Biochemical and Biophysical Research Communications 452 (2014) 14–20

Contents lists available at ScienceDirect

Biochemical and Biophysical Research Communications

journal homepage: www.elsevier.com/locate/ybbrc

Viral and chloroplastic signals essential for initiation and efficiency of translation in *Agrobacterium tumefaciens*

Tauqeer Ahmad, Srividhya Venkataraman, Kathleen Hefferon, Mounir G. AbouHaidar *

Department of Cell and Systems Biology, University of Toronto, St. George Campus, 25 Willcocks Street, Toronto, ON M5S3B2, Canada

ARTICLE INFO

Article history:
Received 14 July 2014
Available online 10 August 2014

Keywords:
Agrobacterium tumefaciens
Translation
T7 SD
RBS
ALMV translation enhancer
rbcL

ABSTRACT

The construction of high-level protein expression vectors using the CaMV 35S promoter in concert with highly efficient translation initiation signals for *Agrobacterium tumefaciens* is a relatively less explored field compared to that of *Escherichia coli*. In the current study, we experimentally investigated the capacity of the CaMV 35S promoter to direct GFP gene expression in *A. tumefaciens* in the context of different viral and chloroplastic translation initiation signals. GFP expression and concomitant translational efficiency was monitored by confocal microscopy and Western blot analysis. Among all of the constructs, the highest level of translation was observed for the construct containing the phage T7 translation initiation region followed by the chloroplastic Rubisco Large Subunit (rbcL) 58-nucleotide 5′ leader region including its SD-like sequence (GGGAGGG). Replacing the SD-like (GGGAGGG) with non SD-like (TTTATTT) or replacing the remaining 52 nucleotides of rbcL with nonspecific sequence completely abolished translation. In addition, this 58 nucleotide region of rbcL serves as a translational enhancer in plants when located within the 5′ UTR of mRNA corresponding to GFP. Other constructs, including those containing sequences upstream of the coat proteins of *Alfalfa Mosaic Virus*, or the GAGG sequence of T4 phage or the chloroplastic atpI and/or PsbA 5′ UTR sequence, supported low levels of GFP expression or none at all. From these studies, we propose that we have created high expression vectors in *A. tumefaciens* and/or plants which contain the CaMV 35S promoter, followed by the translationally strong T7 SD plus RBS translation initiation region or the rbcL 58-nucleotide 5′ leader region upstream of the gene of interest.

1. Introduction

Initiation of translation in *Escherichia coli* involves base pairing between a purine-rich Shine–Dalgarno (SD) domain at the 5′ untranslated region (5′ UTR) of mRNA, and the complementary anti-SD sequence at the 3′ end of 16S rRNA [1]. There are distinct sequence elements of the translation initiation region known to contribute to its efficiency [2]: the initiation codon, the Shine–Dalgarno (SD) sequence [3,4] as well as regions upstream of the SD sequence and downstream of the initiation codon, described as enhancers of translation [5]. The distance between the SD sequence and the initiation triplet has a marked effect on the efficiency of translation [6]. The 6-nucleotide consensus SD AGGAGG core sequence causes the highest level of protein synthesis.

Chloroplasts have their own translation system, which exhibits strong homologies to that of prokaryotes. This is consistent with the presence of a Shine–Dalgarno (SD) sequence (GGAGG) located within 12 nucleotides of the AUG initiation codon of many plastid genes [7]. Moreover, the sequence near the 3′ end of the plastid 16S rRNA contains a highly conserved polypyrimidine-rich region (CCUCC) complementary to the SD sequence as in bacteria. Over 90% of higher plant chloroplast genes encoding polypeptides possess an upstream sequence similar to the bacterial SD sequence. Spacing of these chloroplast SD-like sequences is less conserved, ranging from −2 to −29 nucleotides [8]. Translation of several chloroplast mRNAs is also regulated in response to light as well as to some nuclear-encoded factors. In this regard, it is interesting to study how well chloroplastic translational machinery function in Eubacteria such as *E. coli* and *Agrobacterium tumefaciens*.

The transfer of T-DNA from *Agrobacterium* into the plant genome represents a natural horizontal gene transfer across kingdom barriers and implicates a closer evolutionary relationship between *Agrobacterium* and plants than between any other Eubacterial organism (such as *E. coli*) and plants. The aim of the present study

Abbreviations: CaMV, cauliflower mosaic virus; GFP, green fluorescent protein; NOS, nopaline synthase; SD, Shine–Dalgarno sequence; RBS, ribosome binding site; rbcL, Rubisco Large Subunit.

* Corresponding author. Fax: +1 416 978 5878.
E-mail address: mounir.abouhaidar@utoronto.ca (M.G. AbouHaidar).

http://dx.doi.org/10.1016/j.bbrc.2014.07.135
0006-291X/

is to investigate the sequence determinants responsible for efficient translation in *A. tumefaciens*, which on the one hand is highly similar to *E. coli* in terms of its dependency on the SD sequence for the translation, while on the other hand is also mechanistically similar to chloroplast genes such as the large subunit of the Rubisco in its dependence on the 5′ upstream control region. Also, the essential molecular determinants for the design of an ideal Agrobacterial expression vector are considered.

2. Materials and methods

2.1. Construction of GFP expression plasmids

The binary vector pCAMBIA1300 (CAMBIA, Canberra, Australia) was used in this study. A 35S: sGFP:NOS expression cassette (GenBank EF546437) of size 1.9-kbp was subcloned into this vector through *Hind*III and *Eco*RI sites and designated pC-GFP (Fig. 1A). To create the pCTCR-GFP construct, the translation control region (TCR) [9], comprised of 58 nucleotides of 5′ UTR and 45 nucleotides from the N-terminal coding region of the rbcL gene were synthesized and cloned into pUC57 plasmid (Bio Basic Inc.) Following digestion of pUC57 by *Kpn*I/*Bam*HI and *Xba*I/*Bgl*II, respectively, and gel purification (QIAquick Gel Extraction Kit, QIAgen), rbcL TCR DNA fragments were subcloned into the pC-GFP binary plasmid using the respective restriction sites. All other vectors of the pC-GFP series were produced by ligating double-stranded oligonucleotides into restriction-enzyme digested plasmid DNA with compatible ends (Table 1). Briefly, complementary oligonucleotides synthesized by Eurofins MWG Operon (Huntsville, AL) were mixed in equimolar amounts (50 μM each), boiled and annealed by cooling to room temperature and ligated into a previously restriction enzyme digested pC-GFP vector using T4 DNA ligase (New England Biolabs) according to the manufacturer's protocol. The product of each ligation reaction was used to transform *E. coli* DH5-alpha competent cells and kanamycin (50 μg/ml) resistant bacterial colonies were screened for the presence of the proper recombinant constructs. The presence and accuracy of the inserted gene within the expression cassette in the final recombinant constructs was confirmed by DNA sequencing (The Centre for Applied Genomics, Toronto, Canada) using the GFP-R reverse primer:5′-AAGTCGTGCTGCTTCATGTG-3′.

2.2. Agrobacterium transformation

A modified freeze–thaw method for transformation of *A. tumefaciens* was used as reported previously [10]. After transformation, the cells were resuspended in LB such that all the samples contained a uniform OD_{595} of 1.0. From this, equal culture amounts were in turn taken to perform the downstream RNA, confocal microscopy and Western analyses.

2.3. RNA isolation, reverse transcription and PCR

Total RNA was isolated according to modified method of Abou-Haidar et al. [11] and subjected twice to DNase I treatments (New England Biolabs, NEB). A reverse transcription reaction of each sample was performed on 1 μg of total RNA with 200 units of M-MLV reverse transcriptase (Promega), 200 ng of GFP/16sRNA

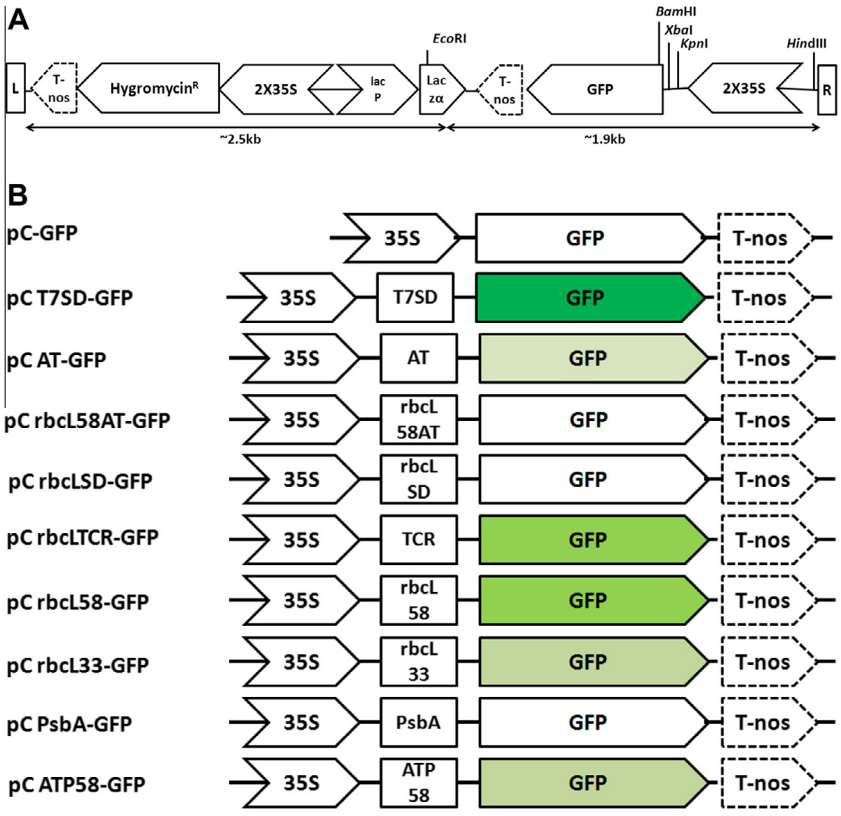

Fig. 1. Physical maps of constructs. (A) Modified pCAMBIA1300 with 35S Promoter, sGFP gene and Tnos terminator cassette. (B) Schematic representation of constructs used in this study. Arrows indicate the direction of transcription and translation. 35S is the CaMV 35S promoter. T-nos: represents the transcription terminator; the box between 35S and GFP contains the different translation initiation contexts. The GFP box is differently colored to reflect the efficiency of its expression. The dark green box representing T7SD shows the highest expression, followed by the rbcL TCR and the rbcL 58 nucleotide 5′ UTR region (light green). Boxes in light green represent marginal GFP expression, while unfilled boxes illustrate no GFP expression. *Note:* Figures not drawn to scale. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 1
Sequences of the translation initiation signals in pCAMBIA vectors. Italic letters indicate restriction site overhangs. Underlined capitalized bold letters indicate SD-sequences. Upper case bold letters indicate start codons. Sequence of the sense and antisense primers used to generate the various constructs is shown.

Vector	Description	Oligonucleotide/DNA sequence (5′ → 3′)
pC T7SD-GFP	Construct with PhageT7 trailer sequence (T7 translational enhance RBS) and is available in pET-X-series	Sense (*Xba*I overhang) *ctag*ttaataattttgtttaacttta**GAAGGAG**atatacatATG*g* Antisense (*Bam*HI overhang) *gatcc*CATatgtatat**CTCCTTC**ttaaagttaaacaaaattatta*a*
pC rbc58-GFP	Construct with only 58 nucleotides of 5′ UTR of rbcL gene	Sense (*Xba*I overhang) *ctag*tgtcgagtagaccttgttgttgtgagaattcttaattcatgagttgta**GGGAGGG**atttATG*g* Antisense (*Bam*HI overhang) *gatcc*CATaaat**CCCTCCC**tacaactcatgaattaagaattctcacaacaacaaggtctactcgac*a*
PC rbc58AT-GFP	Construct with 58 nucleotides of 5′ UTR of rbcL gene where GGGAGGG sequence is replaced with TTTATTT	Sense (*Xba*I overhang) *ctag*tgtcgagtagaccttgttgttgtgagaattcttaattcatgagttgta**TTTATTT**atttATG*g* Antisense (*Bam*HI overhang) *gatcc*CATaaat**AAATAAA**tacaactcatgaattaagaattctcacaacaacaaggtctactcgac*a*
pC rbc33-GFP	Construct with 33 nucleotides of 5′ UTR of *rbcL gene*	Sense (*Xba*I overhang) *ctag*taattcttaattcatgagttgta**GGGAGGG**atttATG*g* Antisense (*Bam*HI overhang) *gatcc*CATaaat**CCCTCCC**tacaactcatgaattaagaatt*a*
pC rbcSD-GFP	Construct with only SD sequence of *rbcL* gene, the 5′ UTR sequence is replaced with non rbcL sequence	Sense (*Kpn*I overhang) *gtac*attgaacagttaagtttccattgatactcgaaagatgtcagcacca**GGGAGGG***g* Antisense (*Bam*HI overhang) *gatcc***CCCTCCC**tggtgctgacatctttcgagtatcaatggaaacttaactgttcaa*t*
pC PsbA-GFP	Construct with 85 nucleotides of 5′′ UTR of *PsbA* gene	Sense (*Xba*I overhang) *ctagt*aaaaagccttccattttctattttgatttgtagaaaactagtgtgcttG**GGAG**tccc**TGATGAT**taaataaacc**AAG**attttaccATG*g* Antisense (*Bam*HI overhang) *gatcc*CATggtaaaat**CTT**ggtttattta**ATCATCA**ggga**CTCC**caagcacactagttttctacaaatcaaaatagaaaatggaaggctttt*a*
pC ATP58	Construct with 58 nucleotides of 5′ UTR of ATP*I* gene	Sense (*Xba*I overhang) *ctagt*agatggttgaatcaaaaaatttgtttaaagttcaatttttca**GAGGGCAAGG****caat**ATGg Antisense (*Bam*HI overhang *gatcc*CATattg**CCTTGCCCTC**tgaaaaaattgaactttaaacaaaatttttgattcaaccatct**a**
pC AT-GFP	Construct with 5′ UTR of capsid protein of *Alfalfa Mosaic Virus* RNA	Sense (*Kpn*I overhang) *gtac*a**g**ttttttatttttaattttctttcaaatacttccaggatctcta**GAg** Antisense (*Bam*HI overhang) *gatc***CTC**tagagatcctggaagtatttgaaagaaaattaaaaataaaaac*t*
pC TCR-GFP	Construct with 58 nucleotides of 5′ UTR and 45 nucleotides from N-terminal coding region of rbcL gene. The required DNA fragment was synthesized. Sequence of only plus strand is given	*ctagt*gtcgagtagaccttgttgttgtgagaattcttaattcatgagttgta**GGGAGGG**atttATGtcaccacaaacagagactaaagcaagtgttggattcaaagct*g*

reverse primer and 500 μM dNTPs, in a final volume of 20 μl as recommended. Primers GFP 5′-ACGTAAACGGCCACAAGTTC-3′ (forward) and GFP 5′-AAGTCGTGCTGCTTCATGTG-3′ (reverse) were used to amplify a 187 bp fragment of the GFP gene. Primers 16SrRNA 5′-AACACATGCAAGTCGAACGC-3′ (forward) and 16SrRNA-R 5′-TAGGCCTTTACCCCACCAAC-3′ (reverse) were used to amplify a 187 bp fragment of *Agrobacterium* 16S rRNA as an internal and comparative control for semi-quantitative PCR.

2.4. Detection of GFP expression

2.4.1. Confocal microscopy

Following Agro-transformation, cell samples each containing OD_{595} to 1.0 were pelleted by centrifugation, and the pellets were resuspended in 10 mM MES (4-Morpholineethanesulfonic acid sodium salt) buffer, pH 5.7. A drop of each cell culture was overlaid on a glass slide and live cell imaging was performed using a confocal microscope (TCS SP5, Leica Microsystems) with a 100× oil objective lens. The 488-nm laser was used for GFP imaging. Differential interference contrast (DIC) microscopy was used for comparative studies of all the constructs. Images were analyzed by Leica Application Suite Advanced Fluorescence (LASAF) software.

2.4.2. SDS–PAGE and Western blot analysis

Following Agro-transformation, the cell samples, each adjusted for an OD_{595} of 1.0, were harvested by centrifugation and protein from each pellet was extracted using TMPDTNU (50 mM Tris, 20 mM $MgCl_2$, 1 mM PMSF, 100 mM DTT, 2% Trition X-100, 0.5% NP-40 and 8 M urea) buffer. Equivalent protein amounts were loaded as determined by the Bradford Protein Assay reagent kit (Bio-Rad, Hercules, CA). Coomassie Brilliant Blue R-250 staining. SDS–PAGE and Western blot analyses were performed according to Sambrook et al. [12].

3. Results and Discussion

GFP expression in ten pCAMBIA constructs (Fig. 1B) containing different translation initiation contexts upstream of the GFP gene was monitored by confocal microscopy and Western blot analysis, after transformation of *A. tumefaciens* (JV3101 strain) with the respective constructs. All constructs uniformly contained the CaMV 35S promoter and the enhanced GFP gene followed by the 3′ NOS terminator. This produces the same GFP transcript levels for all the constructs. The only difference between the constructs was in the sequence of the translation context upstream of the GFP coding sequence, which resulted in differential GFP expression.

3.1. Estimation of equal GFP transcript levels in A. tumefaciens harboring each of the above constructs

Transcription levels of the GFP mRNA for all the constructs were measured by semi-quantitative RT-PCR experiments with 16S rRNA expression levels as the internal control (Fig. 2, Top Panel). The transcriptional efficiency of the CaMV 35S promoter was also compared to that of the ribosomal RNA (rrn) promoter, as the latter uniformly showed similar high-level stable expression in all the cells harboring the respective constructs. PCR reactions in the above experiment were extended only up to 20 cycles in order to enable quantitation of the RNAs at the log phase before cDNA synthesis reached saturation levels. We observed that the GFP mRNA expression was uniform and the transcript levels corresponded to 1/6 dilution of the 16S rRNA in all the cells harboring the respective constructs (Fig. 2, compare Middle and Bottom Panels).

3.2. Major differences in translation initiation requirements between A. tumefaciens and E. coli: High GFP translation levels in A. tumefaciens under the control of phage T7 translational enhancer and RBS

Fig. 1B illustrates a summary of a series of constructs with different ribosomal initiation contexts. Construct pC T7 SD-GFP, which contained the phage T7 translational enhancer along with the Shine-Dalgarno sequence (GAAGGAG) and the ribosome binding site (derived from the 5′ non-coding region of the Novagen expression vector, pET29) upstream of the GFP coding sequence, yielded very high levels of GFP protein (Fig. 3, Panel 2) as observed by strong green signals upon confocal microscopy and by Western blot analysis of the expressed protein at *ca*.27kDa (Fig. 4A, lane 1, pCT7SD-GFP). Surprisingly, this construct demonstrated very poor expression in *E. coli* (data not shown) indicating major differences in the translational machinery between these two microorganisms.

On the other hand, a construct containing solely the phage T4 SD sequence GAGG between the CaMV 35S promoter and the ATG of the GFP gene did not express GFP in *A. tumefaciens*, showing that the T4 SD sequence alone was not sufficient for translation

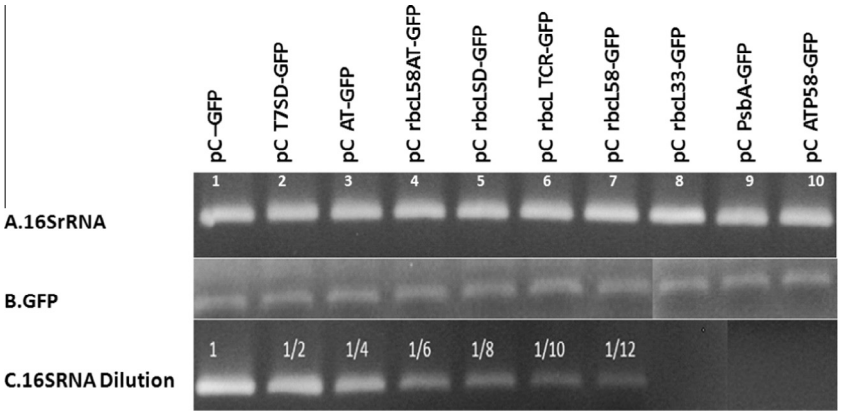

Fig. 2. Quantitation of equivalent GFP transcript levels for all the constructs used in this study. Two percentage agarose-TBE gel analysis of RT-PCR products using primers specific for the 16S rRNA of *Agrobacterium* as well as primers specific for the GFP mRNA (Section 2). Note the relatively higher levels of the cDNA for 16S rRNA (Panel A) compared to that of the GFP mRNA (Panel B); also of note are the equivalent amounts of cDNA for the 16S rRNA in all lanes (Panel A) as well as equivalent amounts of GFP-specific cDNA in all lanes (Panel B), each representing the constructs used in this study. The fractional numbers in Panel C represent the various dilutions of the RT-PCR product for 16S rRNA. Compare the amounts of cDNA in Panel B with those of Panel C: the amounts of GFP cDNA is equivalent with that of the 1/6th dilution of the RT-PCR product for 16S rRNA.

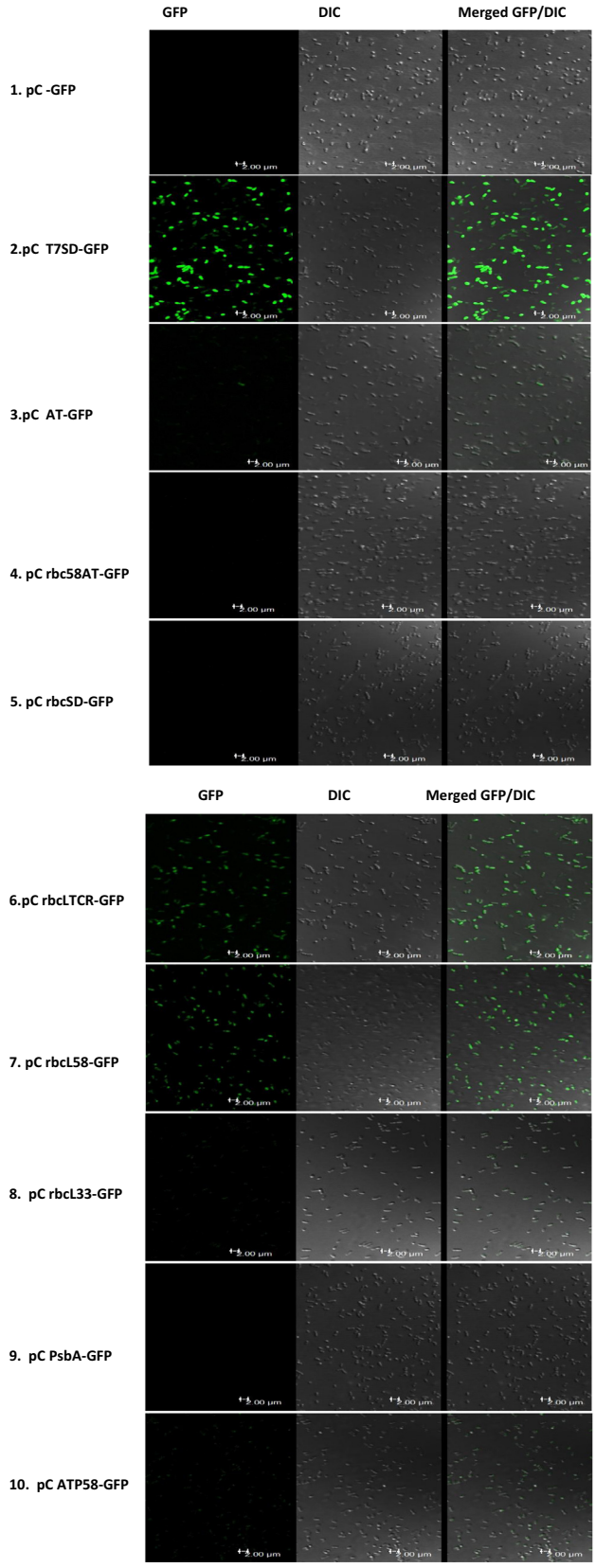

Fig. 3. Confocal microscopy. Detection of green fluorescence due to GFP expression (and translational efficiency) for each of the constructs (Panels 1–10) after transformation into *Agrobacterium* and confocal microscopy. The first image of each panel represents the image with a GFP filter; the middle image that of the DIC filter; and the last image is an overlap of GFP over the DIC picture.

initiation in this organism (Fig. 3, Panel 1 and Fig. 4A, lane 10, pC-GFP), whereas in *E. coli*, the T4 SD sequence alone was sufficient to drive detectable GFP expression [13].

3.3. Effect of the AT-rich sequence from the (AIMV) upstream of the GFP gene on its translation in A. tumefaciens

AIMV CP RNA is one of the most efficiently translated RNAs known [14] and its sequence has been shown to function as a strong translational enhancer [15]. Thirty-three nucleotides containing the 5′ UTR of the AIMV capsid protein gene along with the SD sequence GAGG were cloned upstream of the GFP coding sequence and expressed in *A. tumefaciens*. Data presented in Fig. 3, Panel 3 and Fig. 4A, lane 6, pC AT-GFP, illustrated weak GFP signals as compared to that of the T7 SD construct. This result demonstrated that in *A. tumefaciens*, the T4 SD sequence did not produce enhanced levels of GFP translation even though in combination with the reportedly translationally robust AIMV CP 5′UTR sequence.

3.4. Analysis of 5′ UTR sequences derived from natural chloroplastic genes on translation in A. tumefaciens

Regulation of expression of chloroplastic genes occurs mainly at the level of translation and has several features similar to that of prokaryotes. However, although the SD complementary sequence of the chloroplast 16S rRNA is highly conserved between prokaryotes and plastids [16], the putative SD sequence is poorly conserved in chloroplasts, both in terms of primary sequence and location relative to the start codon [17,18]. Also, plastid gene expression is controlled at the posttranscriptional level by protein factors that are encoded in the nucleus and transported into the chloroplast [19,20], adding a layer of complexity to chloroplast gene expression that is not found in prokaryotes.

In order to compare the above prokaryotic translation initiation sequence context with that of the chloroplast context, and in order to examine the evolutionary closeness of translational regulation between *A. tumefaciens* and chloroplasts as against *E. coli*, we made constructs with 5′ initiation contexts from different chloroplast genes and used them to examine the extent of GFP expression in *A. tumefaciens.*

3.5. Identification of the minimal translation initiation sequence of the rbcL gene required for high level expression in A. tumefaciens

Ribulose1, 5-bisphosphate carboxylase/oxygenase (Rubisco) large subunit (rbcL) is encoded by the chloroplast genome. The 5′-UTR of *rbcL* is highly conserved in the region −1 to −58 and contains a SD sequence (GGAGG) between −4 and −12 [21]. When GFP was cloned downstream of the 5′ translation initiator region of the *rbcL* gene that included the SD-like sequence (GGAGG that is complementary to the CCUCC at the 3′ terminal region of the *Agrobacterium* 16S rRNA), there was no detectable translation of the GFP in *A. tumefaciens* (Fig. 3, Panel 5, Fig. 4A, lane 5, pCrbcSD-GFP), demonstrating that the *rbcL* 5′ translation initiator region (GGGAGGG) by itself is not sufficient for successful translation initiation.

When the 5′ TCR (translation control region containing the 58 nucleotide 5′ leader, the SD-like sequence and the N-terminal coding sequence for the first 14 amino acids) of the *rbcL* gene that was essential for successful translation initiation in chloroplasts [9], was introduced upstream of the ATG of the GFP gene sequence, a robust expression of GFP was detected in *Agrobacterium*, as demonstrated by confocal microscopy (Fig. 3, Panel 6) and by immunoblot analysis (Fig. 4A, lane 2, pC rbcLTCR-GFP).

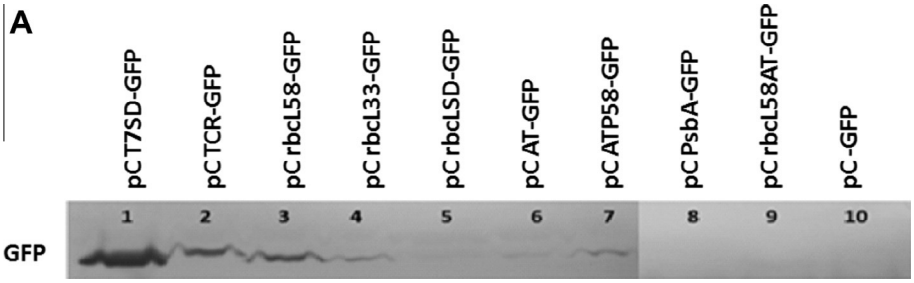

Fig. 4A. Western blotting. Western blots of the enhanced GFP protein (28 kDa) using anti-GFP antiserum and alkaline phosphatase enzyme-linked secondary antibody conjugate. Note the highest level of GFP expression for the pCT7SD-GFP construct (lane 1), followed by that of the pCrbcL TCR-GFP (lane 2) and the pCrbcL58-GFP constructs (lane 3), the latter two in equivalent amounts. The pCrbcL33-GFP (lane 4), pCAT-GFP (lane 6) and pCATP58-GFP (lane 7) constructs exhibit faint bands indicating marginal GFP expression. All other lanes (lanes 5, 8, 9 and 10) are negative for GFP expression.

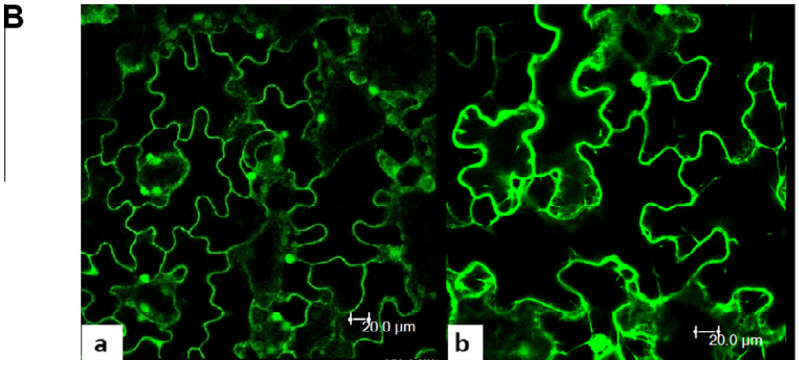

Fig. 4B. Confocal microscopic observation of GFP in *N. tabacum* leaves after 72 h of agro-infiltration with (a) pC rbcL58-GFP and (b) pC-GFP constructs respectively.

Next, we cloned just the 58 nucleotide 5′ UTR of the *rbcL* gene upstream of the GFP gene and transformed it into *A. tumefaciens*. Confocal microscopy (Fig. 3, Panel 7) and Western blot analysis (Fig. 4A, lane 3, pCrbc58-GFP) showed that GFP expression from this construct was equivalent to that of the 5′ TCR. However, it was comparatively less than that observed for the construct containing the highly efficient phage T7 5′ UTR context (Fig. 3, compare Panels 2 and 6, Fig. 4A, compare lane 1, pCT7SD-GFP and lane 2, pCrbcLTCR-GFP). This led us to the conclusion that the 58 nucleotides at the rbcl 5′ UTR was sufficient to initiate efficient translation in *Agrobacteria*. Furthermore, it was observed that these 58 nucleotides serve as translational enhancers when located within 5′-untranslated mRNA leaders (Fig. 4B, a) in plants.

A truncation of the same sequence from 58 to 33 nucleotides from the 5′-terminus resulted in a dramatic reduction of GFP translation (Fig. 3, Panel 8, Fig. 4, lane 4, pCrbcL33-GFP), showing the importance of the 58 base leader sequence for translation in *A. tumefaciens*. In another experiment when SD-like (GGGAGGG) of the 58 base leader sequence was mutated to the TTTATTT sequence, translation was totally abolished (Fig. 3, Panel 4, Fig. 4A, lane 9, pCrbcL58AT-GFP), indicating that the SD-like sequence and context sequence are important for successful translation.

3.6. Comparison of the 5′ UTRs of both rbcL and Psb A genes for translation initiation in A. tumefaciens

Examination of the translational requirements for successful protein expression in *A. tumefaciens* was performed using the *psbA* gene that encodes the D1 protein of photosystem II. A construct that contained an 85 nucleotide 5′ UTR including the RBS, AU-rich region and ATG of the *psbA* gene [21] upstream of the GFP gene was prepared and transformed into *Agrobacterium*. Results showed no detectable GFP expression as judged by confocal microscopy (Fig. 3, Panel 9) and by Western blot analysis (Fig. 4A, lane 8, pC psbA-GFP). This indicated that there is a major difference in the translatability of the GFP protein between the 5′ non-coding sequences of the rbcL gene and the *psbA* gene, even though both are encoded by the chloroplastic genome and are known to be involved in photosynthesis. The *rbcL* gene with the SD-like sequence 10 nucleotides away from the AUG codon (along with its 58 nucleotide 5′ leader) satisfies the requirement for successful translation in *A. tumefaciens*, whereas the *psbA* gene SD-like sequence is much farther away (40 nucleotides upstream) within an unfavorable 5′ sequence context which does not allow for GFP expression.

3.7. 5′ UTR of the chloroplastic atp1 gene supports low GFP translation levels in A. tumefaciens

The atpI gene, which encodes the CFo-IV subunit of the ATP synthase complex [22] is an important chloroplastic gene which possesses an SD-like sequence at an ideal distance: 5 nucleotides upstream of the start codon. When the SD-like sequence along with the 58 nucleotide 5′ translational determinant of the *atpI* gene in chloroplasts [23] was cloned upstream of the GFP coding sequence and expressed in *A. tumefaciens*, a low level of GFP expression was observed (Fig. 3, Panel 10; Fig. 4A, lane 7, pC ATP58-GFP). This result indicates that recognition of the translational context in *A. tumefaciens* is dependent on factors other than just the correctly positioned SD-like sequence, and that the upstream sequence that works in chloroplasts may not work in *A. tumefaciens*. Therefore, of all the chloroplastic constructs used in this study, the *rbcL* 58 with the ideal spacing of the SD-like

sequence from the initiation codon (10 nucleotides) and ideal upstream sequence was the most robust in supporting GFP expression in *A. tumefaciens*.

In the light of the above findings, it would be interesting to examine if there is any other chloroplastic gene besides the *rbcL* gene that can be translated at the same level, if not higher than that of the *rbcL* gene product in *A. tumefaciens*. Results from such further experiments would enable us to make a firm conclusion with regard to both the cis- and trans-acting factors of the Agrobacterial translation machinery. It would also help establish the nature of the evolutionary relationship between *A. tumefaciens* and chloroplasts, as much of the studies in this regard have so far been predominantly performed using *E. coli* as the major Eubacterial organism.

The current study reveals unique translation initiation requirements for high-level protein expression in *A. tumefaciens*. This together with the high strength 35S promoter that shows enhanced transcription levels would enable the design of unique, robust protein expression vectors for *A. tumefaciens* using binary vectors such as pCambia. This system also facilitates transgene design for high-level expression of recombinant proteins using a binary vector in *A. tumefaciens* prior to further downstream applications such as the generation of transgenic plants and plastid-based expression. Thus, preliminarily enhanced translation in *A. tumefaciens* can be used as a predictor of high-level protein synthesis in transgenic plants considering the time-consuming nature of the latter process.

Acknowledgment

We thank both NSERC - Canada and OGS for each partially providing funding for T.A.

References

[1] L. Gold, Post-transcriptional regulatory mechanisms in *Escherichia coli*, Annu. Rev. Biochem. 57 (1988) 199–233.
[2] V. Vimberg, A. Tats, M. Remm, T. Tenson, Translation initiation region sequence preferences in *Escherichia coli*, BMC Mol. Biol. 8 (2007) 100.
[3] J. Shine, L. Dalgarno, The 3′-terminal sequence of *Escherichia coli* 16S ribosomal RNA: complementarity to nonsense triplets and ribosome binding sites, Proc. Natl. Acad. Sci. U.S.A. 71 (1974) 1342–1346.
[4] R.K. Shultzaberger, R.E. Bucheimer, K.E. Rudd, T.D. Schneider, Anatomy of *Escherichia coli* ribosome binding sites, J. Mol. Biol. 313 (2001) 215–228.
[5] J.E. Brock, R.L. Paz, P. Cottle, G.R. Janssen, Naturally occurring adenines within mRNA coding sequences affect ribosome binding and expression in *Escherichia coli*, J. Bacteriol. 189 (2007) 501–510.
[6] H. Chen, M. Bjerknes, R. Kumar, E. Jay, Determination of the optimal aligned spacing between the Shine–Dalgarno sequence and the translation initiation codon of *Escherichia coli* mRNAs, Nucl. Acids Res. 22 (1994) 4953–4957.
[7] P.C. Bonham-Smith, D.P. Bourque, Translation of chloroplast-encoded mRNA: potential initiation and termination signals, Nucl. Acids Res. 17 (1989) 2057–2080.
[8] N.W. Gillham, J.E. Boynton, C.R. Hauser, Translation regulation of gene expression in chloroplasts and mitochondria, Annu. Rev. Genet. 28 (1994) 71–93.
[9] H. Kuroda, P. Maliga, Sequences downstream of the translation initiation codon are important determinants of translation efficiency in chloroplasts, Plant Physiol. 125 (2001) 430–436.
[10] R. Hofgen, L. Willmitzer, Storage of competent cells for Agrobacterium transformation, Nucl. Acids Res. 16 (1988) 9877.
[11] M.G. AbouHaidar, H. Xu, K.L. Hefferon, Plant virology protocols: from in vivo isolation to transgenic resistance, Methods Mol. Biol. 8 (1998) 131–143.
[12] J. Sambrook, E.F. Fritsch, T. Maniatis, Molecular Cloning: A Laboratory Manual, second ed., Cold Spring Harbor Press, New York, USA, 1989.
[13] A. Golshani, J. Xu, V. Kolev, M.G. AbouHaidar, I.G. Ivanov, Inability of *Agrobacterium tumefaciens* ribosomes to translate in vivo mRNAs containing non-Shine–Dalgarno translational initiators, Z. Naturforsch. 57c (2002) 307D312.
[14] M. Kozak, Comparison of initiation of protein synthesis in procaryotes, eucaryotes, and organelles, Microbiol. Rev. 47 (1983) 1–45.
[15] D.R. Gallie, D.E. Sleat, J.W. Watts, P.C. Turner, T.M.A. Wilson, A comparison of eukaryotic viral 5′-leader sequences as enhancers of mRNA expression in vivo, Nucl. Acids Res. 15 (1987) 8693–8711.
[16] B.L. Maidak, J.R. Cole, T.G. Lilburn, C.T. Parker Jr., P.R. Saxman, R.J. Farris, G.M. Garrity, G.J. Olsen, T.M. Schmidt, J.M. Tiedje, The RDP-II (Ribosomal Database Project), Nucl. Acids Res. 29 (2001) 173–174.
[17] T. Hirose, M. Sugiura, Functional Shine–Dalgarno-like sequences for translational initiation of chloroplast mRNAs, Plant Cell Physiol. 45 (2004) 114–117.
[18] M. Sugiura, T. Hirose, M. Sugita, Evolution and mechanism of translation in chloroplasts, Annu. Rev. Genet. 32 (1998) 437–459.
[19] C. Raynaud, C. Loiselay, K. Wostrikoff, R. Kuras, J. Girard-Bascou, F.A. Wollman, Y. Choquet, Evidence for regulatory function of nucleus-encoded factors on mRNA stabilization and translation in the chloroplast, Proc. Natl. Acad. Sci. U.S.A. 104 (2007) 9093–9098.
[20] L. Wobbe, C. Schwarz, J. Nickelsen, O. Kruse, Translational control of photosynthetic gene expression in phototrophic eukaryotes, Physiol. Plant. 133 (2008) 507–515.
[21] J. Kim, J.E. Mullet, Ribosome binding sites on chloroplast rbcL and psbA mRNAs and light-induced initiation of Dl translation, Plant Mol. Biol. 25 (1994) 437–448.
[22] T. Miyagi, S. Kapoor, M. Sugita, M. Sugiura, Transcript analysis of the tobacco plastid operon rps2/atpI/H/F/A reveals the existence of a non-consensus type II (NCII) promoter upstream of the atpI coding sequence, Mol. Gen. Genet. 257 (1998) 299–307.
[23] J.J. Baecker, J.C. Sneddon, M.J. Hollingsworth, Efficient translation in chloroplasts requires element(s) upstream of the putative ribosome binding site from atpI, Amer. J. Bot. 96 (2009) 627–636.