rarely used (reviewed by Ehresmann *et al*, 1987). By comparing data from *in vivo* and *in vitro* methylated RNA, one can determine if some stops are related to proteins binding, since RNA treated *in vitro* is exempt of protein interaction. Comparisons of methylation patterns in wild type and mutant strains altered in the *trans*-acting translational regulatory factors can reveal the *cis*-elements and RNA structures they might form, which are important for this regulation (Higgs *et al*, 1999; Senecoff and Meagher, 1992). The only problem with this technique arises from the fact that naturally modified nucleotides block the elongation of the reverse transcriptase and that stable structures might prove to be difficult to melt by the enzyme (Ehresmann *et al*, 1987; Higgs *et al*, 1999; Senecoff and Meagher, 1992; Senecoff and Meagher, 1993; Wells *et al*, 2000).

In *E. gracilis*, the complete 5' UTR of the *rbcL* (encodes the large subunit of the ribulose-1,5-bisphosphate carboxylase) mRNA is required for maximal translation efficiency (Koo and Spremulli, 1994). It is only 55 nt in length and contains no SD-like sequence near the AUG start codon. The structure is believed to be the most important regulatory determinant, and would act as in a prokaryote, since mRNAs containing no

stable stem-loop are expressed more efficiently (Koo and Spremulli, 1994). However, it seems to be an exception, since most chloroplast mRNA translation is thought to be regulated via activator elements, such as a stem-loop structures that will be used as a binding sites by *trans*-acting factors to promote the translation, or be used as a recognition site to block ribonucleases activity therefore promoting RNA stability (Higgs *et al*, 1999; Mayfield *et al*, 1994; reviewed by Zerges, 2000). Genes containing mutations that weakened predicted stem-loop structures in the 5' UTR of *C. reinhardtii rps7* chloroplast mRNA, without affecting the mRNA level, was fused to the *aadA* reporter gene (Aminoglycosid-3'-Adenyltransferase confers spectinomycin / streptomycin resistance) and did not promote translation of the reporter, while compensatory mutations that restore these structures also restored translation of the chimeric reporter mRNA (Fargo *et al*, 1999). The modifications in the predicted secondary and tertiary structures that affected translation initiation were consistent with a change in the UV-melting profiles, while the mutations affecting the stability showed a similar profile to that of wild type (Fargo *et al*, 2000). A putative stem-loop structure that serve as a binding site and could be required for D1 synthesis has been reported using deletion mutants in the *C. reinhardtii psbA* 5' UTR. However, this structure is not part of the mature mRNA and its deletion did not affect considerably the translation (Mayfield *et al*, 1994). Using the *mfold* program (Mathews *et al*, 1999), Fargo *et al* (1999) observed that deleting the *psbA* SD-like element resulted in global changes in the predicted secondary structure, while replacements retained some of the wild-type folding pattern, which is in agreement with the translation efficiency. Computer folding and *in vivo* DMS probing allowed other researchers to characterise three elements in the 5' UTR that are required for translation

of the *petD* mRNA in *C. reinhardtii* (Higgs *et al*, 1999). Element I is required for stability and forms a small stem-loop structure near the 5' end. *In vitro* versus *in vivo* DMS methylation of element II revealed that it is unstructured and interacts with proteins that possibly bring it to form a secondary structure. Element III forms a stem-loop both *in vivo* and *in vitro*. Evidence suggests that element II and III interact together at the tertiary structure level. Mutations that are predicted to destabilize this alternate tertiary structure showed decreased translation (Higgs *et al*, 1999).

1.6 Translation of the psbC mRNA

As seen above, the chloroplast *psbC* gene of *C. reinhardtii* encodes P6 (the homologue of CP43 in vascular plants), which represent the 43 kDa chlorophyll-binding subunit of the PSII core reaction center. The *psbC* mRNA contains the longest 5' UTR in the *C. reinhardtii* chloroplast of 547 nt. Translation initiates at a GUG codon. A 98 nt stem-loop structure is predicted between positions 222-320 (with respect to the 5' end), and has been shown to be essential for translation (Rochaix *et al*, 1989; Zerges *et al*, 1997). A SD-like sequence is also present 14-10 nt from the GUG and its replacement reduces the level of translation (Zerges *et al*, submitted). A systematic deletion mutant series has been undertaken to further characterize regions of importance for translation of this mRNA. Various deletion mutants of a chimeric 5' UTR fused to the *aadA* reporter gene were assessed for their ability to confer spectinomycin resistance and for chimeric mRNA accumulation (Zerges *et al*, submitted). The first 93 nt region at the 5' end is required for transcription and/or mRNA stability. Two regions were identified as essential for translation: the stem-loop region (222 – 320) and an AU-rich region 30 nt

(519 – 530) just upstream of the SD-like sequence (534 – 538), near the GUG. A region spanning nucleotides 378 to 518 is dispensable, while two other regions flanking the stem-loop (94 – 222 and 320 – 378) are partially required for translation (Zerges *et al*, submitted).

Three nuclear loci are required for translation initiation of the *psbC* mRNA (Figure 3). They are named *TBC1*, *TBC2* and *TBC3* to indicate that they are required for translation (T) of the mRNA encoding the PSII (B) subunit P6 (C) (Zerges *et al*, 1997). Pulse labelling of newly synthesized chloroplast proteins, in the presence of cycloheximide to inhibit translation in the cytosol, showed that *tbc1-F34* and *tbc2-F64* nuclear mutants completely abolish the translation, while *tbc3-rb1* mutant only slightly reduces P6 synthesis. However, *tbc3-rb1* also partially suppresses the mutant phenotype of *tbc1-F34*, but not that of *tbc2-F64*. A chloroplast suppressor mutant, *psbC-F34sul*, also partially restores the defect caused by *tbc1-F34*. It consists of the transversion of a T to an A at position 226 in the stem of the predicted 222-320 stem-loop structure. This mutation enlarges a predicted bulge at the base of the stem. Another chloroplast mutation, *psbC-FuD34*, is defective in translation. It corresponds to two point mutations in the predicted 222-320 stem-loop but on the opposite strand of the stem to the strand affected by the *psbC-F34sul* mutation. The removal of one C at position 415 eliminates the same bulge affected by *psbC-F34sul*, while the insertion of two T's at position 309-310 eliminates a second predicted bulge in the stem. The overall stability of the 222-320 stem-loop is predicted to be higher in *psbC-Fud34* (ΔG –38.6 kcal at 25°C, and a predicted Tm of 75.7°C) compared to the wild type (ΔG –31.1 kcal; Tm = 64.0°C), while the *psbC-F34sul* is predicted to slightly destabilize the 222-320 stem-loop (ΔG –26.8

kcal; Tm 62.7°C). These genetic data suggest that melting of this stem-loop structure might be required for translation. The bulges might also be essential by representing recognition site for *trans*-acting factors. Furthermore, the *TBC*-dependent factor might interact with the stem-loop since *psbC-F34sul* partially suppresses the requirement for *TBC1*. Moreover, *tbc3-rb1* partially restores the translation from the UTR bearing the *psbC-FuD34* mutation, a deletion of the stem loop. *TBC3* might also interact with *TBC1* and the 5' UTR to control the translation at the mRNA level (Rochaix *et al*, 1989; Zerges *et al*, 1997; Zerges and Rochaix, 1994).

UV cross-linking experiments revealed the binding of a 46 kDa protein with the 5' UTR, but only in *tbc2-F64* strain (Zerges and Rochaix, 1994). *TBC2* might have to inactivate the binding activity of this 46 kDa protein. *TBC2* has recently been cloned (Auchincloss *et al*, 2002). It encodes a protein of 114.8 kDa localized in the chloroplast stroma, and is part of a high molecular weight complex of 400 kDa that does not appear to contain RNA since its size is not altered following RNase treatment. Moreover, it is not associated with polysomes, meaning that it interacts only transiently or indirectly with ribosomes. *TBC2* shares some similarities (37%) with the maize chloroplast protein *CRP1*, which is required for the processing and translation of *petA* and *petD* mRNAs (Auchincloss *et al*, 2002).

In this study, I used different approaches to uncover the importance of *cis*-elements, and more specifically the structural motifs within the 5' UTR that affect the translation of the *psbC* mRNA. Preliminary studies using primer extension analysis did not allow me to identify a precursor form of the mRNA with an extension at the 5' end, as it is the case

for *psbA*, *psbB* and *psbD*. However, it does not imply that *psbC* is not processed. Several strong and consistent stops during the footprint analysis point out the importance of some structural differences in the stem-loop region between the wild type and mutants, especially the strains containing the *psbC-FuD34* mutation in the stem-loop region. Mobility experiments undertaken also revealed differences in the mobility and the melting pattern of some mutants compared with the wild type, particularly in *psbC-FuD34*. As its predecessors, this study suggests an important role for the 222-320 stem-loop in the translation of the *psbC* mRNA.

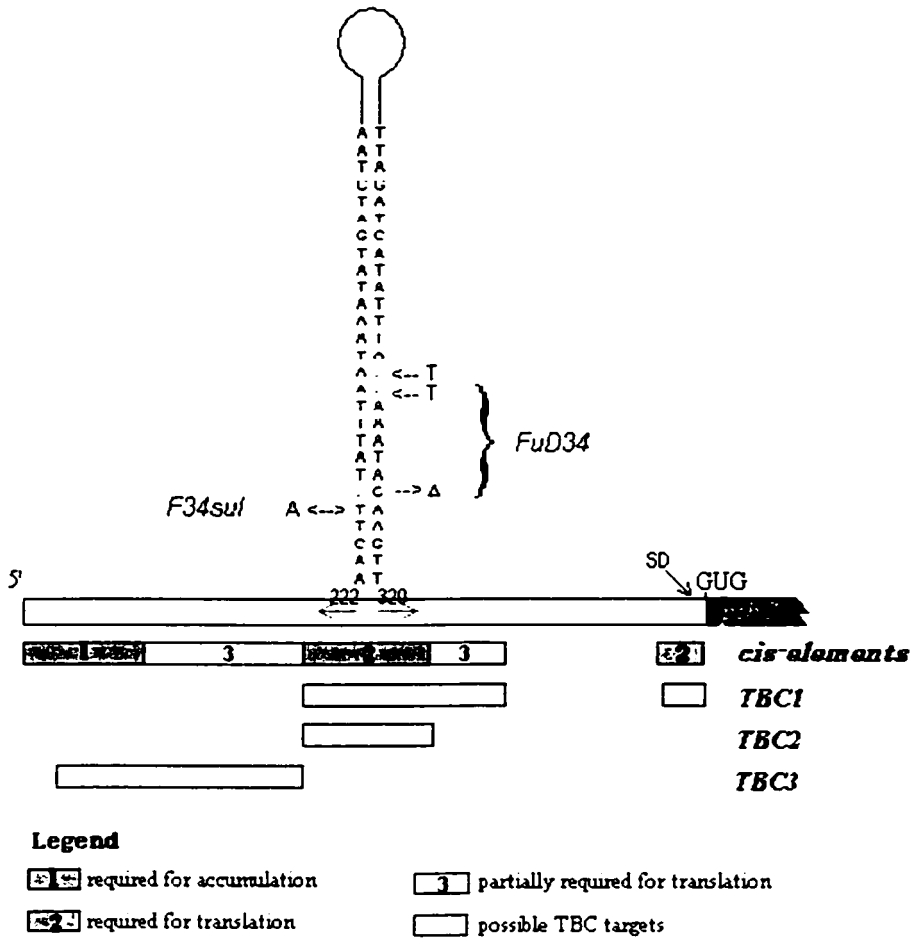

Figure 3. Schematic representation of the genetic of *psbC* mRNA translation. The *psbC* 5' UTR is 547 nt long and has a GUG start codon. A predicted stem-loop structure from position 222-320 (with respect to the 5' end) is required for the translation. Region near the GUG is also required and contains the SD-like sequence. The first 92 nt at the 5' end are required for transcription or RNA stability. Other regions are partially required or dispensable. Predicted target regions of the three nucleus-encoded *TBC* factors are indicated. *psbC-FuD34* and *psbC-F34sul* chloroplast mutations and their respective alteration on the stem-loop are indicated. The dots in the stem represent the absence of complementary base-pairing, which create bulges in the stem (Rochaix *et al*, 1989; Zerges *et al*, submitted).

2. MATERIALS AND METHODS

2.1 Strains and cultures conditions

All the *C. reinhardtii* strains used in this work are described in Table 1. Cells were cultured in Tris-acetate-phosphate medium (TAP) (Harris, 1989), under a light intensity of 100 $\mu E\ m^{-2}s^{-1}$. Strains were maintained in light to reduce selection for phenotypic revertants. To confirm the PSII deficient phenotypes of the mutants, these strains were tested for growth, under bright light, on high-salt minimal medium (HSM), which lacks a reduced carbon source.

The series of *psbC*-5' UTR deletion mutants used in mobility experiments are represented in figure 5. They consist of chimeric 5' UTRs fused to the *aadA* reporter gene (Zerges *et al*, submitted). However, only their respective 5' UTR region was PCR amplified (see below) and used in this study. *E. coli* strain DH5α was used during the molecular cloning step described hereafter.

2.2 RNA extraction and DMS footprinting

Extreme care was taken during manipulation of RNA to prevent contamination and degradation. Total RNA was extracted from 2 - 5 X 10^6 cells/ml, using TRI Reagent (Sigma). Typically, for a 25 ml culture, cells were harvested at 5000 rpm for 5 min, resuspended in 1 ml of 1X TAP and centrifuged again at 7000 rpm for 2 min in a 1.5 ml microtube. The supernatant was then removed and the pellet was stored at -80°C for at least 30 min. 500 µl of TRI Reagent was added to the frozen pellet, which was then resuspended by vortexing. Glass beads were added (the equivalent of 50 µl) and the

sample was mixed again by vortexing for 15 sec. The sample was incubated for 5 min at room temperature (around 21°C) and then centrifuged at 12,000 rpm 15 min at 4°C, followed by another incubation period of 5 min at 21°C. 100 µl of chloroform (without isoamyl alcohol) was added and the sample was vortexed for 15 sec, followed by another 5 min incubation period at 21°C, and then centrifuged again at 12,000 rpm for 15 min. The aqueous phase was transferred to a new microtube and the RNA was precipitated using 250 µl of cold iso-propanol and centrifuged at 12,000 rpm for 15 min after a 5 min incubation period. The pellet was washed with 75% (vol/vol) ethanol then centrifuged again (12,000 rpm 15 min at 4°C) and resuspended in 20 µl of dH_2O (DEPC-treated HPLC grade water) and stored at -20°C or -80°C.

In vivo DMS treatments (7.9 M dimethyl sulfate, Aldrich) were modified from those described previously for *C. reinhardtii* (Higgs *et al*, 1999). 25 µl of DMS was added to a 25 ml culture (in TAP), gently mixed and incubated for 5 min at 21°C, then 250 µl of β-mercaptoethanol was added to quench the reaction. Total RNA from *in vivo* DMS-treated culture was then extracted as described above. For *in vitro* DMS treatments, 40 µg of total RNA (prepared from untreated cells) was allowed to adopt its stable structure in 200 µl of DMS buffer (10 mM Tris-HCl pH 7.5; 10 mM $MgCl_2$; 3 mM $CaCl_2$) by heating the sample at 75°C 5 min and then cooling during 30 min at room temperature. 1 µl of DMS was diluted in 9 µl of ethanol and the 10 µl were added and gently mixed, then incubated 5 min before adding 20 µl of β-mercaptoethanol. The sample was ethanol precipitated with 2.5 volumes of 95% (vol/vol) ethanol, 0.1 volume of 3M sodium acetate (pH 5.6) and incubated 30-45 min at -20°C (or 15 min at -80°C). It was then centrifuged

(12,000 rpm 15 min at 4°C) and washed with 75% (vol/vol) ethanol and then centrifuged again at 12,000 rpm 15 min at 4°C, and then resuspended in dH_2O.

For primer extension 10 µg of RNA were ethanol precipitated with 1 pmol of [^{32}P]-5' labelled primer. Oligonucleotides used as primers (Figure 4) were labelled using T4 polynucleotide kinase (Fermentas) and [γ-^{32}P] ATP (Amersham). The primers were 5' end labelled as recommended by the manufacture (Fermentas) and purified by G-50 spin column as described previously (Sambrook and Russell, 2000). The RNA-primer mix was resuspended in 22 µl of dH_2O with 8 µl of hybridization buffer (3M NaCl; 0.5 M HEPES; 2mM EDTA). The hybridization mix was denatured and 85°C 5 min in a dry bath, then allowed cooling slowly in the dry bath to room temperature (approx. 90 min). Then it was precipitated with 2.5 volumes of 95% (vol/vol) ethanol, 0.1 volume of 3M sodium acetate (pH 5.6) and incubated 30-45 min at -20°C (or 15 min at -80°C) prior centrifugation (12,000 rpm 15 min at 4°C) and washed with 75% (vol/vol) ethanol and centrifuged at 12,000 rpm 15 min at 4°C. The pellet was resuspended in 25µl of extension buffer (2.8 µl 5mM dNTPs; 5 µl 5X reverse transcriptase buffer (Promega or Invitrogen); 12 units of Ribonuclease inhibitor; 1.25 µl 1 mg/ml BSA; 2 µl DTT; 2 µl 200 mM $MgCl_2$; dH_2O). Addition of 50 to 100 units of M-MuLV reverse transcriptase (Promega) or 100 units of Superscript II reverse transcriptase (Invitrogen) started the reaction. Either the reaction was started immediately at 45°C for 90 min, or a temporal temperature gradient was applied from 25°C to 45°C in 20-30 min, and incubated at 45° during 60 min. The reaction was allowed to extend further and terminated by a gradual temperature increase from 45°C to 85°C during 20 min. The sample was precipitated by addition of 2.5 volumes of 95% (vol/vol) ethanol and 100 µl 3M ammonium acetate and

incubated at 45 min at -20°C prior centrifugation at 12,000 rpm 15 min at 4°C. The pellet was washed with 75% (vol/vol) ethanol prior centrifugation (12,000 rpm 15 min at 4°C). The pellet was resuspended in 10 µl of RNA loading dye (95% (vol/vol) formamide; 0.025% (wt/vol) SDS; 0.025% (wt/vol) bromophenol blue; 0.025% (wt/vol) xylene cyanol FF; 1mM EDTA), and 4 µl of this sample were denatured at 95°C for 5 min then chilled on ice prior the loading. The [^{32}P]cDNA were electrophoresed in 8% denaturing polyacrylamide gels with 8M urea (PAGE) (56 cm X 38 cm X 4 mm) by application of 1800-2000 Volts at a maximum of 50 W and 50 mA until the xylene cyanol was approximately 10 cm from the bottom of the gel (approximately 90 min). The gel was then fixed in 10% (vol/vol) methanol and 10% (vol/vol) acetic acid for 5 to 10 min and dried on a Wattman 3M paper for 60 min at 60°C under vacuum. Signals were visualized by autoradiography using Kodak Biomax MS or Fuji RX x-ray film with a β-screen. Exposures were from 24h to 72h at -80°C. Sequencing ladders obtained by chain-termination reactions (Sanger *et al.* 1977) derived from using each of the [^{32}P]primers of the wild-type 5' UTR (NTL template, see below) were electrophoresed in parallel for each primer, to be able to identify positions where the reverse transcriptase stops on the nucleotide sequence of the 5' UTR. Sequencing was done using a T7 DNA polymerase sequencing kit (Amersham).

Table 1. Strains used in this study

Strain[†]	Phenotype	Nuclear genotype	Chloroplast genotype
JVD.4A^{+}	Wild type (WT)	*TBC1;TBC2;TBC3*	psbC-WT
F34.2A^{-}	PSII deficient	*tbc1-F34;TBC2;TBC3*	*psbC-WT*
F64.2	PSII deficient	*TBC1;tbc2-F64;TBC3*	*psbC-WT*
RB1.2A	Partial PSII activity	*TBC1;TBC2;tbc3-rb1*	*psbC-WT*
FuD34.3	PSII deficient	*TBC1;TBC2;TBC3*	*psbC-FuD34*
F34suI	Partial PSII activity	*TBC1;TBC2;TBC3*	*psbC-F34suI*
RB1.4A	Partial PSII activity	*TBC1;TBC2;tbc3-rb1*	*psbC-FuD34*

[†] All those strains have been described in previous studies (Rochaix *et al*, 1989; Zerges *et al*, 1997).

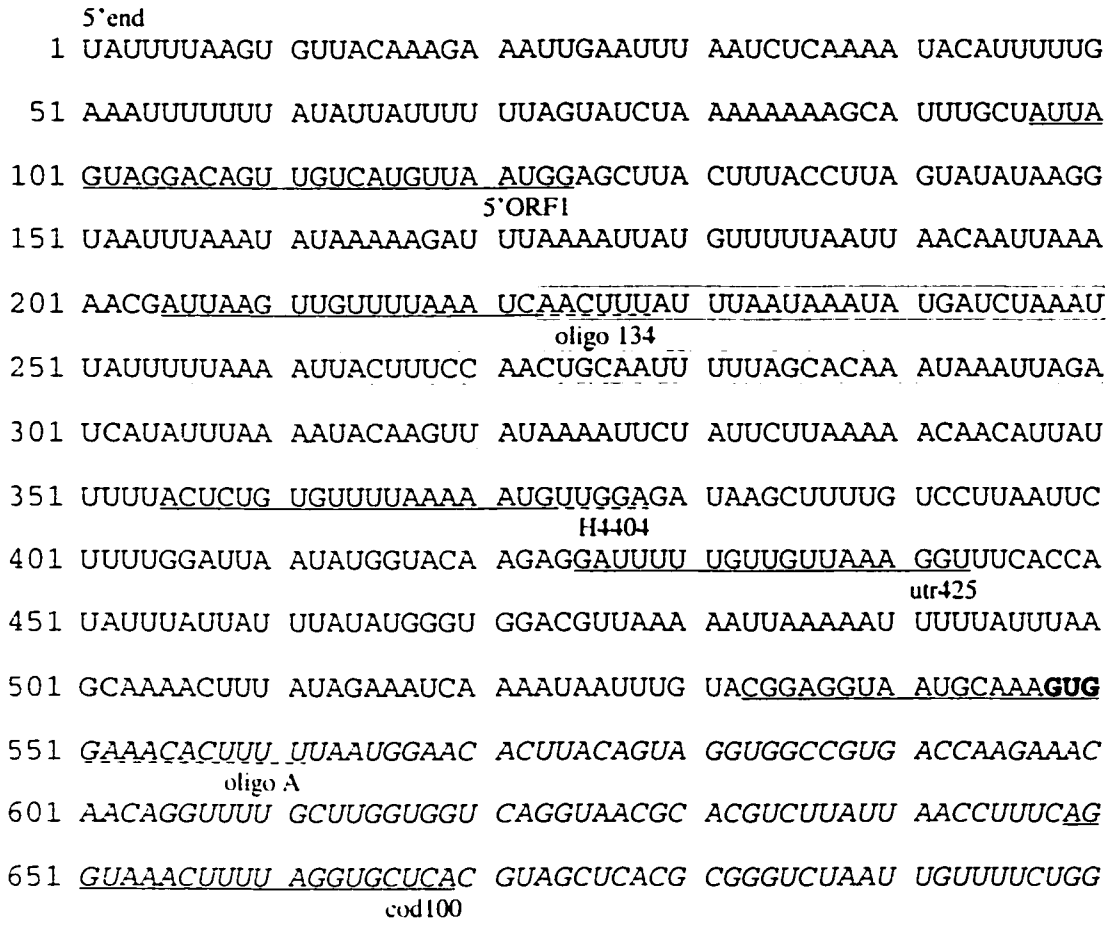

Figure 4. Oligonucleotides map. Sequence of the *psbC* 5' UTR mRNA and part of its coding sequence are represented, with the localization of the oligonucleotides (underlined) used for the primer extension (dotted lines represent a part of the oligo sequence that does not correspond to the mRNA sequence, but rather contains a restriction site). The coding sequence is shown in *italic*, the GUG start codon is shown in **bold** character and the stem-loop region (222-320) is boxed. Note that oligonucleotides are complementary to this mRNA sequence.

2.3 PCR, RNA synthesis, labelling and RNA purification

Plasmid clones (pKS and pBS) containing inserts of the *psbC* 5' UTR of wild type, *psbC-FuD34*, *psbC-F34sul* and a set of deletion mutants (Figure 5), were amplified by PCR (*Taq* DNA polymerase from Fermentas). The forward primer (5'-GGATCCGTAATACGACTCACTATAGGG*ATTTTAAGTGTTACA*) contained the T7 promoter (underlined) and annealed at the first 15 nucleotides at the 5' end of the 5' UTR (in italics). The reverse primer was the oligo A (Figure 4). A 1.2 kb cloned fragment with the 3' end of *psbC* (parts of the coding and non-coding sequences and 3' flanking sequences) was used to generate a 225 nt fragment containing the *psbC* 3' UTR. The forward primer (5'-GTAATACGACTCACTATAGGGTCATGGTTAGCTTGTTCG) also contained the T7 promoter (underlined), while the reverse primer was 5'-GGATCCTAGTAAATAACAAATCTG. PCR products were then used as templates to synthesize RNA with T7 RNA polymerase following the manufacturer's protocol (Promega). The reaction was terminated by addition of 5 units of DNase I per μg of template DNA and incubating at 37°C for 15 min, and then extracted with phenol : chloroform : isoamyl alcohol (25:24:1) and twice with chloroform alone. The aqueous phase was then precipitated by addition of 2.5 volumes 95% (vol/vol) ethanol and 0.1 volume of 3M sodium acetate (pH 5.6) and incubated -20°C for 30 min prior centrifugation at 12,000 rpm 15 min at 4°C. The pellet was washed twice with 75% (vol/vol) ethanol and centrifuged at 12,000 rpm 15 min at 4°C. The pellet was resuspended in dH_2O. For each 1 μg of template DNA, 8 μg of RNA transcript were usually obtained. RNA was either 5' end labelled using T4 polynucleotide kinase (Fermentas) or labelled during synthesis as described above except that 2 pmol of [α-

^{32}P]dCTP was added (ICN) in order to synthesise a low specific activity RNA for the TGGE. RNA was gel purified from an 8% denaturing polyacrylamide gels with 8M urea, as described previously (Sambrook and Russell, 2000). The gel fragment containing the RNA of interest was identified by autoradiography and was excised with a clean razor blade. The RNA was eluted overnight at 37°C in 0.3 M sodium acetate (pH 7.0); 1 mM EDTA; 0.1% SDS, and then precipitated by addition of 2.5 volumes of 95% (vol/vol) ethanol and 0.1 volume of 3M sodium acetate (pH 5.6) and incubated -20°C for 30 min prior centrifugation at 12,000 rpm 15 min at 4°C. The pellet was washed with 75% (vol/vol) ethanol and centrifuged again at 12,000 rpm 15 min at 4°C. The pellet was resuspended in dH_2O. Purification of non-labelled RNA was done using 1% agarose gel electrophoresis. The slice from the gel containing the RNA was visualized by staining with ethidium bromide and illumination with UV light. The RNA was recovered by filtration through a polyester filter floss disposed at the bottom of a 0.5 ml microtube with a hole at the bottom. The solution containing the RNA was collected by centrifugation in a 1.5 ml microtube and then frozen, phenol extracted (twice), chloroform extracted (twice) and the ethanol precipitated. The pellet was resuspended in dH_2O. RNA could be stained with ethidium bromide and approximately quantitated, but anomalously low absorbance of 260 nm light was detected. Controls were done using DNA instead of RNA during the purification, and UV absorbance level was in agreement with the level of recovery that I expected (85%). Ethidium bromide staining "drop assays" were used to confirm that I had the same recovery level with the RNA.

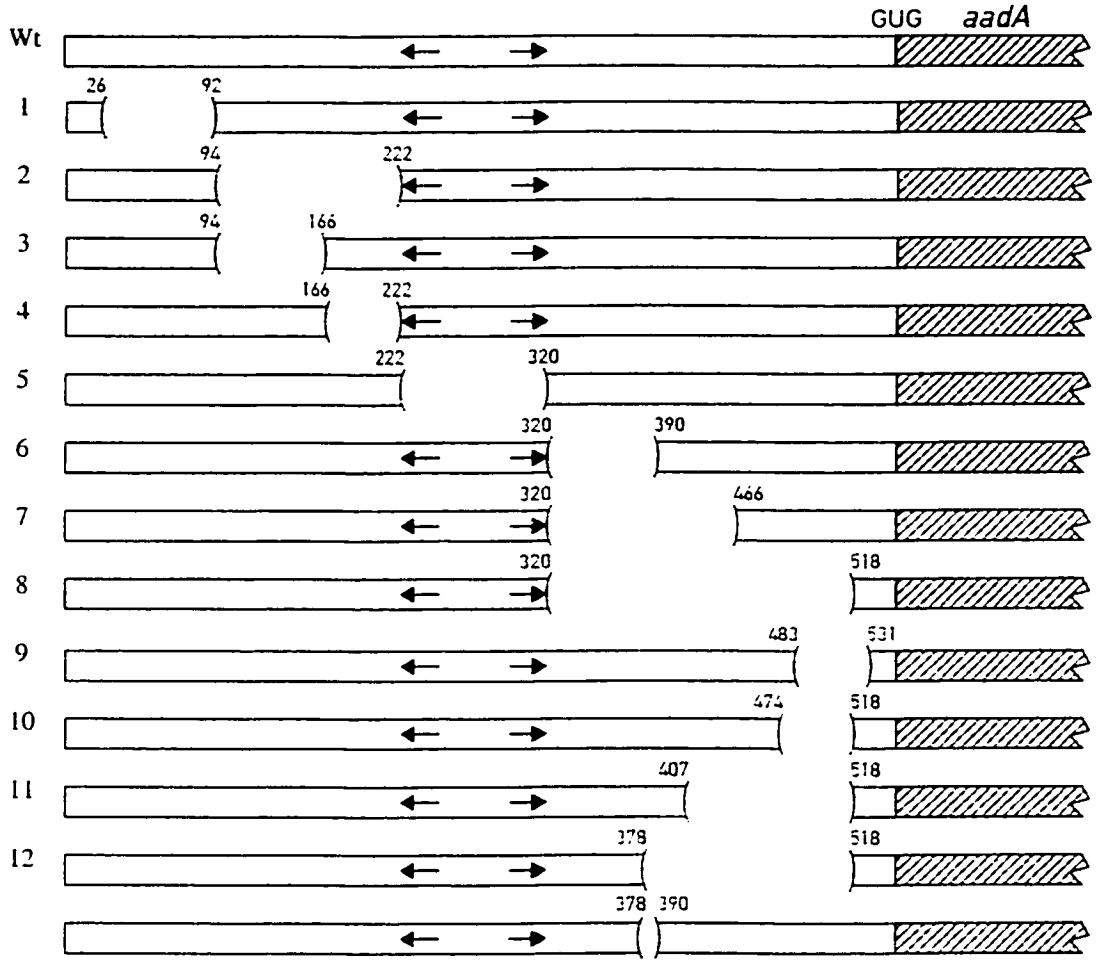

Figure 5. Map of the *psbC* 5' UTR deletion mutants. These mutants have been previously fused to the aadA reporter gene to identify important *cis*-elements of the translation (Zerges *et al*, submitted). PCR amplification has been used to select only the 5'UTR. In the present study, they will be referred using the deletion region, by example: *psbC-Δ222-320*.

2.4 Temperature gradient gel electrophoresis (TGGE)

The TGGE system (Figure 6) was based on a model described previously (Wartell *et al*, 1990). A 60:1 acrylamide/N,N'methylene-bis-acrylamide ratio is used to cast a 6% native gel (1.5 mm thick and 10 cm long by 13 cm wide). The gel is placed between two aluminium blocks. Channels at the left-hand side of these aluminium blocks are connected to a cooling circulating water bath, while a similar system on the other side is heated by water circulating from a bath of higher temperature. Pumps provide constant flow in the channels. The temperature gradient is therefore established perpendicular to the direction of the electrophoresis and is highly reproducible and linear. Thermocouple probes connected to a digital display were used to set the temperature range from 25°C to 65°C and to assess the linearity of the perpendicular gradient (3°C/cm). 1X TBE was used as the running buffer in both chambers. It was previously shown to decrease by only 0.68 pH over a 50°C increase (from 20° to 70°C) in this type of experiment (Birmes *et al*, 1990). Approximately 10 ng of low specific activity ^{32}P-RNA was denatured at 95°C, allowed to fold to its proper conformation during 1h, and then loaded across the gel. The gel was then pre-run without any gradient for 25 min to allow the RNA to be properly inserted in the gel between the two aluminium blocks. Then the cooling and heating water were allowed to establish the temperature gradient (the gradient forms in approximately 15 min) before the RNA was allowed to adopt conformations in relation to the temperature at each horizontal position across the gel for 1 hour. The gel was exposed to an electric field of 8.5 V/cm (maximum setting were 85 V; 2 W; 20 mA) for approximately 4 h. The RNAs were then fixed by soaking in 10% (vol/vol) methanol and 10% (vol/vol) acetic acid for 5 to 10 min and then adhered to a Wattman 3M paper and

dried 120 min at 60°C on a gel dryer under vacuum. Autoradiography was taken by exposing the gel on a Kodak Biomax MS x-ray film with a β-screen for 72h at -80°C.

Temperature melting points were measured as previously described by plotting the slope (baseline of each structural states) between two structural transition and identifying the equidistant point (Szewczak *et al*, 1998).

Changes in mobility were also measured using a similar setting, but instead of forming a gradient of temperature across the native gel, one temperature was maintained constant across the gel, and standard wells were used to restrict each sample to a single lane. Native gels were then run at 25°, 35°, 45° 55° and 65°C with all the deletion mutants on the same gel in separate lanes. RNA labelling and autoradiography were done as previously described.

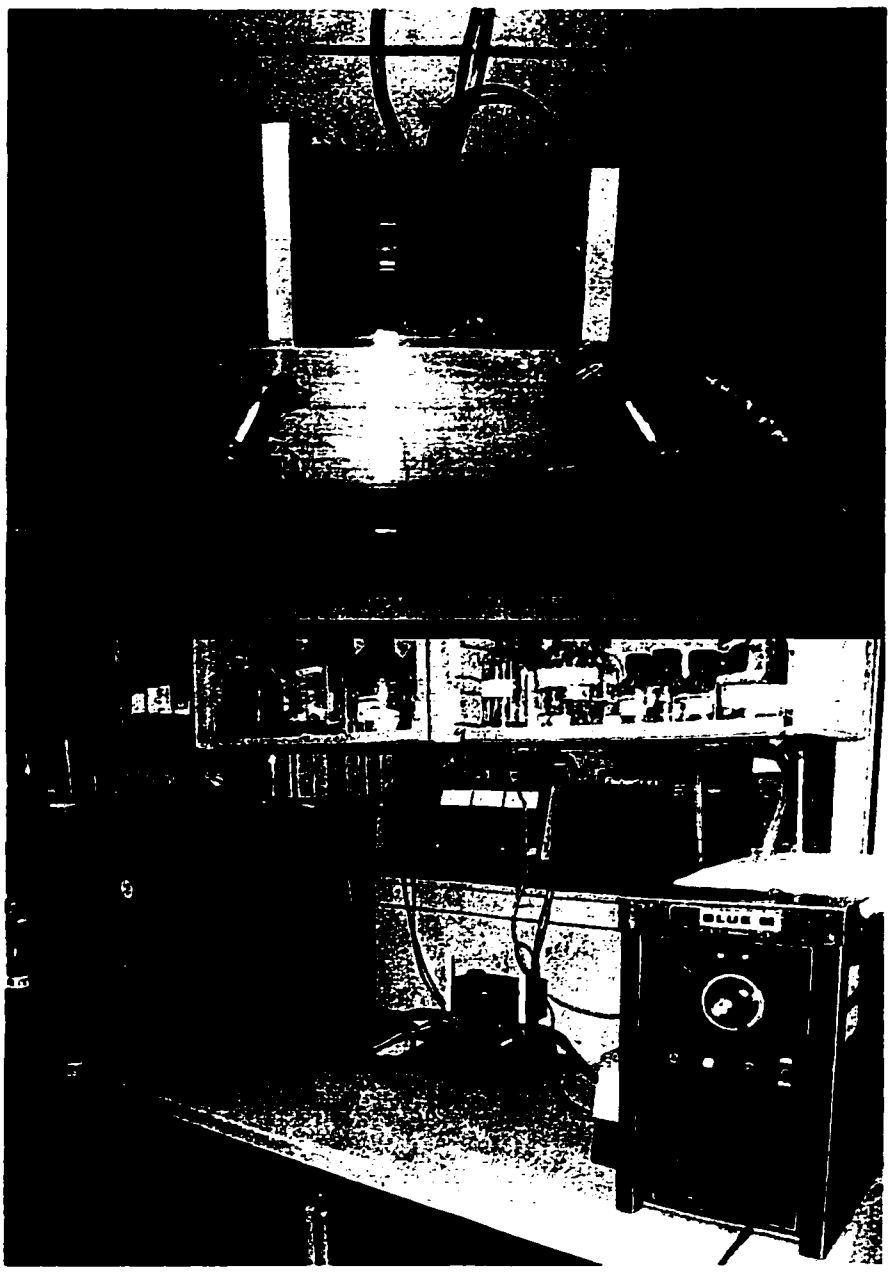

Figure 6. Pictures of the Temperature gradient gel electrophoresis (TGGE) apparatus and setup. The front view (top) shows the gel inserted between two aluminium blocks and clamped to the gel box. Pumping and tubing system (bottom) allow water from the cooling bath (left) and heating bath (right) to circulate at each extremity of the aluminium block therefore creating a horizontal gradient of temperature. Water enters a the top corner of the back aluminium block, circulate in a channel along the edge inside the block and then enter the front block (using the U tube) by the bottom corner and exit at the top corner of the block and empty in its respective bath.

3. RESULTS

3.1 Primer extension analysis addressed whether or not the *psbC* mRNA has a minor larger form with a 5' terminal extension

Primer extension is a useful tool to identify mRNA 5' ends. A [^{32}P] oligonucleotide primer is hybridized to a specific site of the mRNA, within 50-500 nt of the 5' terminus. Reverse transcriptase is then used to extend a cDNA from this oligonucleotide primer to the 5' end of the mRNA. Usually, most of the reverse transcriptase will stop at the 5' end of a given RNA. However, premature stops can be caused by structures in the mRNA. The length of the cDNA, determined by denaturing PAGE and autoradiography, corresponds to the distance between the 5' end of the primer and the 5' end of the mRNA. In the case of *psbA*, *psbB* and *psbD*, primer extension results identified minor forms of the mRNA with 5' ends derived from sequences located further upstream. This longer mRNA form is thought to represent the pre-mature form of the mRNA and its processing is required for translation to occur (Bruick and Mayfield, 1998; Nickelsen *et al*, 1999; Vaistij *et al*, 2000). When analyzed by primer extension, some longer *psbC* 5' ends can be seen depending of which primer one uses, however these 5' ends are inconsistent from one primer to another (data not shown). So far, only oligo 5'ORF1 and oligo 134 have shown an identical upstream 5' end (Figures 7). They hybridize to the *psbC* 5' UTR at 124 and 228 nt, respectively, from the 5' end. This second 5' end is situated 9 nt from the more abundant mature 5' end and seems to be present only in the *RB1.4A* mutant, and possibly in *psbC-FuD34*. The reason why we do not see it in the other strains and even in wild type might be that more primer successfully extended to

the 5' end in the *RB1.4A* compared with the other strains (fewer downstream stops and a more abundant 5' end stop in *RB1.4A*), although the 5' ends generated in all the strains with the oligo 5'ORF1 are darker than the *RB1.4A* 5' end generated by oligo 134. This raises the possibility that the *TBC3* factors might interact with the 222-320 stem-loop (directly or indirectly) and affect a processing event, although this event does not affect the accumulation of mature *psbC*.

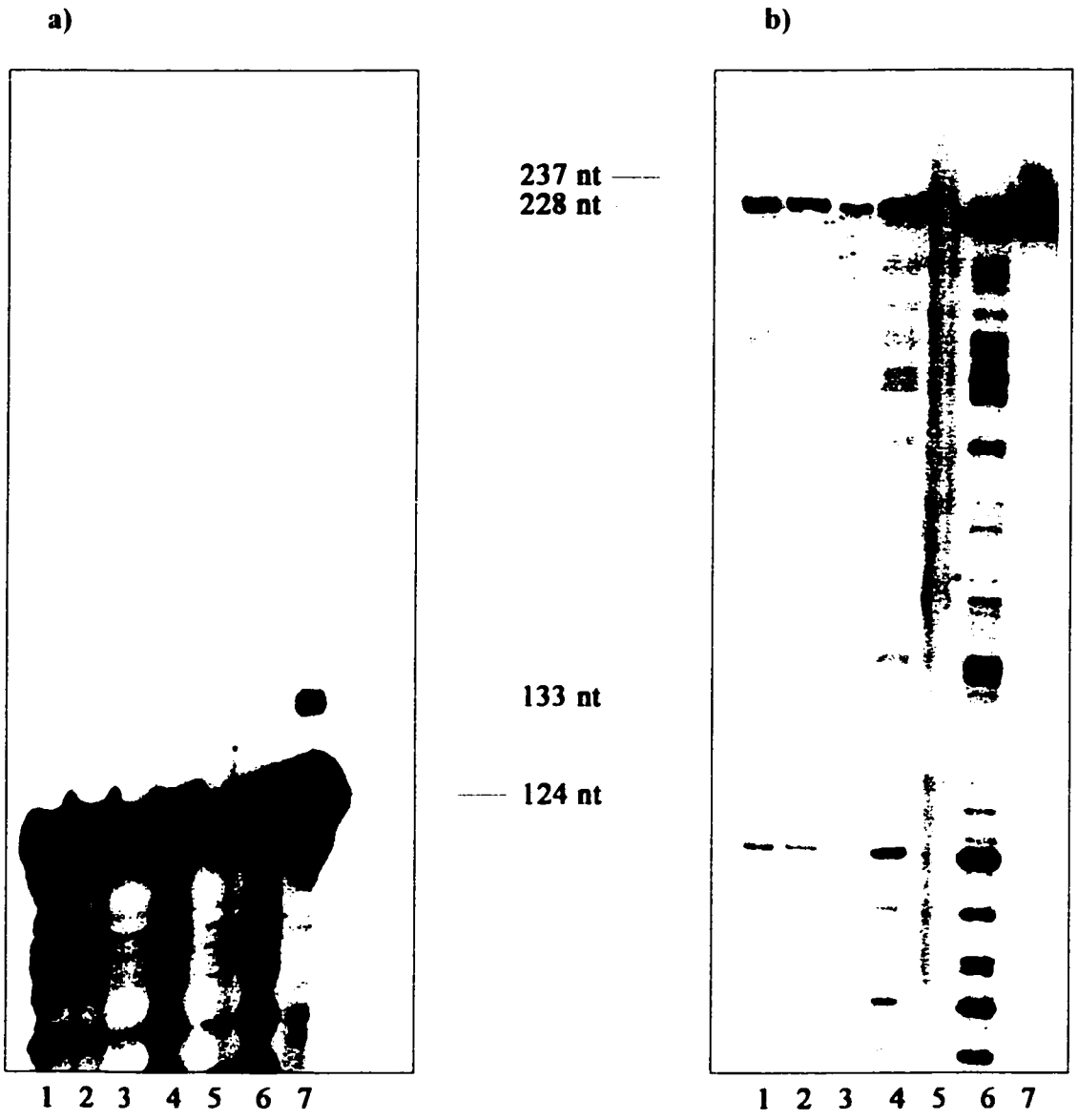

Figure 7. 5' end mapping of *psbC* wild type and mutants using primer extension. Total RNA from wild type (lane 1) and *tbc1-F34* (lane 2), *tbc2-F64* (lane 3), *tbc3-rb1* (lane 4), *psbC-F34sul* (lane 5), *psbC-FuD34* (lane 6) and *tbc3-rb1:psbC-FuD34* (lane 7) mutants, where analyzed for the presence of multiple 5' ends using oligo 5'ORF1 **(a)** and oligo 134 **(b)** as primers. A second 5' end is clearly present 9 nt from the mature 5' end in the *tbc3-rb1:psbC-FuD34* mutant (lane 7).

3.2 Reverse transcriptase stops prematurely at several A's and Us involved in stable structures

Stops can also occur before the 5' end and might be caused by modified bases and/or stable structures (reviewed by Ehresmann *et al*, 1987). Those early stops are a common problem with analysis of chloroplast mRNAs by primer extension (Higgs *et al*, 1999). Analysis of the 22 major stops observed in *psbC* 5' UTR with oligo utr425 (Figure 8) revealed that 12 correspond to A's (54.5%) and 10 are Us (45.5%). This could result from the fact that *psbC* 5' UTR is AU rich (38.6% A's and 41.3% Us). However, most of the stops observed when extending with utr425, can be seen when using oligo A (Figure 10). One surprising observation is that four reverse transcriptase stops (U151, A341, U364 and U378) are present only in *psbC-FuD34* while three others (U144, A229 and A370) seen on the wild type mRNA are absent in *psbC-FuD34* (Figure 8, lane 5). This is the case at position U144 where all the strains stop, except for *psbC-FuD34* where a stop is detected at position U151. The 133-153 region is predicted to form a small stable stem-loop (Figure 9a). All the strains stop at the single-stranded U144 in the loop, while on the mRNA with the *psbC- FuD34* mRNA, an reverse transcriptase stop is observed at a double-stranded U151 in the middle of the stem (Figure 9a). Several reverse transcriptase stops occur in the 222-320 stem-loop (Figure 9b). The stop at position A229 is present only on the *psbC* mRNA from wild type or strains carrying the *tbc3-rb1* mutation in the stem of the 222-320 stem-loop, near the positions altered by the *psbC-F34sul* and *psbC-FuD34* mutations. Another stop at position A260 occurs right below the loop. Stop A287 is in a small predicted bulge just before the loop. Stop A309 is at the *psbC-FuD34* mutation site. However, it is present in every strain including *psbC-*

FuD34. Stops between positions 334 to 384 are part of a more or less stable region that contains several duplexes (base-paired region) interspaced with internal loops and bulges. In this region, most of the stops are predicted to be single-stranded A or Us. Finally the region 385-424 forms another stable stem-loop (Figure 9c) that contain two bulges where U391 and U395 stop occur.

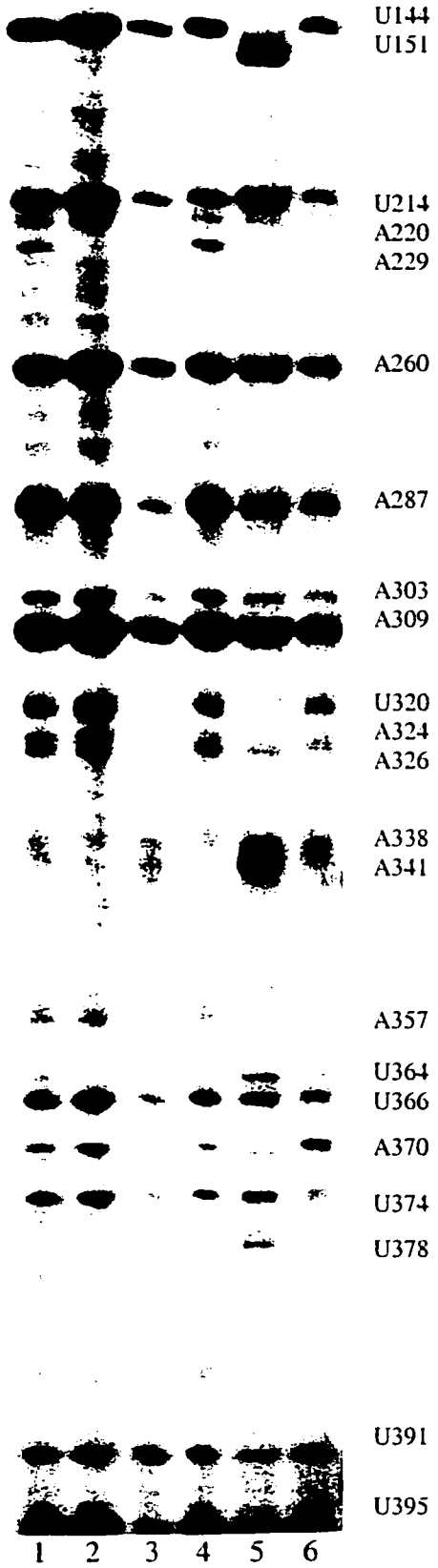

Figure 8. Primer extension of *psbC* wild type and mutant strains. Total RNA from wild type (lane 1), and *tbc1-F34* (lane 2), *tbc2-F64* (lane 3), *tbc3-rb1* (lane 4), *psbC-FuD34* (lane 5), *psbC-F34sul* (lane 6) mutants, was analyzed by primer extension with oligo 425. Reverse transcriptase pauses at several nucleotides (indicated at the right of the picture) before reaching the 5' end (not shown). Most of them are involved in predicted stable structures (see figure 9), and some specific to *psbC-FuD34* mutant.

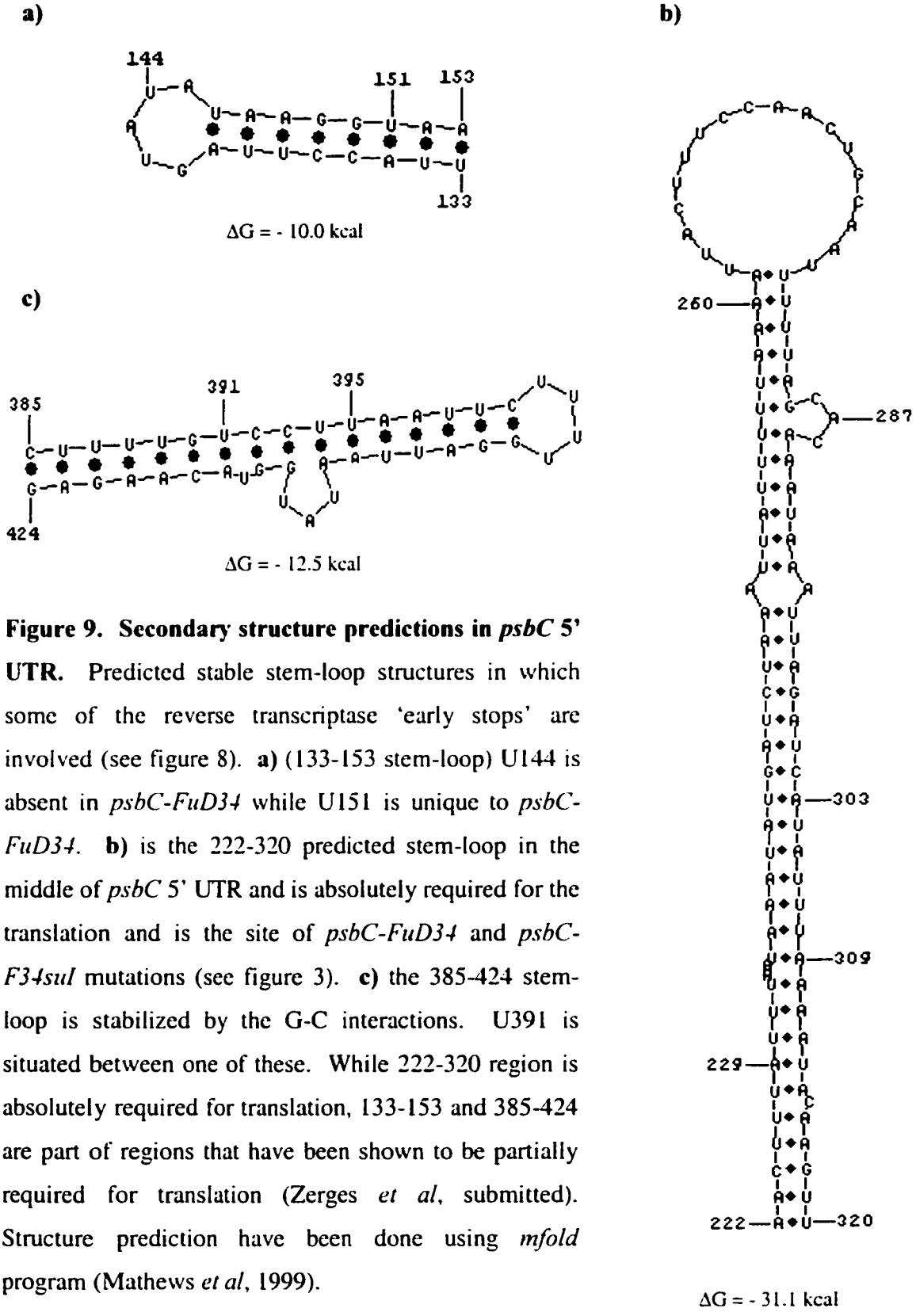

Figure 9. Secondary structure predictions in *psbC* 5' UTR. Predicted stable stem-loop structures in which some of the reverse transcriptase 'early stops' are involved (see figure 8). **a)** (133-153 stem-loop) U144 is absent in *psbC-FuD34* while U151 is unique to *psbC-FuD34*. **b)** is the 222-320 predicted stem-loop in the middle of *psbC* 5' UTR and is absolutely required for the translation and is the site of *psbC-FuD34* and *psbC-F34sul* mutations (see figure 3). **c)** the 385-424 stem-loop is stabilized by the G-C interactions. U391 is situated between one of these. While 222-320 region is absolutely required for translation, 133-153 and 385-424 are part of regions that have been shown to be partially required for translation (Zerges *et al*, submitted). Structure prediction have been done using *mfold* program (Mathews *et al*, 1999).

3.3 DMS footprinting

Dimethyl sulfate (DMS) footprinting was performed using RNA isolated from a wild-type strain. Since primer extension on *psbC* 5' UTR shows a high number of stops that are caused by structures, and not by DMS methylation, this type of experiment requires rigorous optimization. The same pattern of reverse transcriptase stops was observed on an RNA corresponding to the *psbC* 5' UTR transcribed *in vitro* and, therefore, these stops are not due to modification of bases (data not shown). Only my analyses of RNA from wild-type strains yielded some preliminary results. Figure 10 shows the primer extension analysis done on untreated RNA (lane 1), and DMS treated RNA *in vivo* (lane 2 and 3) and *in vitro* (lanes 4 and 5) using oligo A (see Figure 4). The first observation is the presence of numerous structure-induced reverse transcriptase stops (described above). Most correspond to the stops previously described for oligo 425. The stronger stop in lane 1 corresponds to region A287 (the bulge in the 222-320 stem-loop, see Figure 9c) and a stop at A315 was present only in *psbC-Fud34* strain this time in this experiment. Moreover, while the untreated sample was extended to its 5' end, the DMS treated samples revealed a stop at A287, but the intensity decreases soon after. This is due to the increasing amount of stops during the extension, caused in part by the presence of methylated bases near the beginning of the extension (bottom of the x-ray film). Based on computer prediction it is unlikely that this region is composed of only single-stranded A's and C's. Moreover, since this region does not contain only A and C residues, several stops also represent G's and T's.

Few stops differ between treated and untreated lanes: nine positions were methylated *in vivo*; two were methylated only *in vitro*, while at least one is clearly missing *in vitro*

over the 190 nucleotides mapped (Figure 10). C315 is the base that is removed in *psbC-FuD34*. It is also predicted to form an internal bulge in the wild type and *psbC-F34sul* 5' UTRs. Following DMS methylation pattern analysis, a predicted secondary structure was selected based on the fact that all nine *in vivo* methylated nucleotides were found to be single-stranded (Figure 11). Two bases methylated *in vitro*, A378 and C385, are predicted to be involved in Watson-Crick base-pairing. However, the predicted probability of A378 of being single-stranded is 35% and the adjacent GU base pair (G379-U497) has 45% chance of being single-stranded. These probabilities were computed using *mfold* 3.1 and correspond to the **%p-num** (measure of the level of promiscuity in the association of any given nucleotide with or helix alternative complementary pairs) and the **ss-count** (measure of the propensity of bases to be single-stranded) (Mathews *et al*, 1999). For C385 the probability is only of 4%. DMS footprinting might also help to reveal tertiary interaction such as pseudoknots, where single-stranded regions of two adjacent structures hybridize (reviewed by Pleij and Bosch, 1989). Overall, the predicted secondary structures examined agreed within 56% to 100% (average of 71%) with the methylation pattern observed.

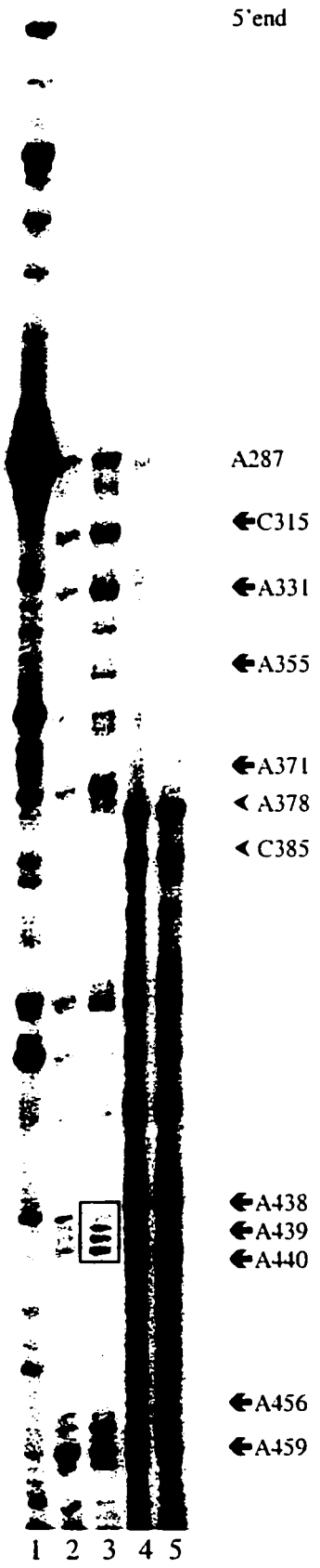

Figure 10. DMS footprinting analysis of wild type *psbC* 5' UTR. Primer extension of total extracted RNA was done using oligo A. Untreated RNA (lane 1), are compared with *in vivo* DMS treated RNA (lane 2 and 3) and *in vitro* DMS treated RNA (lane 4 and 5) (see material & methods). Methylated nucleotides are identified at the right. The untreated RNA presents multiple 'early stops' (they correspond to those seen in figure 8 with oligo 425). Two stops (A378 and C385) have been specifically methylated *in vitro*. Treated RNA rarely extends efficiently over one or two hundred nucleotides (in this case almost 300 nt) probably due to the abundance of downstream stops. Legend: ← (arrow) *in* vivo methylation stops; ◂ (arrowhead) *in vitro* methylation stops; ▯ (boxed region) delimit A438-A440 stops)

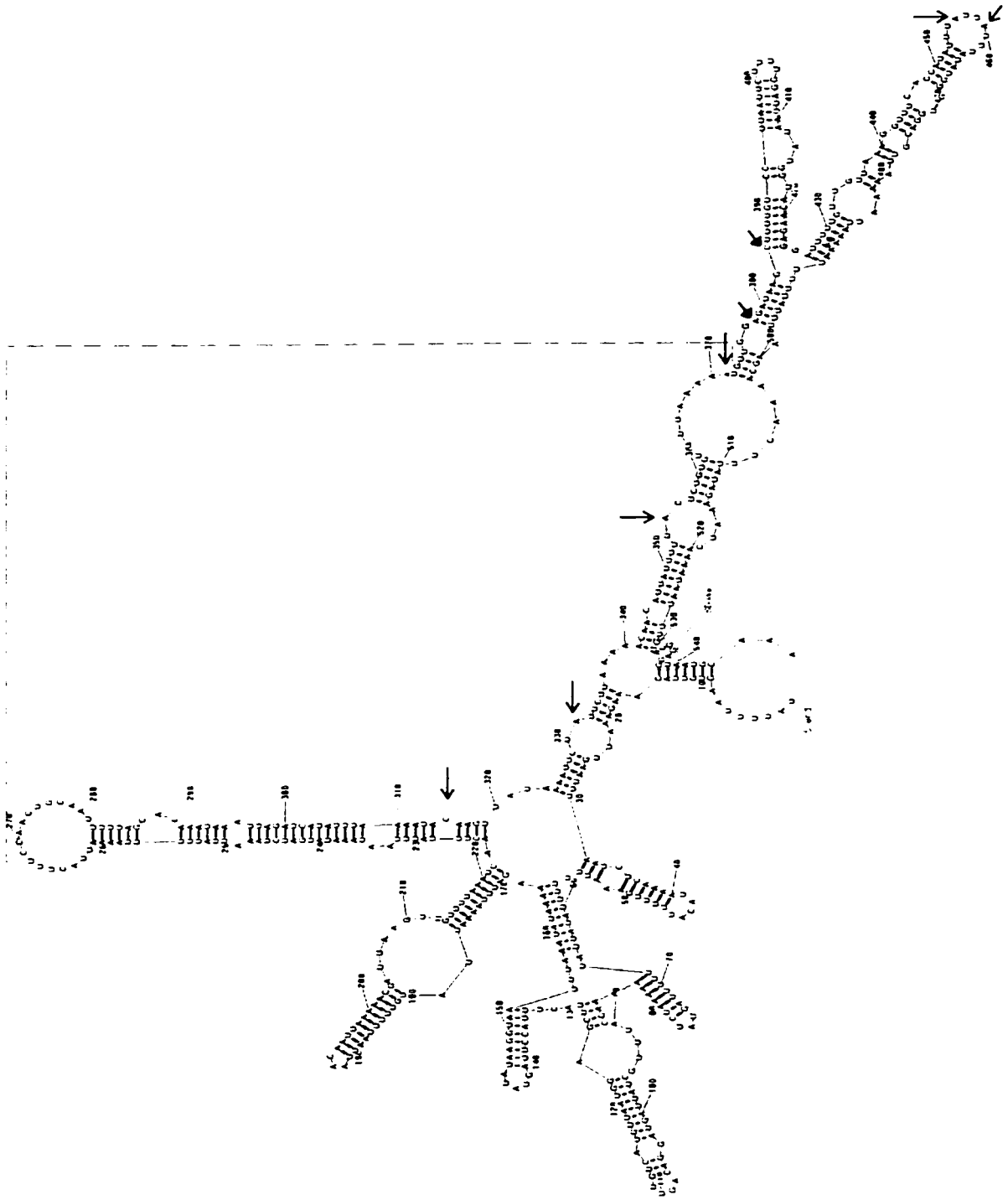

Figure 11. Predicted *psbC* 5' UTR secondary structure. Nucleotides methylated *in vivo* are identified by open arrows (↓), while the two methylated *in vitro* nucleotides are marked by closed arrows (➧). The single-stranded Shine-Dalgarno like sequence (GGAG) is boxed, and doted lines identify a possible pseudoknot.

3.4 Change in RNA gel mobility and temperature gradient gel electrophoresis

When run on a native gel, an *in vitro* synthesized RNA corresponding to the *psbC* 5' UTR with the *psbC-FuD34* mutation was first observed to migrate slower than the wild type and *psbC-F34sul* 5' UTRs (data not shown). The RNA mobilities of wild type, *psbC-FuD34*, *psbC-F34sul* 5' UTRs and a set of *psbC* 5' UTR deletion mutants (Figure 5) were electrophoresed on a series of five native gels with a low cross-linking ratio (60:1), in order to amplify the effect of their structure on electrophoretic mobility. Each gel was electrophoresed at a temperature from 25° up to 65°C (Figure 12). The electrophoretic mobilities of these RNAs were compared to those of an RNA molecular weight standard. The distance between the 400 MW RNA standard and each 5' UTR was plotted with respect to temperature, and compared to the wild type pattern (Figure 13). The first observation is that all 5' UTRs show a slower mobility at 25°C compared to their expected size, as seen on a denaturing polyacrylamide gel or on a native acrylamide gel at 65°C. Interestingly, the RNA with *psbC-FuD34* shows a dramatically aberrant mobility at temperatures 35°C, with lesser effects seen at 25°C. Moreover, the 5' UTR with the deletion *psbC-Δ320-466* shows two bands also at 35°C, possibly reflecting the existence of two distinct conformations. In addition, a deletion of the 222-320 stem-loop region causes more variation relative to wild type, as observed with deletion mutants *psbC-Δ166-222*, *psbC-Δ320-466* and *psbC-Δ320-466*. All these deletion mutants affect *psbC* translation (Zerges *et al*, submitted) and affect regions that are predicted to have secondary structures by the *mfold* program (Mathews *et al*, 1999). This is similar to *psbC-FuD34* that is predicted to stabilize the stem-loop, while the *psbC-F34sul* UTR has electrophoresed with the wild-type UTR at all temperatures tested.

The 5' UTRs of wild-type sequence or baring the mutations, *psbC-FuD34*, *psbC-Δ320-390* or *psbC-Δ320-466* were further analyzed using the temperature gradient gel electrophoresis system described in Material and Methods (see Figure 6). TGGE results are shown on Figure 14. As discussed (see Introduction), the electrophoretic mobility of all molecules within a gel is directly proportional to temperature, regardless of the RNA structure. Therefore, one must correct for this negative slope in determinations of the effects of RNA structure on mobility. The wild-type 5' UTR shows an almost linear gradual negative slope, which parallels the molecular weight marker RNAs (Figure 14 b). In contrast, the three mutant 5' UTRs displayed more complex relationships of mobility with temperature. In the case of the 5' UTR with *psbC-FuD34* a deviation from the constant slope seen for the molecular weight marker is observed at temperatures flanking 40°C, while deviations for the two deletion mutants are observed at 32°C for *psbC-psbC-Δ320-466* and at 34°C for *psbC-psbC-Δ320-466*. These deviations could indicate that the RNA undergoes a melting of tertiary and secondary structures, and that transition states are detected with partially unfolded structures at intermediate temperatures. Two conformations are visible for *psbC-Δ320-466* as previously observed with the native gel at a 35°C constant temperature. The samples were compared with an RNA corresponding to the 3' UTR of the *psbC* mRNA, and the distance was reported over temperature on Figure 15. One point to keep in mind is that while the slope of the mobility *versus* temperature is linear and negative, the distance separating the 5' UTRs from the 3' UTR increased. This is because the electrophoretic mobility of RNA is related to the log of its molecular weight; a longer molecule will run slower than a shorter one. This also seems to be amplified by the increasing mobility that affect

electrophoresis at higher temperatures. This was tested by running the labelled RNA ladder on a TGGE. The variation from start to end (i.e. the distance between two samples of different molecular weight at 25°C and the distance between the same samples at 65°C) was seen to increase following a logarithmic pattern from the higher to the lower molecular weight standards (not shown). On Figure 15, we can see that the 5' UTR with *psbC-FuD34* is only slightly slower at 25°C than wild type, but ends at a similar position at 65°C, as expected since they differ by only one nucleotide in length. TGGE are sensitive enough to detect a single nucleotide change, which might explain the slower mobility of *psbC-FuD34* at 25°C (Bevilacqua and Bevilacqua, 1998). However, *psbC-Δ320-466* has a smaller size, but start at a position similar to wild type and ends at its size-expected position. Melting temperature of an RNA structure can be determined using TGGE by plotting the slopes before and after a structural transition on a TGGE gel image. The center of this structural transition slope represents the melting temperature (Szewczak *et al*, 1998). As shown on Figure 16, the 5' UTR with the *psbC-FuD34* mutation shows a higher Tm (43.9°C) compared to wild type (33.8°C). *psbC-Δ390-466* (Tm 38.1°) also shows a noteworthy difference compared to wild type. It should be noted that these TGGE experiments were performed in the absence of Mg^{2+} ions, which are known to stabilize both secondary and tertiary structures (Gluick *et al*, 1997; Laing *et al*, 1994). It is therefore important to note that structures responsible for the aberrant migration of the 5' UTR are independent of Mg^{2+} ions.

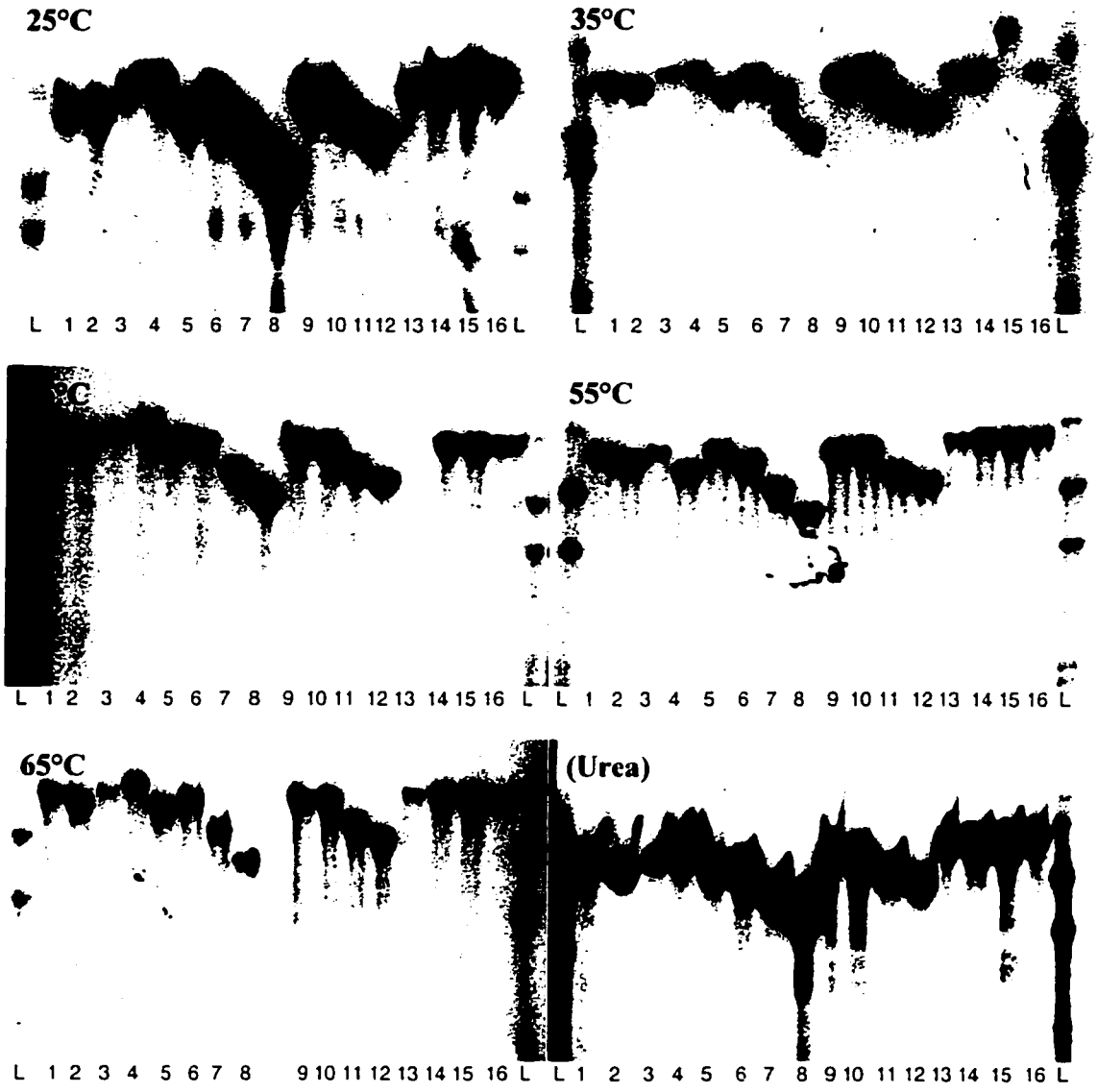

Figure 12. RNA mobilities on a series of native gels at constant temperature. Wild type (14), *psbC-FuD34* (15), *psbC-F34sul* (16), and a series of deletion mutant (see numbering on the left side of Figure 5 in Materials and Methods) where run on native gels, with the RNA molecular weight ladder (lane L). The temperature of each gel is indicated in the top right corner. The denaturing gel contained 8M urea and was run at 50°C.

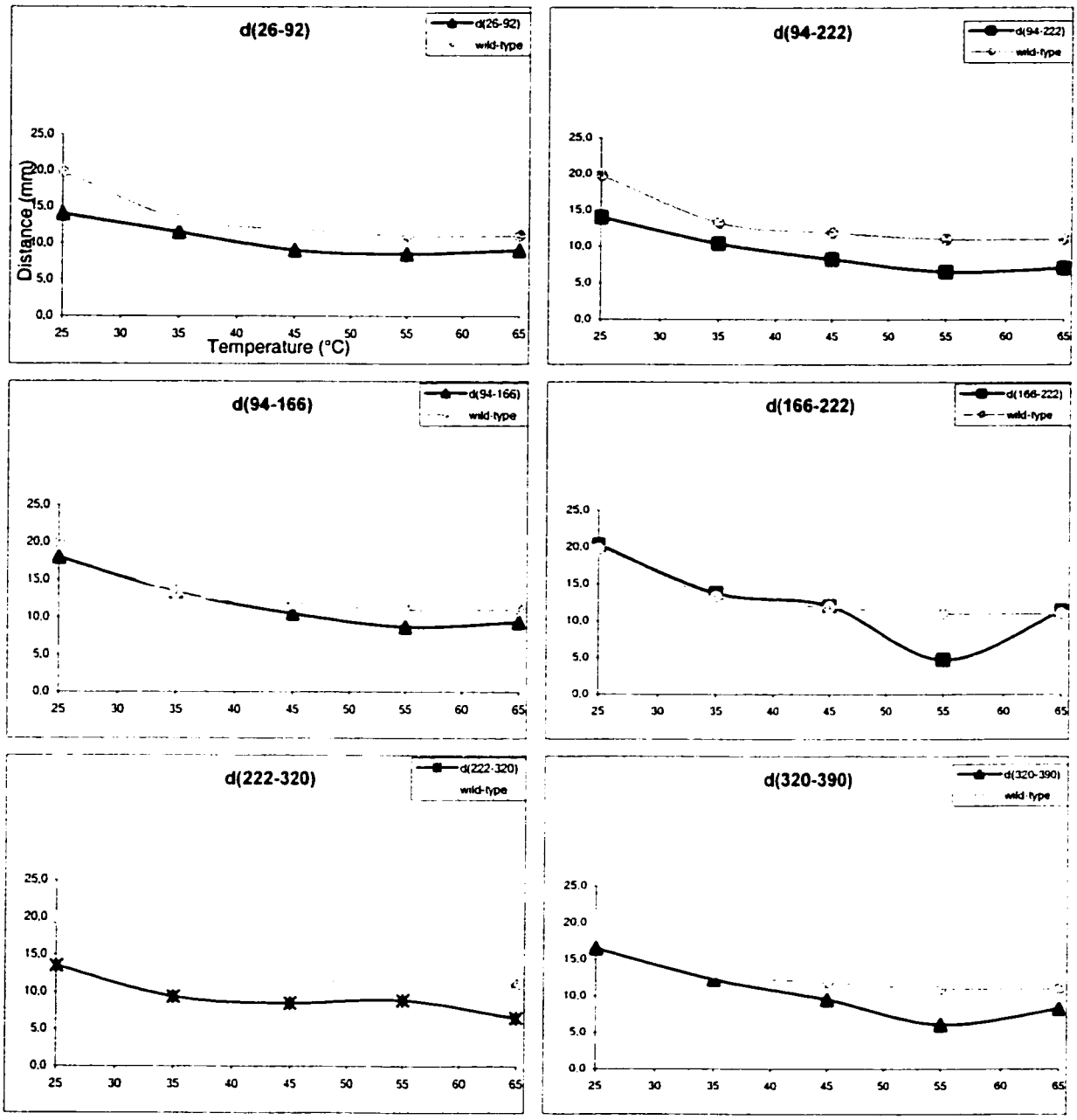

Figure 13. Mobilities of *psbC* 5' UTR wild type and mutants. Since electrophoretic mobility of RNAs increase with temperature, the distance between various *psbC* 5' UTR samples (see figure 5 in materials & methods) and a 400 nt molecular weight RNA standard (Fermentas) was measured using five native gels run at constant temperature (25°, 35°, 45°, 55° and 65°C). Distance between the 5' UTR and the 400 nt standard are in mm on the **Y** axis, and temperature in degrees Celsius on the **X** axis. (Following on next page...)

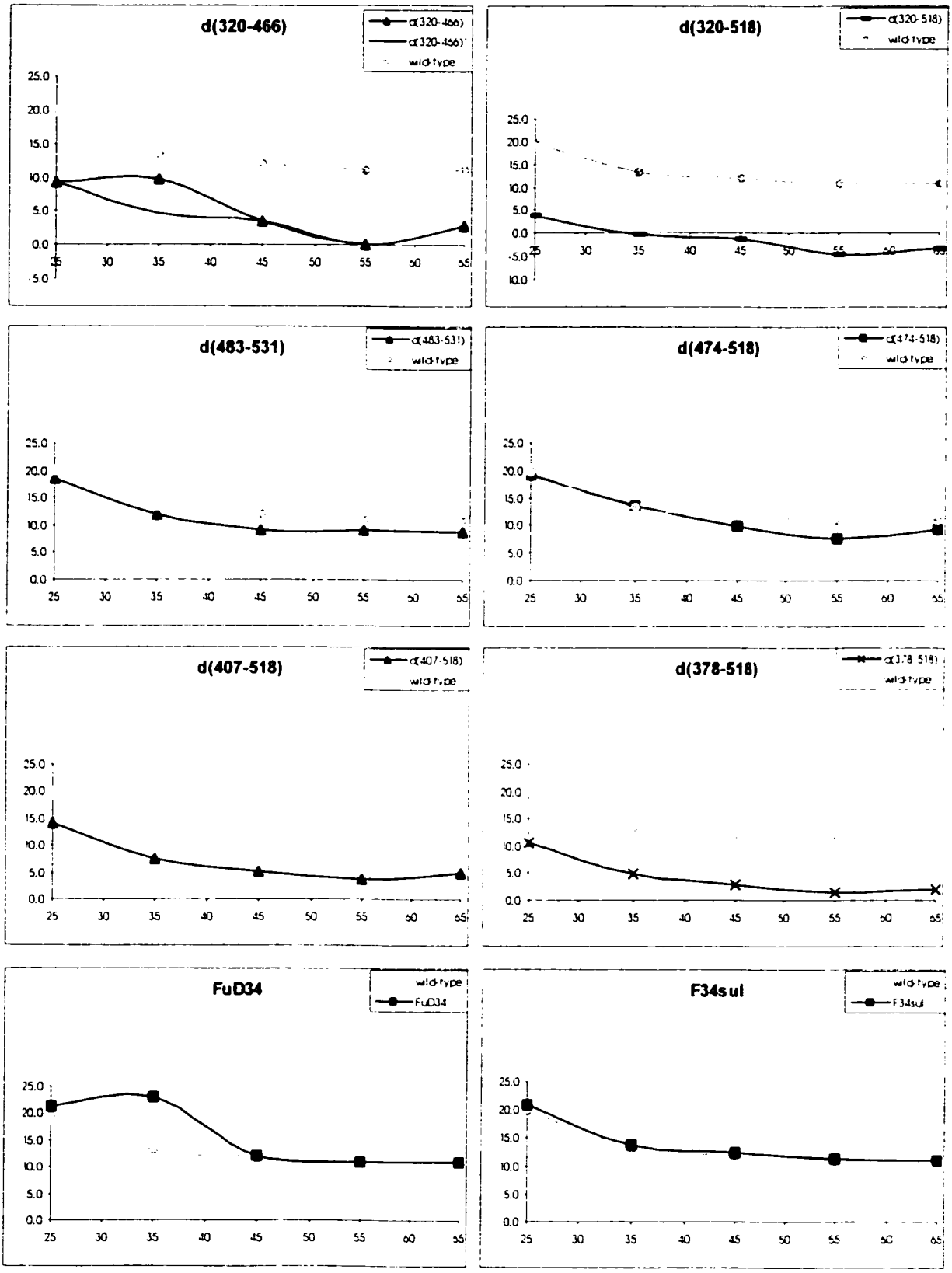

Figure 13 (Suite 2/2)

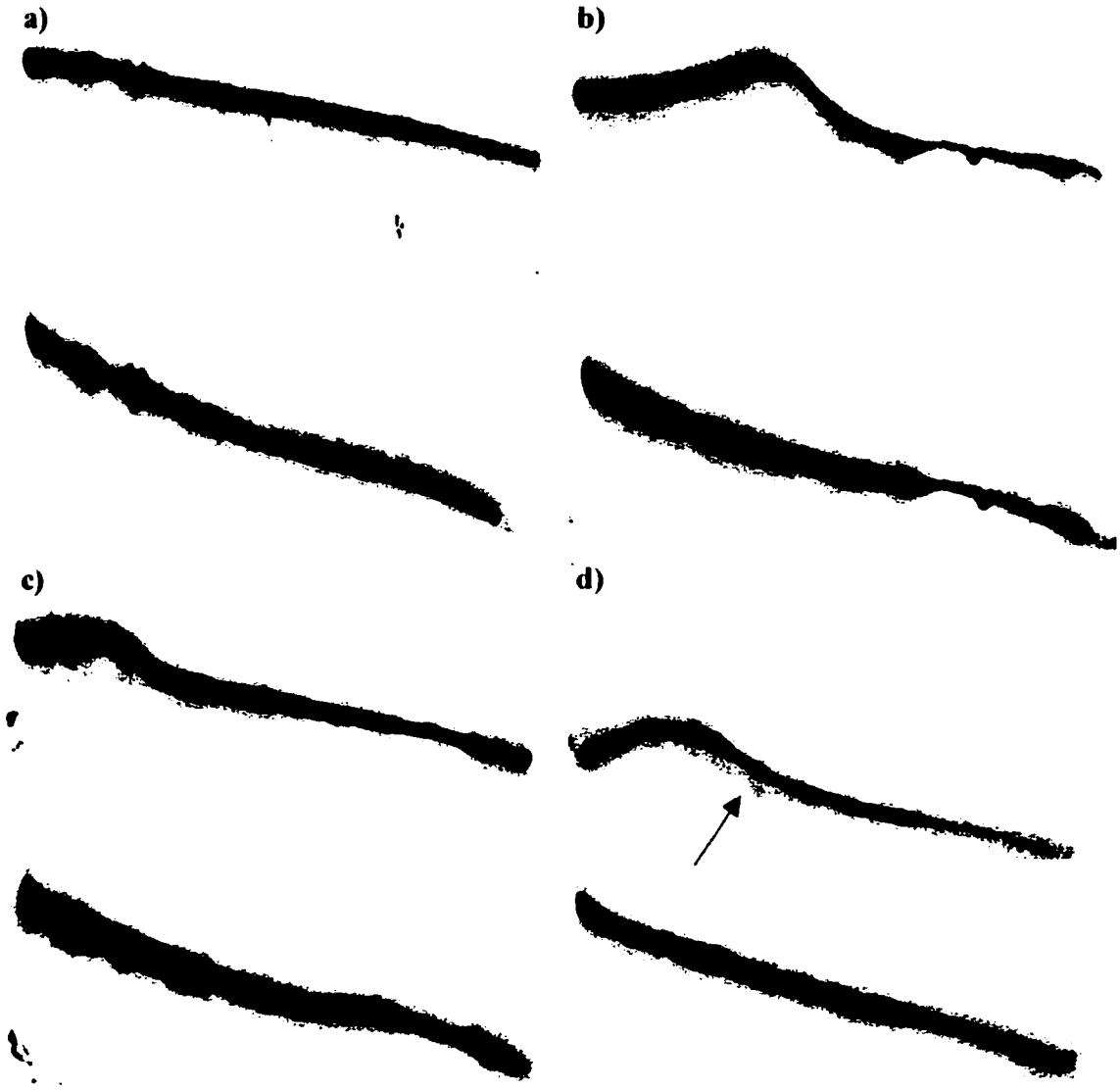

Figure 14. Temperature gradient gel electrophoresis. TGGE was used to analyze the structure of *psbC* 5' UTR with wild-type sequence (561 nt) **(a)**, *psbC-FuD34* (562 nt) **(b)**, *psbC*-Δ320-390 (491 nt) **(c)** and *psbC*-Δ320-466 (415 nt) **(d)**. A control corresponding to the *psbC* 3' UTR (225 nt) was included on each gel. The temperature gradient was from 25°C to 65°C (from right to left). An arrow indicates the presence of the *psbC*-Δ320-466 double conformation.

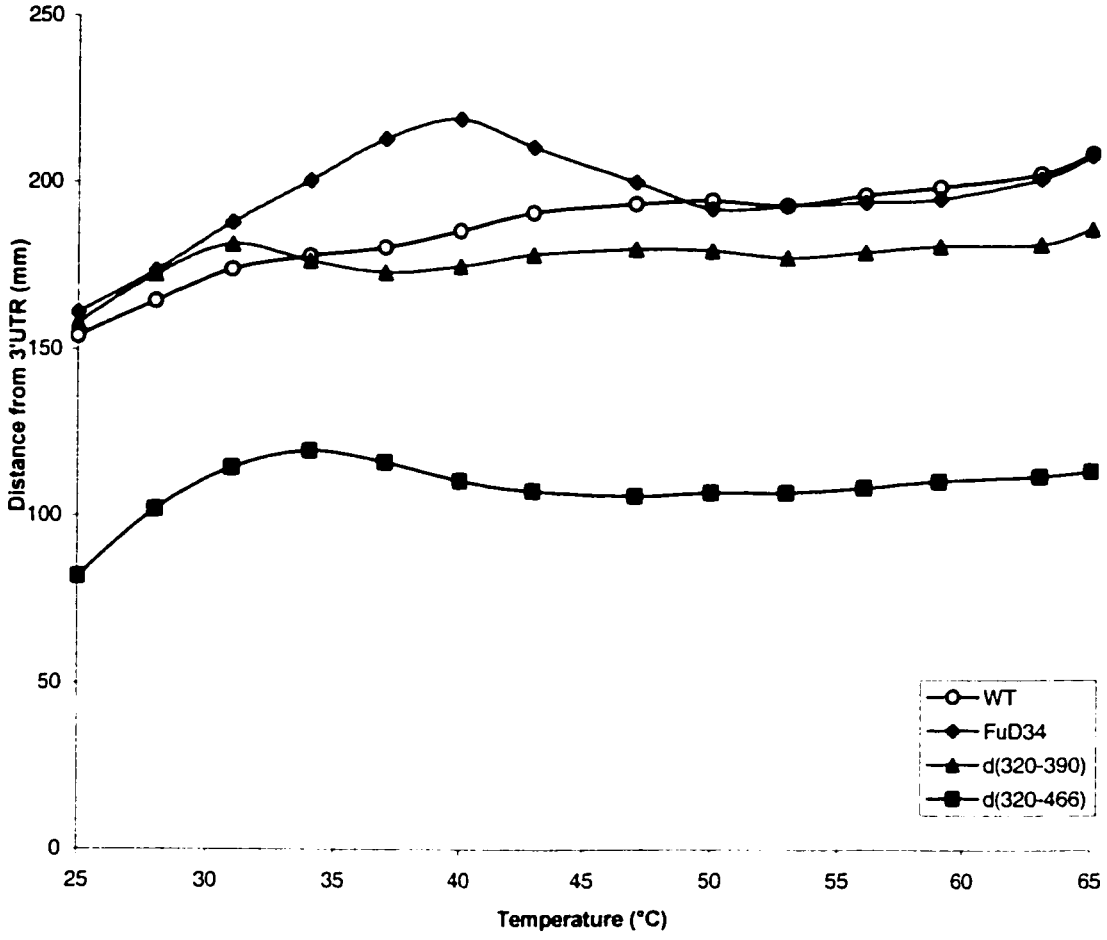

Figure 15. Representation of the distance migrated during TGGE. The distance between the four RNA samples tested and the 3' UTR template was determined at 10 mm intervals along the temperature gradient (corresponding to 3°C increments).

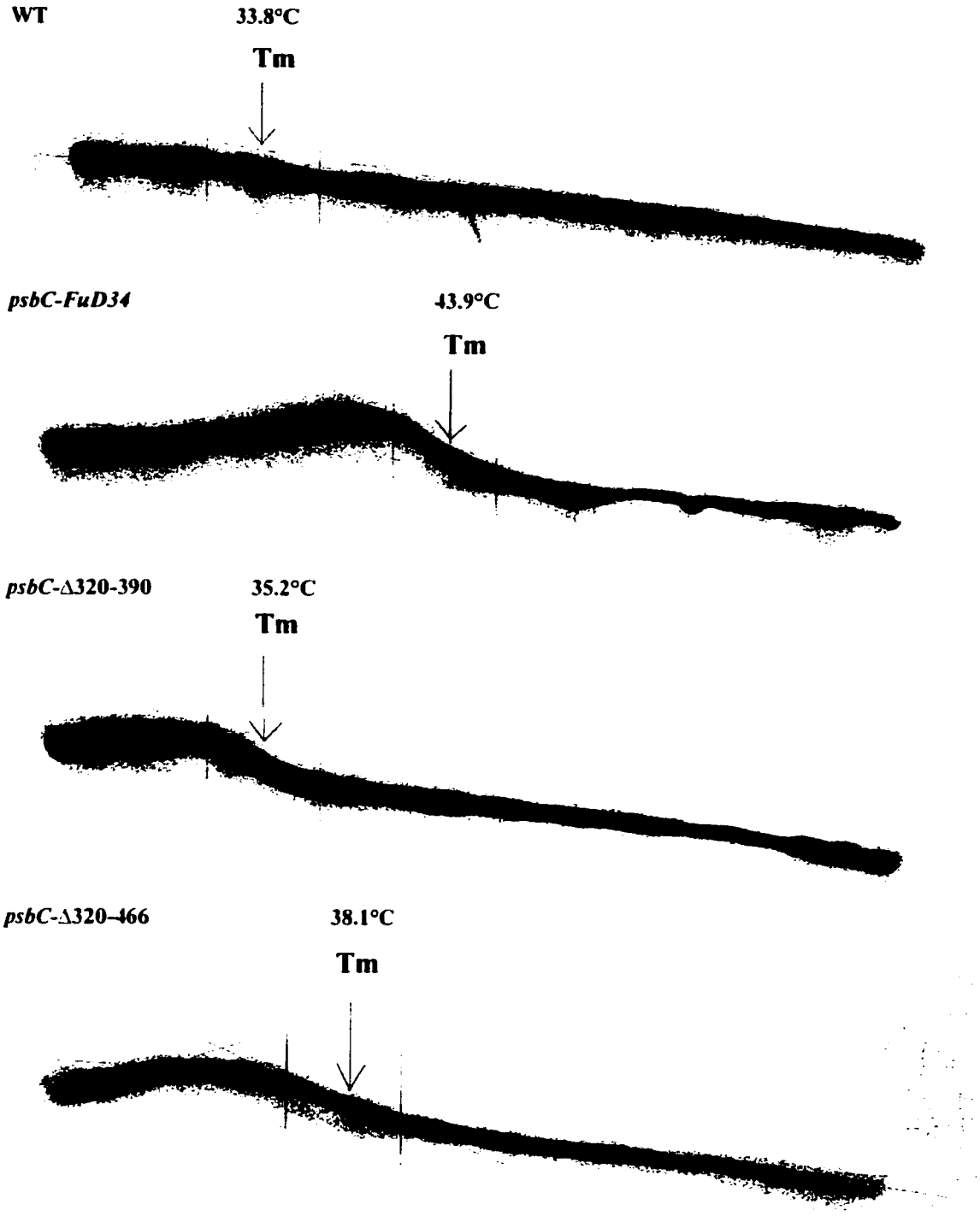

Figure 16. Structure melting temperatures were estimated from the TGGE results. The transition slope between two states allows us do determine the Tm of a given RNA structure.

3.5 mRNA synthesis and purification experiments abnormalities suggest highly stable secondary structure

PCR amplification done to produce DNA template for RNA synthesis, revealed an abnormality. A second and smaller fragment (300 bp smaller) was visible in wild type, *psbC-FuD34* and *psbC-F34sul*. However, this fragment was much stronger in *psbC-FuD34* (Figure 17). Since *psbC-FuD34* affects the 222-320 region by increasing the stem-loop stability, I hypothesize that one of the PCR primer stops along the stem of this structure, possibly at the *psbC-FuD34* mutation site, giving a fragment of approximately 300 bp. Alternatively, although less probable, it could be an artefact attributable to contamination present in every template tested. Another possibility will involve the primers hybridizing elsewhere. Attempts to clone and sequence this smaller fragment were unsuccessful.

As stated in Material and Methods, the recovery of 5' UTR synthesized mRNA was very low (10 to 20%) when purified using denaturing polyacrylamide gel electrophoresis and a significant loss occurred due to radiolabeled RNA which did not enter the gel. Therefore agarose gel was used to purify the synthesized mRNA. However, when a sample of the purified fragment was run on a second agarose gel (1.2% agarose at 12V/cm), two fragments distant from about 100 bp from each other were observed (Figure 18a). When purified separately, they were still running at their previous size (Figure 18b). Then the two fragments were run on a denaturing polyacrylamide gel electrophoresis, stained with ethidium bromide and visualized by UV light (Figure 18c). The larger fragment was seen running at its expected size; while the shorter was visible at the top of the well (some might have been washed out during the staining step). The

possibility of having a circular RNA that could present such behaviour (Dr David Stern, *personal communication*) was tested by phosphorylating the 5' end in the presence of [γ-^{32}P]ATP. As expected, the RNA fragment was radiolabeled and represents most of the radiolabeled material at top of the wells. Therefore, this RNA fragment does not join its 5' and 3'ends to form a circular molecule. DNase and RNase digestions were done as controls to make sure the fragment was RNA, as it was sensitive to the latter but not the former enzyme.

Figure 17. *psbC* 5' UTRs amplified by PCR show a minor smaller product. When run on an agarose gel (0.8 to 1%), minor PCR products can be observed. However, the 5' UTR with the *psbC-FuD34* mutation (lanes 5, 6 and 7) template generates more of the minor product of approximately 300 bp, than does either the wild-type sequence (lanes 1 and 2) or the 5' UTR with the *psbC-F34sul* mutation (lanes 3 and 4) do. Each lane has the PCR products derived from an independent plasmid preparation as template for the PCR reaction. (The 100 bp DNA ladder is shown in lane 8).

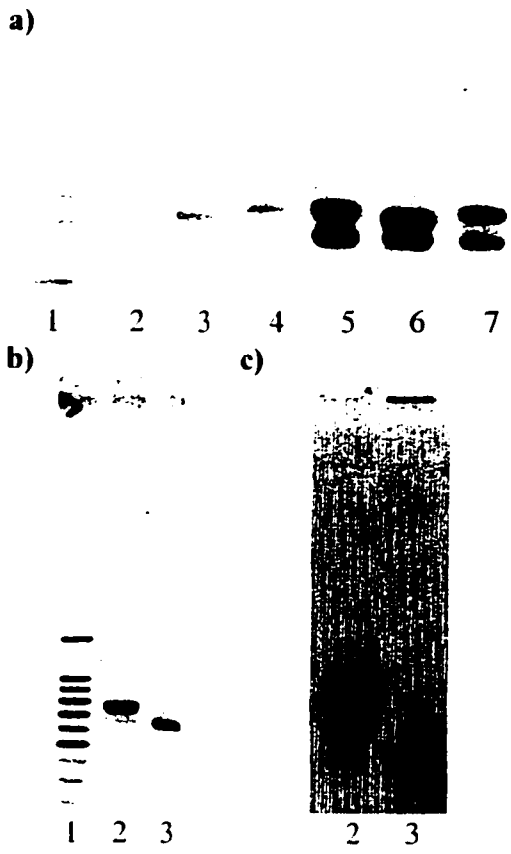

Figure 18. Synthesized 5' UTR present a second smaller fragment. **a)** A second smaller fragment (by approximately 100 bp) can be seen on 1.5% agarose gel for wild type (lanes 5), *psbC-FuD34* (lanes 6), and *psbC-F34sul* (lanes 7). Lanes 2 (wild type), 3 (*psbC-FuD34*) and 4 (*psbC-F34sul*) represent RNA purified by denaturing polyacrylamide, while lanes 5, 6 and 7 represent RNA purified by agarose gel. Lane 1 is the 100 bp DNA ladder. **b)** Wild type was purified from a) and visualized again on 1% agarose gel. The two fragments run independently (lanes 2 and 3). **c)** The two fragments from **b)** were then run on a denaturing polyacrylamide gel and stained with ethidium bromide. While the bigger fragment run at its expected size (lane 2), the smaller fragment is seen at the top of its well (lane 3).

4.0 DISCUSSION

4.1 Processing of the 5' UTR have still to be identified in *psbC*

One goal of this thesis was to determine whether the *psbC* mRNA has a minor form with a 5' terminal extension. Minor mRNA forms with 5' terminal extensions have been described previously for *psbA* (Bruick and Mayfield, 1998), *psbB* (Vaistij *et al*, 2000) and *psbD* (Nickelsen *et al*, 1999), encoding PSII reaction center subunits D1, P5 and D2, respectively in *C. reinhardtii*. These longer mRNA forms are thought to represent precursors that are processed by exo- or endoribonucleases to generate the 5' end of the mature mRNA. Processing events are thought to be required for mRNA stability and translation (reviewed by Monde *et al*, 2000). Indeed, mutants have been described that do not translate the mRNAs of *psbA*, *psbB* and *psbD* and accumulate only the longer form (Bruick and Mayfield, 1998; Nickelsen *et al*, 1999; Vaistij *et al*, 2000). Moreover, 3' end processing is required for translation of *atpB* and *rbcL* mRNAs in *C. reinhardtii* chloroplasts, thus supporting the idea that the 3' and 5' ends functionally interact to promote translation (Rott *et al*, 1998). Besides, D1, D2 and P5 appear to be translated in a coordinate way, while P6 is translated independently (de Vitry *et al*, 1989). Therefore, it is possible that processing is involved only in this coordinate translational regulation, and that the *psbC* mRNA is not processed in the same way.

My inability to consistently detect a longer *psbC* mRNA form during the course of this study does not mean that *psbC* mRNA is not processed. Primer extension analysis of the minor forms of the mRNA of *psbB* and *psbD* required extremely strong signals for the mature 5' end, which was not obtained in my experiments. Therefore, a possible

minor mRNA form could have been missed because the assay lacked sufficient sensitivity. Alternatively, 5' processing of *psbC* mRNAs could occur immediately after transcription and at a faster rate than for other mRNAs, such that the minor precursor form does not accumulate at all. Lastly, the inability to detect a non-processed RNA could be due to its rapid degradation after transcription.

While nuclease protection assays are used to confirm and complement primer extension results, previous attempts to hybridize RNA or DNA probes larger than oligonucleotides have been unsuccessful (Zerges, *personal communication*). A mutant that does not accumulate the *psbC* mRNA has been described (Sieburth *et al*, 1991) and analysis of its 5' and 3' termini could reveal the precursor form of the mRNA.

4.2 Primer extension and DMS footprinting indicate evidences that the *psbC* 5' UTR is highly structured 5' UTR

RNA can be probed for structure with dimethyl sulfate (DMS), which only methylates unpaired A and C bases at positions that arrest reverse transcriptase (reviewed by Ehresmann *et al*, 1987; Senecoff and Meagher, 1993). Thus, methylated A's and C's can be revealed using primer extension and identified as positions that are not involved in base-pairing. These data can be used as constraints with the *mfold* program to restrict the prediction of RNA structures to those that have bases involved in base-pairing, as revealed by this technique. This technique has identified the *in vivo* higher-order structure of three *cis*-elements that are required for translation of *petD* mRNA in the *C. reinhardtii* chloroplast (Higgs *et al*, 1999).

In the present study, I relied on nuclear and chloroplast mutants that have been previously described to interact with the 5' UTR of the *psbC* mRNA and thereby control translation from it (Rochaix *et al*, 1989; Zerges *et al*, 1997; Zerges and Rochaix, 1994). In particular, the ability to probe the accessibility of bases *in vivo* should allow the identification of structures and bound proteins that are altered by the *trans*-acting mutations in mutants for *TBC1*, *TBC2* and *TBC3* (see Introduction). The goal was partially achieved by the analysis of preliminary results with wild-type strains. Only a few reverse transcriptase stops were observed to result from methylation of the 5' UTR. Two reasons might account for this low amount of methylated nucleotides. First, the *psbC* 5' UTR could be highly structured and lack many single-stranded A's and C's that are accessible for DMS methylation. The second reason is the difficulty that I encountered to optimize the conditions for DMS treatment and RNA isolation. In addition, many reverse transcription stops were observed on non-methylated RNA, both from cells and transcribed *in vitro* from a plasmid template. Therefore, structures in the RNA stop reverse transcriptase and make it difficult to identify methylated positions. I first observed that increasing concentration of DMS alone or increasing the incubation time alone did not affect the methylation pattern (as was also observed by Dr David Higgs, *personal communication*). Afterwards, I found that a combination of 10 mM DMS final concentration, 5 to 7 minutes of incubation and 0.8 μg/ml RNA yielded methylation stops strong enough to be distinguished from the background stops. Temperature and salt concentration directly affect the secondary and tertiary structures adopted by RNAs (Draper and Gluick, 1995). Also, RNAs tend to adopt sub-optimal conformations in solution, possibly trapped intermediates in the folding process (Draper

and Gluick, 1995; Zuker, 1994). The DMS buffer that I originally used was designed to theoretically allow the mRNA structure to adopt a stable conformation. This can work well with RNA treated *in vitro* since we can denature with high temperature treatment and then allow the RNA to fold during subsequent gradual cooling to adopt its native conformation and interact with ions (especially the Mg^{2+}). No ions are known to interfere with the DMS (reviewed by Ehresmann *et al*, 1987).

Interestingly, structures predicted by computer were supported by the *in vivo* methylation pattern (Figure 11). The A's and C's predicted to be unpaired were methylated and paired A's and C's were not. One exception is A378, which is predicted to be involved in Watson-Crick base-pairing by *mfold*, but is unpaired *in vitro*, because it is methylated by DMS. However, *in vivo* it is not accessible to DMS and therefore, either paired our bound to protein. A378 is part of a short region with a sequence (from position 373 to 378) complementary to a region of the 222-320 stem-loop (region 268-273 at the top of the apical loop). Base-pairing between both regions could lead to the formation of a pseudoknot structure (Figure 11). Such an interaction requires that both regions be single-stranded (reviewed by Pleij and Bosch, 1989). Results from *in vivo* methylation patterns refute the formation of this pseudoknot, at least *in vitro*. However, it shows that A378 could be single-stranded, contrary to its predicted secondary structure. Therefore, the fact that both regions are not methylated *in vivo* make it possible that this pseudoknot does form *in vivo*, but its formation requires specific factors such as ions or protein interactions not present *in vitro* (Gluick *et al*, 1997; Gluick and Draper, 1994). Pseudoknot formation has been shown to be required for binding of a translational

regulatory factor to the 5' UTR of the bacteriophage T4 gene 32, which encodes a protein that is involved in DNA replication, recombination and repair (Shamoo *et al*, 1993).

Primer extension on non-methylated RNA revealed some natural reverse transcriptase obstacles, which differ between wild type and mutant 5' UTRs. Differences between reverse transcriptase stops on the wild type and the *psbC-FuD34* 5' UTRs bring us to consider a computer-predicted small stable stem-loop (region 133-153, see Figure 9a). Since *psbC-FuD34* is predicted to stabilize the stem of the 222-320 stem-loop, it is possible that both the 133-153 and 222-320 stem-loops interact in wild type. Such an interaction would require the intervention of a *trans*-acting factor, because no complementary sequence exists between these two regions. It is possible that a regulatory factor interacts with both structures (and specific regions within these structures), but the increased stability of the 222-320 stem-loop in *psbC-FuD34* prevents such interaction. However, this hypothesis does not explain why an interaction between the two stem-loops would prevent the premature arrest. Alternatively, it could be that this interaction involves non Watson-Crick base-pairing, and that *trans*-acting factors are required to repress this interaction.

Theoretically, we should not see any differences between the primer extension of the *psbC* mRNA isolated from the *TBC* mutants and wild type, since these mutants do not affect the sequence of the 5' UTR. One potential explanation is that the *TBC* mutants are affected in the post-transcriptional modification of bases, making them prone to reverse transcriptase stop. I found the same pattern of reverse transcription stops on the *psbC* 5' UTR synthesized from a plasmid template *in vitro* (data not shown), indicating that a structure and not modified bases are responsible for these blocks.

4.3 Variations in 5' UTR RNA electrophoretic mobility reveal that structural conformation and stability of the 222-320 stem-loop might be a key factor in *psbC* translation.

Temperature gradient gel electrophoresis (TGGE) has proven to be a good complement to RNA structural probing by chemicals and ribonucleases (Szewczak *et al.*, 1998). For example, a previous study used a gel-shift assay in combination with TGGE to study protein binding interactions with the spinach chloroplast *psbA* mRNA (Klaff *et al*, 1997). TGGE has also been used to study tertiary interactions that are required for stability and auto-catalytic splicing activity of the group I rRNA intron in *Tetrahymena* (Szewczak *et al*, 1998). Here TGGE was used to observe if the structure of the *psbC* 5' UTR is related to its ability to drive translation. While unfolding of the wild-type 5' UTR shows a regular and smooth structural transition on the TGGE, results shown in Figure 14 reveal that the unfolding pattern of several mutant 5' UTRs are quite different from the wild type. For example, the 5' UTR with the *psbC-FuD34* mutation, which affects the 222-320 stem-loop sequence and blocks the translation, has a dramatically altered melting profile. Usually, structured RNA shows an increase in electrophoretic mobility relative to the unfolded form, even at the tertiary structure level (Szewczak *et al*, 1998). The wild-type *psbC* 5' UTR can be perceived as a RNA that has no structure under the conditions of this experiment or one that unfolds in small discreet steps. Unfolding of the three mutant 5' UTRs happens at higher temperature, due to higher structural stability. The conformations adopted appear to be less hydrodynamic, and therefore decreasing the mobility (Szewczak *et al*, 1998). Retardation is especially pronounced for RNA molecules with internal loops, such as a spliced intron lariat

(Domdey *et al*, 1984). In the case of the mutants showing different TGGE patterns, complex structural transitions involving tertiary and secondary structure interactions could unfold faster, therefore reducing their mobility. Alternatively, it is possible that tertiary and secondary structures possess intermediate transition states conformation that hampered the RNA mobility in mutants. Deleting (Figure 13, *psbC-Δ222-320*) or replacing (Figure 12 lanes 13) the 222-230 stem-loop did not seem to affect the structure based on the mobilities of these RNAs. The shift in RNA mobility and Tm caused by particular structural conformations adopted by the mutants 5' UTRs during their unfolding, indicate that the global stability of secondary and tertiary structure is altered. None of these deletions is predicted to prevent or change the 222-320 stem-loop formation. Using varying amounts of Mg^{2+} would be required to investigate the thermodynamics of tertiary structures because this ion type is known to stabilize certain secondary and tertiary structures (Laing *et al.*, 1994).

Another striking feature is the presence of dual isoforms of some of the mutant 5' UTRs. The most evident is found in *psbC-Δ320-466*, which adopts two conformations at 35°C (Figures 12 and 14). Moreover, this higher mobility form seems to follow a wild-type transition sooner than its slower isoform. Careful observation of the *psbC-FuD34* TGGE pattern (Figure 14) also reveals a minor secondary melting process that does not undergo the mobility shift seen at 35°C for the major isoform. Analysis of this mutant RNA on denaturing polyacrylamide gels revealed a subpopulation with an extremely high apparent molecular weight (Figure 18), which is often indicative of RNA with a covalently closed loop, like spliced intron lariat structure (Domdey *et al*, 1984). Thus, these two gel systems revealed the existence of more than one form of the 5' UTR baring

the *psbC-FuD34* mutation. A series of DMS probing experiments at different temperatures and salt concentrations would allow a better interpretation of the TGGE results and could ultimately allow us to determine the relevant structure(s) to translational control and the interactions with the *TBC1*, *TBC2* and *TBC3* gene products.

Melting point temperature difference between the wild type (33.8°C) and *psbC-FuD34* (43.9°C) 5' UTRs enhance the impact of the 222-320 stem-loop stability on the mobility. Since the only difference reside in three nucleotides change in the 222-320 stem-loop, it is probable that this stem-loop as some role to play in *psbC* 5' UTR structural stability and is involved in tertiary interactions that affect its mobility. Computer predictions of the 5' UTR structure, gave 5' UTR secondary structures that start to unfold around 34°C. The 222-320 stem-loop region of the *psbC-FuD34* 5' UTR is predicted to unfold first by the apical loop around 54°C, while the wild-type stem-loop structure is predicted to unfold by the base of the stem beginning at 48°C (or at 43°C for the *psbC-F34sul* stem-loop). As the temperature increases around 34°C, only minor changes appear to occur in the overall 5' UTR secondary structure of wild type, *psbC-FuD34* and *psbC-F34sul*, and they might not affect the mobility as much as what we observe on a TGGE. Major secondary structural changes that should cause a shift in mobility are predicted to occur at 55°C, which is similar to the melting temperature previously determined in *rps7* chloroplast mRNA, based on changes of the absorbance of UV light at 260 nm with temperature (Fargo *et al*, 1999). Therefore, the low relative mobility isoforms of the mutant 5' UTRs could correspond to stabilized or novel tertiary structure(s), which typically melt at a lower temperature than do most secondary structure (Laing and Draper, 1994). Moreover, stabile tertiary structures involving distant bases on

the primary sequence could be expected to have a more drastic effect on mobility. This is especially true for long and complex mRNAs (Laing and Draper, 1994; Szewczak *et al*, 1998; Zuker, 1994). It should be noted, however, that these deviations from the expected mobility could be due to both secondary and tertiary structures.

It is possible that the central stem-loop is involved in a tertiary interaction that is required for translation. Since the transition occurs at a relatively low temperature, it suggests that it could reflect a physiologically relevant structural transition that regulates translation. While this study does not prove the presence of the 222-320 stem-loop, the *psbC-FuD34* mutation stabilizes a structure, based on the strong reverse transcriptase stops within this region and results of TGGE. This could mean that this mutation stabilizes a structure that inhibits translation, by preventing a structural interaction from occurring and blocking the system in the "off" state. Conformation changes in the mRNA have been described to play a role in translation. For example, the bacteriophage Mu *mom* mRNA is regulated at the level of translation (Wulczyn and Kahmann, 1991). The *Mom* protein protects the phage DNA from its host endonucleases. The *mom* mRNA contains a GUG initiation codon sequestered in a 53 nt stable stem-loop and its translation is regulated by the *Com* protein, which activates translation by melting this structure. Contrary to the major stem-loop structure in the 5' UTR of the *psbC* mRNA, however, mutations that destabilizes the structure led to constitutive *Com*-independent translation.

We do not know if these regions interact with protein. Evidence suggests that at least one of the *TBC* factors interacts directly or indirectly with this stem-loop to control the translation (see Introduction and Zerges *et al*, 1997; Zerges and Rochaix, 1994). One

potential role for one of these factors might be to set the 5' UTR in an appropriate conformation in order to allow the 30S ribosomal subunit, or another factor involved in translation, to bind the mRNA. Chloroplast translation processes are often compared to mitochondria. The 5' UTRs of mitochondrial mRNAs are also AU rich, of various lengths, and they interact with mRNA specific translational regulators (reviewed by Fox, 1996). The case of the *cox2* mRNA (encoding the cytochrome *c* oxidase subunit II) is particularly interesting, since a small stem-loop is involved in a tertiary interaction triggered by the nuclear factor *Pet111p* that is required for translation activation (Dunstan *et al*, 1997). It is suggested that this interaction could recruit and position the small ribosomal subunit or stabilize its binding to the initiation codon. A similar pathway could be used by *psbC* and involve the 222-320 stem-loop with one of the *TBC* factors to control translation.

5.0 CONCLUSION

In this study, DMS footprinting and TGGE were used to analyze wild type and mutant *psbC* 5' UTRs for RNA structures that affect its ability to promote translation. Differences in mobilities between wild type and mutants 5' UTRs bring evidence for the importance of a structure involving the 222-320 stem-loop and possibly neighbouring sequences in promoting translation from the GUG initiation codon. For example, both approaches provided evidence that *psbC-FuD34*, a mutation affecting a translational element in the middle of the 547 nt 5' UTR, stabilizes an RNA structure. This could correspond to a stabilization of a stem-loop structure, which has been predicted from the sequence (Rochaix *et al*, 1989). Other preliminary data from *in vivo* footprinting and structure probing with DMS indicate that the *psbC* 5' UTR is highly structured and revealed that some additional RNA structures are caused by mutations of the 5' UTR. Two regions have the potential to interact with the 222-320 stem-loop and therefore, are good candidates for further investigations. Future studies should use other approaches to characterize further this structure and address with biochemical assays and genetic experiments the precise roles of these structures and *TBC1*, *TBC2* and *TBC3* gene products in translation control.

6.0 REFERENCES

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 the chloroplast psbC mRNA in *Chlamydomonas reinhardtii*. Journal of Cell Biology *157*, 953-62.

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

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

Bennoun, P. and Beal, D. (1998). New Digital Imaging Instrument for Measuring Fluorescence and Delayed Luminescence. 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. 450-458.

Bevilacqua, J.M. and Bevilacqua, P.C. (1998). Thermodynamic analysis of an RNA combinatorial library contained in a short hairpin. Biochemistry *37*, 15877-15884.

Birmes, A., Sattler, A., Maurer, K.-H., and Riesner, D. (1990). Analysis of the conformational transitions of proteins by temperature-gradient gel electrophoresis. Electrophoresis *11*, 795-801.

Bonham-Smith, P.C. and Bourque, D.P. (1989). Translation of chloroplast-encoded mRNA: potential initiation and termination signals. Nucleic Acids Research *17*, 2057-80.

Brion, P., Michel, F., Schroeder, R., and Westhof, E. (1999). Analysis of the cooperative thermal unfolding of the *td* intron of bacteriophage T4. Nucleic Acids Research *27*, 2494-2502.

Brodersen, D.E., Clemons, W.M. Jr, Carter, A.P., Wimberly, B.T., and Ramakrishnan, V. (2002). Crystal structure of the 30S ribosomal subunit from *Thermus thermophilus*: structure of the proteins and their interactions with 16 S RNA. Journal of Molecular Biology *316*, 725-68.

Bruick, 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*, 1145-1153.

Cavalier-Smith, T. (2000). Membrane heredity and early chloroplast evolution. Trends in Plant Science *5*, 174-82.

Chen, X., Kindle, K., and Stern, D. (1993). Initiation codon mutations in the Chlamydomonas chloroplast *petD* gene result in temperature-sensitive photosynthetic growth. EMBO J *12*, 3627-35.

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*, 1295-305.

Craig, A.W., Haghighat, A., Yu, A.T., and Sonenberg, N. (1998). Interaction of polyadenylate-binding protein with the eIF4G homologue PAIP enhances translation. Nature *392*, 520-3.

Danon, A. (1997). Translational regulation in the chloroplast. Plant Physiology *115*, 1293-1298.

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

de Smit, M.H. and van Duin, J. (1990). Secondary structure of the ribosome binding site determines translational efficiency: a quantitative analysis. Proceedings of the National Academy of Sciences USA *87*, 7668-72.

de Vitry, C., Olive, J., Drapier, D., Recouvreur, M., and Wollman, FA. (1989). Posttranslational events leading to the assembly of photosystem II protein complex: a study using photosynthesis mutants from *Chlamydomonas reinhardtii*. Journal of Cell Biology *109*, 991-1006.

Domdey, H., Apostol, B., Lin, R.J., Newman, A., Brody, E., and Abelson, J. (1984). Lariat structures are *in vivo* intermediates in yeast pre-mRNA splicing. Cell *39*, 611-21.

Drager, R.G., Girard-Bascou, J., Choquet, Y., Kindle, K.L., and Stern, D.B. (1998). In vivo evidence for 5'-->3' exoribonuclease degradation of an unstable chloroplast mRNA. Plant Journal *13*, 85-96.

Drager, R.G., Higgs, D.C., Kindle, K.L., and Stern, D.B. (1999). 5' to 3' exoribonucleolytic activity is a normal component of chloroplast mRNA decay pathways. Plant Journal *19*, 521-31.

Draper, D.E. and Gluick, T.C. (1995). Melting studies of RNA unfolding and RNA-ligand interactions. Methods in Enzymology *259*, 281-305.

Draper, D.E., Xing, Y., and Laing, L.G. (1995). Thermodynamics of RNA unfolding: stabilization of a ribosomal RNA tertiary structure by thiostrepton and ammonium ion. Journal of Molecular Biology *249*, 231-8.

Dunstan, H.M., Green-Willms, N.S., and Fox, T.D. (1997). In vivo analysis of *Saccharomyces cerevisiae* COX2 mRNA 5'-untranslated leader functions in mitochondrial translation initiation and translational activation. Genetics *147*, 87-100.

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

Eibl, C., Zou, Z., Beck, a., Kim, M., Mullet, J., and Koop, H.U. (1999). In vivo analysis of plastid *psbA*, *rbcL* and *rpl32* UTR elements by chloroplast transformation: tobacco plastid gene expression is controlled by modulation of transcript levels

and translation efficiency. Plant Journal *19*, 333-45.

Erickson, J.M. (1998). Assembly of Photosystem II. 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. 255-285.

Esposito, D., Higgs, D.C., Drager, R.G., Stern, D.B., and Girard-Bascou, J. (2001). A nucleus-encoded suppressor defines a new factor which can promote *petD* mRNA stability in the chloroplast of *Chlamydomonas reinhardtii*. Current Genetics *39*, 40-48.

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 reporter mRNAs both in the *Chlamydomonas reinhardtii* chloroplast and in *Escherichia coli*. Molecular and Cellular Biology *19*, 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 General Genetics *264*, 291-299.

Fargo, D.C., Zhang, M., Gillham, N.W., and Boynton, J.E. (1998). Shine-Dalgarno-like sequences are not required for translation of chloroplast mRNAs in *Chlamydomonas reinhardtii* chloroplasts or in *Escherichia coli*. Molecular General Genetics *257*, 271-282.

Fox T.D. (1996). Genetics of Mitochondrial Translation. In Translational Control. Hershey J.W.B., Mathews 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.

Gluick, T.C. and Draper, D.E. (1994). Thermodynamics of folding a pseudoknotted mRNA fragment. Journal of molecular biology *241*, 246-262.

Gluick, T.C., Wills, N.M., Gesteland, R.F., and Draper, D.E. (1997). Folding of an mRNA pseudoknot required for stop codon readtrough: Effects of mono- and divalent ions on stability. Biochemistry *36*, 16173-16186.

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

Harris, E.H. (1989). The *Chlamydomonas* Sourcebook: A Comprehensive Guide to Biology and Laboratory Use (San Diego, CA: Academic Press).

Harris, H.E. (1998). Introduction to *Chlamydomonas*. 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. 1-11.

Hauser, C.R., Gillham, N.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), pp. 197-217.

Higgs, D.C., Shapiro, R.S., Kindle, K.L., 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*, 8479-8491.

Hirose, T., Kusumegi, T., Tsudzuki, T., and Sugiura, M. (1999). RNA editing sites in tobacco chloroplast transcripts: editing as a possible regulator of chloroplast RNA polymerase activity. Molecular General Genetics *262*, 462-467.

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

Hsu-Ching, C. and Stern, D.B. (1991). Specific binding of chloroplast proteins *in vitro* to the 3' untranslated region of spinach chloroplast *petD* mRNA. Molecular and Cellular Biology *11*, 4380-4388.

Innis, M.A., Myambo, K.B., Gelfand, D.H., and Brow, M.A. (1988). DNA sequencing with *Thermus aquaticus* DNA polymerase and direct sequencing of polymerase chain reaction-amplified DNA. Proceedings of the National Academy of Sciences USA *85*, 9436-9440.

Joliot, P., Beal, D., and Delosme, R. (1998). In vivo Measurements of Photosynthetic Activity. 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. 443-449.

Kapazoglou, A., Sagliocco, F., and Dure, L.3. (1995). PSII-T, a new nuclear encoded lumenal protein from photosystem II. Targeting and processing in isolated chloroplasts. Journal of Biological Chemistry *270*, 12197-202.

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*, 1468-79.

Knapp, G. (1989). Enzymatic approaches to probing of RNA secondary and tertiary structure. Methods in Enzymology *180*, 192-212.

Koo, J.S. and Spremulli, L.L. (1994). Effect of the secondary structure in the *Euglena gracilis* chloroplast ribulose-bisphosphate carboxylase/oxygenase messenger RNA on translation initiation. Journal of Biological Chemistry *269*, 7501-7508.

Kozak, M. (1999). Initiation of translation in prokaryotes and eukaryotes. Gene *234*, 187-208.

Kraus, B.L. and Spremulli, L.L. (1986). Chloroplast initiation factor 3 from *Euglena gracilis*. Identification and initial characterization. Journal of Biological Chemistry *261*, 4781-4.

Laing, L.G. and Draper, D.E. (1994). Thermodynamics of RNA folding in a conserved ribosomal RNA domain. Journal of Molecular Biology *237*, 560-576.

Laing, L.G., 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.

Lefebvre, P.A. and Silflow, C.D. (1999). *Chlamydomonas*: The cell and its genomes. Genetics *151*, 9-14.

Levy, H., Kindle, K.L., and Stern, D.B. (1999). Target and specificity of a nuclear gene product that participates in mRNA 3'-end formation in *Chlamydomonas* chloroplasts. Journal of Biological Chemistry *274*, 35955-62.

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

Makrides, S.C. (1996). Strategies for achieving high-level expression of genes in *Escherichia coli*. Microbiological Review *60*, 512-38.

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., dePamphilis, C.W., Cui, L., Miller, W., Harris, E.H., Stern, D.B. *in press*. The *Chlamydomonas reinhardtii* plastid chromosome: Islands of genes

in sea of repeats. Plant Cell.

Mayfield, S.P., Cohen, A., Danon, A., and Yohn, C.B. (1994). Translation of the *psbA* mRNA of *Chlamydomonas reinhardtii* requires a structured RNA element contained within the 5' untranslated region. Journal of Cell Biology *127*, 1537-45.

McCarthy, J.E. and Brimacombe, R. (1994). Prokaryotic translation: the interactive pathway leading to initiation. Trends in Genetics *10*, 402-7.

Mets, 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.

Monde, R.A., Schuster, G., and Stern, D.B. (2000). Processing and degradation of chloroplast mRNA. Biochimie *82*, 573-82.

Morais, F., Barber, J., and Nixon, P.J. (1998). The chloroplast-encoded alpha subunit of cytochrome b-559 is required for assembly of the photosystem two complex in both the light and the dark in *Chlamydomonas reinhardtii*. Journal of Biological Chemistry *273*, 29315-20.

Muhlrad, D., Decker, C.J., and Parker, R. (1995). Turnover mechanisms of the stable yeast PGK1 mRNA. Molecular and Cellular Biology *15*, 2145-56.

Nickelsen, J. (2000). Mutations at three different nuclear loci of *Chlamydomonas* suppress a defect in chloroplast *psbD* mRNA accumulation. Current Genetics *37*, 136-142.

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*. Plant Cell *11*, 957-970.

Ossenbuhl, F. and Nickelsen, J. (2000). *cis*- and *trans*-acting determinants for translation of *psbD* mRNA in *Chlamydomonas reinhardtii*. Molecular and Cellular Biology *20*, 8134-8142.

Pain, V.M. (1996). Initiation of protein synthesis in eukaryotic cells. European Journal of Biochemistry *236*, 747-71.

Pleij, C.W.A. and Bosch, L. (1989). RNA pseudoknots: structure, detection, and prediction. Methods in Enzymology *180*, 289-303.

Redding, K. and Peltier, G. (1998). Reexamining the validity of the Z-scheme: Is photosystem I required for oxygenic photosynthesis in *Chlamydomonas*. 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. 349-362.

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

Rochaix, J.D. (2001a). Posttranscriptional control of chloroplast gene expression. From RNA to photosynthetic complex. Plant Physiology *125*, 142-144.

Rochaix, J.D. (2001b). Assembly, function, and dynamics of the photosynthetic machinery in *Chlamydomonas reinhardtii*. Plant Physiology *127*, 1394-8.

Rochaix, J.D., Kuchka, M., Mayfield, S., Schirmer-Rahire, M., 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*. EMBO J *8*, 1013-1021.

Rott, R., Levy, H., Drager, R.G., Stern, D.B., and Schuster, G. (1998). 3'-Processed mRNA is preferentially translated in *Chlamydomonas reinhardtii* chloroplast. Molecular and Cellular Biology *18*, 4605-4611.

Rozen, F., Edery, I., Meerovitch, K., Dever, T.E., Merrick, W.C., and Sonenberg, N. (1990). Bidirectional RNA helicase activity of eukaryotic translation initiation factors 4A and 4F. Molecular and Cellular Biology *10*, 1134-1144.

Rutherford, A.W. (1989). Photosystem II, the water-splitting enzyme. Trends in Biochemical Sciences *14*, 227-232.

Sakamoto, W., Sturm, N.R., Kindle, K.L., and Stern, D.B. (1994). *petD* mRNA maturation in *Chlamydomonas reinhardtii* chloroplasts: role of 5' endonucleolytic processing. Molecular and Cellular Biology *14*, 6180-6186.

Sambrook, J. and Russell D.W. (2000). Molecular Cloning: A Laboratory Manual 3rd ed (New York: Cold Spring Harbor Laboratory).

Sanger, F., Nicklen, S., and Coulson, A.R. (1977). DNA sequencing with chain-terminating inhibitors. Proceedings of the National Academy of Sciences USA *74*, 5463-7.

Senecoff, J.F. and Meagher, R.B. (1992). In vivo analysis of plant RNA structure: soybean 18S ribosomal and ribulose-1,5-bisphosphate carboxylase small subunit RNAs. Plant Molecular Biology *18*, 219-34.

Senecoff, J.F. and Meagher, R.B. (1993). In vivo analysis of plant 18S ribosomal RNA structure. Methods in Enzymology *224*, 357-372.

Shamoo, Y., Tam, A., Konigsberg, W.H., and Williams, K.R. (1993). Translational repression by the bacteriophage T4 gene 32 protein involves specific recognition of an RNA pseudoknot structure. Journal of Molecular Biology *232*, 89-104.

Sieburth, L., Berry-Lowe, S., and Schmidt, G.W. (1991). Chloroplast RNA stability in *Chlamydomonas*: Rapid degradation of *psbB* and *psbC* transcripts in two nuclear mutants. Plant Cell *3*, 175-189.

Simpson, C. and Stern, D. (2001). *Chlamydomonas reinhardtii* as a model system for dissecting chloroplast RNA processing and Decay Mechanisms. Methods in Enzymology *342*, 384-407.

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

Stern, D.B. and Gruissem, W. (1987). Control of plastid gene expression: 3' inverted repeats act as mRNA processing and stabilizing elements, but do not terminate transcription. Cell *51*, 1145-1157.

Sugiura, M., Hirose, T., and Sugita, M. (1998). Evolution and mechanism of translation in chloroplasts. Annual Review of Genetics *32*, 437-459.

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.

Tinoco, I.jr. and Bustamante, C. (1999). How RNA folds. Journal of Molecular Biology *293*, 271-281.

Vaistij, F., Goldschmidt-Clermont, M., Wostrikoff, K., and Rochaix, J.D. (2000). Stability determinants in the chloroplast *psbB/T/H* mRNAs of *Chlamydomonas reinhardtii*. Plant Journal *21*, 469-482.

van Wijk, K.J., Roobol-Boza, M., Kettunen, R., Andersson, B., and Aro, E.M. (1997). Synthesis and assembly of the D1 protein into photosystem II: processing of the C-terminus and identification of the initial assembly partners and complexes during photosystem II repair. Biochemistry *36*, 6178-86.

Wang, C.C., Roney, W.B., Alston, R.L., and Spremulli, L.L. (1989). Initiation complex formation on *Euglena* chloroplast 30S subunits in the presence of natural mRNAs. Nucleic Acids Research *17*, 9735-47.

Wartell, R.M., Hosseini, S.H., and Moran, C.P.Jr. (1990). Detecting base pair substitutions in DNA fragments by temperature-gradient gel electrophoresis. Nucleic Acids Research *18*, 2699-2705.

Weeks, D.P. (1992). *Chlamydomonas*: an increasingly powerful model plant cell system. Plant Cell *4*, 871-878.

Wells, S.E., Hughes, J.M.X., Igel, A.H., and Ares, M.Jr. (2000). Use of dimethyl sulfate to probe RNA structure *in vivo*. Methods in Enzymology *318*, 479-493.

Wulczyn, F.G. and Kahmann, R. (1991). Translational stimulation: RNA sequence and structure requirement for binding of COM protein. Cell *65*, 259-269.

Yamaguchi, K. and Subramanian, A.R. (2000). The plastid ribosomal proteins. Identification of all the proteins in the 50S subunit of an organelle ribosome (chloroplast). Journal of Biological Chemistry *275*, 28466-82.

Yamaguchi, K., von Knoblauch, K., and Subramanian, A.R. (2000). The plastid ribosomal proteins. Identification of all the proteins in the 30S subunit of an organelle ribosome (chloroplast). Journal of Biological Chemistry *275*, 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 the National Academy of Sciences USA *95*, 2238-43.

Zerges, W. (2000). Translation in chloroplast. Biochimie *82*, 583-601.

Zerges, W. (2002). Does complexity constrain organelle evolution? Trends in Plant Science *7*, 175-182.

Zerges, W., Auchincloss, A.H., and Rochaix, J.D. (submitted). Translation of the chloroplast *psbC* mRNA is controlled by interactions between RNA elements in the 5' Untranslated region and three nuclear gene products in *Chlamydomonas reinhardtii*.

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*, 3440-3448.

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*, 5268-5277.

Zerges, W. and Rochaix, J.D. (1998). Low density membranes are associated with RNA-binding proteins and thylakoids in the chloroplast of *Chlamydomonas reinhardtii*. Journal of Cell Biology *140*, 101-10.

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