Terms and Conditions of Use of Digitised Theses from Trinity College Library Dublin

Copyright statement

All material supplied by Trinity College Library is protected by copyright (under the Copyright and Related Rights Act, 2000 as amended) and other relevant Intellectual Property Rights. By accessing and using a Digitised Thesis from Trinity College Library you acknowledge that all Intellectual Property Rights in any Works supplied are the sole and exclusive property of the copyright and/or other IPR holder. Specific copyright holders may not be explicitly identified. Use of materials from other sources within a thesis should not be construed as a claim over them.

A non-exclusive, non-transferable licence is hereby granted to those using or reproducing, in whole or in part, the material for valid purposes, providing the copyright owners are acknowledged using the normal conventions. Where specific permission to use material is required, this is identified and such permission must be sought from the copyright holder or agency cited.

Liability statement

By using a Digitised Thesis, I accept that Trinity College Dublin bears no legal responsibility for the accuracy, legality or comprehensiveness of materials contained within the thesis, and that Trinity College Dublin accepts no liability for indirect, consequential, or incidental, damages or losses arising from use of the thesis for whatever reason. Information located in a thesis may be subject to specific use constraints, details of which may not be explicitly described. It is the responsibility of potential and actual users to be aware of such constraints and to abide by them. By making use of material from a digitised thesis, you accept these copyright and disclaimer provisions. Where it is brought to the attention of Trinity College Library that there may be a breach of copyright or other restraint, it is the policy to withdraw or take down access to a thesis while the issue is being resolved.

Access Agreement

By using a Digitised Thesis from Trinity College Library you are bound by the following Terms & Conditions. Please read them carefully.

I have read and I understand the following statement: All material supplied via a Digitised Thesis from Trinity College Library is protected by copyright and other intellectual property rights, and duplication or sale of all or part of any of a thesis is not permitted, except that material may be duplicated by you for your research use or for educational purposes in electronic or print form providing the copyright owners are acknowledged using the normal conventions. You must obtain permission for any other use. Electronic or print copies may not be offered, whether for sale or otherwise to anyone. This copy has been supplied on the understanding that it is copyright material and that no quotation from the thesis may be published without proper acknowledgement.
A LEAF-SPECIFIC PHAGE T7 RNA POLYMERASE-BASED SYSTEM FOR TRANSGENE EXPRESSION IN TOBACCO CHLOROPLASTS.

Alan Magee

A thesis submitted to the University of Dublin for the degree of Doctor of Philosophy

Smurfit Institute of Genetics, University of Dublin, Trinity College, Dublin 2, Ireland. October, 2000
In memory of my parents
Blanche and Michael Magee
DECLARATION

This thesis has not been previously submitted to this or any other University for examination for a higher degree. The work presented here is entirely my own. This thesis may be made available for consultation within the University library. It may be photocopied or loaned to other libraries for the purposes of consultation.

Alan Magee
ACKNOWLEDGEMENTS

Firstly, I would like to thank my supervisor Tony Kavanagh for giving me the opportunity to carry out this interesting work, to attend EuropaBio '99 in Munich and for his help and guidance throughout my PhD.

Thanks to Nga for help inside and outside the lab.

Thanks to Sinead, Paul, Nga, Seamus and Jim for their friendship and help.

A Special thanks to Edele for her support and understanding. Thanks also to all my family especially my sisters Una, Mary and Alice and my uncle Tom. Thanks to my sister Blanche for giving me somewhere to live at the start of my PhD.

Thanks to my friends especially Neal and Adrian.

Thanks to Isabelle for her help at the start of my PhD.

Thanks to Louis and Dave for help.

Thanks to Jullian Hibberd at the University of Cambridge, Downing Street for demonstration of the particle bombardment procedure.
This thesis describes the development of a chloroplast-localised gene expression system based on the phage T7 RNA polymerase. In order to direct T7 RNAP expression in a leaf-specific manner and to target it to the chloroplast, we constructed three chimeric genes (st8, st14 and st25) in which tobacco \textit{rbcS} sequences encoding the Rubisco small subunit (SSU) transit peptide and either the first 8, 14 or 25 amino acids of mature SSU were fused in-frame with the coding sequence of T7 RNAP. Transgenic tobacco plants expressing the ST8, ST14 and ST25 chloroplast-targeted fusion proteins in their leaves were generated using the \textit{Agrobacterium} system for nuclear transformation. These transgenic plants were shown to express a chloroplast-localised transcriptional activity which was specific for the phage T7 DNA template. All transgenic lines grew normally, the transgenes were inherited in a Mendelian fashion in the progeny, and homozygous lines expressing the ST14 and ST25 fusion proteins were identified. We further determined that a large proportion of all three ST fusion proteins was localised within the chloroplast. In addition we determined that the ST25 fusion protein had an \textit{in vivo} half-life of at least 16 days in the dark and that the transcription of eight plastid genes in the ST25 line appeared to be unaltered by the presence of the T7 RNAP activity.

A suite of modular vectors was produced in order to facilitate the efficient introduction of T7 RNAP-dependent transgene expression cassettes into the chloroplast genome. This required the introduction of multiple cloning sites into chloroplast transformation vectors to facilitate the insertion of expression cassettes. In addition, the phage T7 gene 10 (T7G10) expression cassette was modified by the introduction of the plastid \textit{rpsl6} 3' UTR transcript stability element upstream of the T7 transcription termination signals. We introduced multiple cloning sites flanking the modified T7G10 cassette so that chimeric genes based on this cassette could be introduced into the modified chloroplast transformation vectors. We introduced chimeric genes encoding either adult human hemoglobin (HbA), single-chain camel antibodies or thermostable glycosyl hydrolases from \textit{Pyrococcus furiosus} into the modified T7G10 and investigated their expression in \textit{E. coli}.

Transient T7 RNAP-dependent expression of \textit{celB} in leaves of a nuclear transgenic line expressing the ST14 fusion protein was demonstrated using the biolistic transformation procedure. The thermostable \(\beta\)-galactosidase activity encoded by \textit{celB} was detected histochemically and was not detected in leaves of bombarded
wild-type tobacco plants. The biolistic procedure was also used to produce stable chloroplast transformants containing a dicistronic adult human hemoglobin operon in the modified T7G10 expression cassette. These plants grew normally and their progeny showed uniform maternal inheritance of spectinomycin resistance. A sexual cross between the chloroplast transformed line (as the female parent) and a nuclear transgenic line expressing chloroplast-targeted T7 RNAP (as the male parent) produced progeny with a novel yellow-bleached phenotype.
## ABBREVIATIONS

- **bp**: base pairs
- **BAP**: benzylaminopurine
- **BSA**: bovine serum albumin
- **CAT**: chloramphenicol acetyltransferase
- **cAbs**: single chain camel antibodies
- **CaMV**: Cauliflower Mosaic Virus
- **DEPC**: diethyl-pyrocarbonate
- **E. coli**: *Escherichia coli*
- **EDTA**: ethylenediamine tetraacetic acid
- **GFP**: Green Fluorescent Protein
- **GUS**: β-glucuronidase
- **HEPES**: N-[2-Hydroxyethyl]piperazine-N'-[2-ethansulfonic acid]
- **IPTG**: β-D-isopropyl-thiogalactopyranoside
- **kb**: kilobase
- **Kd**: kilodalton
- **kPa**: kiloPascal
- **LSU**: Rubisco large subunit
- **LB**: Luria-Bertani
- **mSSU**: mature Rubisco small subunit
- **µg**: microgram
- **µl**: microlitre
- **MAP**: methionine amino peptidase
- **mg**: milligram
- **ml**: millilitre
- **mM**: millimolar
- **M**: molar
- **MS**: Murashige and Skoog medium
- **MU**: methylumbelliferone
- **MUG**: 4-methyl umbelliferyl glucuronide
- **NAA**: α-naphthalene acetic acid
- **NADH**: β-nicotinamide adenine dinucleotide, reduced form
- **NBX**: MS medium with hormones and vitamins
- **nt**: nucleotide
- **NEP**: nucleus encoded plastid RNAP
- **orf**: open reading frame
- **PAGE**: polyacrylamide gel electrophoresis
- **PEP**: plastid encoded plastid RNAP
- **RbcS**: Rubisco small subunit
- **RNAP**: RNA polymerase
- **rpm**: revolutions per minute
- **Rubisco**: Ribulose-1, 5-bisphosphate carboxylase
- **SDS**: sodium dodecyl sulfate
- **SSU**: Rubisco small subunit
- **SSU-T7**: SSU-T7 RNAP gene fusion
- **ST**: SSU-T7 RNAP fusion protein
- **TIS**: transcription initiation site
- **tsp**: total soluble protein
- **UTR**: untranslated region
- **X-Gal**: 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside
- **X-Gluc**: 5-bromo-4-chloro-3-indolyl β-D-glucuronide
# TABLE OF CONTENTS

## CHAPTER 1
### GENERAL INTRODUCTION

<table>
<thead>
<tr>
<th>Section</th>
<th>Page No.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1. GENE EXPRESSION IN CHLOROPLASTS</td>
<td>1</td>
</tr>
<tr>
<td>1.1.1. Structure and function of the chloroplast genome</td>
<td>1</td>
</tr>
<tr>
<td>1.1.2. Transcription in higher plant chloroplasts</td>
<td>3</td>
</tr>
<tr>
<td>1.1.2.1. Chloroplast RNA polymerase activities</td>
<td>3</td>
</tr>
<tr>
<td>1.1.2.2. Chloroplast promoters</td>
<td>5</td>
</tr>
<tr>
<td>1.1.2.3. The role of PEP and NEP</td>
<td>8</td>
</tr>
<tr>
<td>1.1.2.4. Transcriptional regulation of plastid gene expression</td>
<td>9</td>
</tr>
<tr>
<td>1.1.3. Post-transcriptional control of plastid gene expression</td>
<td>10</td>
</tr>
<tr>
<td>1.1.3.1. Plastid mRNA stability</td>
<td>10</td>
</tr>
<tr>
<td>1.1.3.2. RNA processing and editing</td>
<td>11</td>
</tr>
<tr>
<td>1.1.4. Chloroplast translation</td>
<td>12</td>
</tr>
<tr>
<td>1.1.4.1. Components of the translational apparatus</td>
<td>12</td>
</tr>
<tr>
<td>1.1.4.2. Light-dependent regulation of translation in chloroplasts</td>
<td>13</td>
</tr>
<tr>
<td>1.1.4.3. The role of the Shine-dalgarno sequence in translational initiation</td>
<td>15</td>
</tr>
<tr>
<td>1.2. PLASTID TRANSFORMATION IN HIGHER PLANTS</td>
<td>17</td>
</tr>
<tr>
<td>1.2.1. Transformation procedures</td>
<td>17</td>
</tr>
<tr>
<td>1.2.2. Selectable markers</td>
<td>18</td>
</tr>
<tr>
<td>1.2.3. Transformation vectors</td>
<td>20</td>
</tr>
<tr>
<td>1.2.4. Transformation efficiency</td>
<td>20</td>
</tr>
<tr>
<td>1.2.5. Transformation in species other than tobacco</td>
<td>22</td>
</tr>
<tr>
<td>1.2.6. Genetic improvement of crop plants</td>
<td>22</td>
</tr>
<tr>
<td>1.2.6.1. Advantages of plastid transformation over nuclear transformation</td>
<td>22</td>
</tr>
<tr>
<td>1.2.6.2. Feasibility of producing recombinant proteins in chloroplasts</td>
<td>23</td>
</tr>
<tr>
<td>1.3. BACTERIOPHAGE T7 RNA POLYMERASE BASED EXPRESSION SYSTEMS</td>
<td>24</td>
</tr>
</tbody>
</table>
CHAPTER 2

PRODUCTION OF TRANSGENIC TOBACCO PLANTS EXPRESSING GENE FUSIONS ENCODING A CHLOROPLAST-TARGETED T7 RNA POLYMERASE

2.1. INTRODUCTION

2.2. MATERIALS AND METHODS

2.2.1. Bacterial strains and plasmids

2.2.2. Cloning Enzymes and Molecular cloning techniques

2.2.3. E.coli growth media

2.2.4. Transformation of E.coli

2.2.5. Purification of plasmid DNA (boiling method)

2.2.6. Oligonucleotides and sequencing

2.2.7. Triparental mating

2.2.8. In vitro plant material

2.2.9. Transformation of tobacco

2.2.10. Preparation of tobacco protein extracts

2.2.11. Preparation of Percoll gradients

2.2.12. Preparation of intact chloroplasts

2.2.13. Thermolysin treatment of intact chloroplasts

2.2.14. Determination of protein and chlorophyll concentrations

2.2.15. Western blotting

2.2.16. Densitometry of protein gels and western blots

2.2.17. T7 RNAP transcription assays

2.2.18. Fluorometric GUS assay

2.2.19. Northern analysis

2.3. RESULTS

2.3.1. Construction of chimeric genes encoding ST fusion proteins

2.3.1.1. Construction of pBinTS25 and pBinTGS25

2.3.1.2. Construction of pBinTGS8 and pBinTGS14

2.3.2. Analysis of transgenic tobacco plants expressing ST fusion proteins
2.3.2.1. Generation of transgenic tobacco plants
2.3.2.2. Western blot detection of ST fusion protein expression
2.3.2.3. T7 RNAP activity determinations in transgenic tobacco plants
2.3.2.3. Determination of GUS activity in transgenic lines
2.3.2.4. Tissue specificity of T7 RNAP expression
2.3.2.5. Stability of the ST fusion protein in tobacco
2.3.2.6. Inheritance of transgenes
   2.3.2.6.1. Inheritance of the st genes in the F\textsubscript{1} progeny of selfed primary transformants
   2.3.2.6.2. Inheritance of kanamycin resistance in the F\textsubscript{1} progeny of selfed primary (F\textsubscript{0}) transformants
   2.3.2.6.3. Inheritance of GUS activity in the seed progeny of selfed primary (F\textsubscript{0}) transformants
2.3.2.7. Sub-cellular location of ST fusion proteins
   2.3.2.7.1. Immunodetection of ST fusion proteins in untreated and thermolysin treated gradient-purified intact chloroplasts
   2.3.2.7.2. Detection of T7 RNAP activity in untreated and in thermolysin-treated gradient-purified intact chloroplasts
   2.3.2.7.3. Chlorophyll content of total leaf protein and intact chloroplast preparations
   2.3.2.7.4. Rubisco LSU and ST fusion protein levels in intact and broken chloroplasts
   2.3.2.8. A comparative northern blot analysis of chloroplast transcription in non-transgenic and transgenic tobacco expressing ST fusion proteins.
   2.3.2.8.1. Identification of T7 promoter-like sequences in the plastid genome of tobacco.

2.4. DISCUSSION

2.4.1. ST fusion protein expression and activity in transgenic tobacco
2.4.2. Leaf tissue specificity of the SSU promoter
2.4.3. Stability of ST fusion protein in transgenic tobacco
2.4.3. Transgene inheritance pattern and identification of homozygous lines
2.4.4. Sub-cellular location of ST fusion proteins
2.4.4.1. Detection of ST fusion protein and T7 RNAP activity in untreated and thermolysin-treated intact chloroplasts

2.4.4.2. Efficiency of targeting of ST fusion proteins into chloroplasts

2.4.5. Northern analysis of transcription in transgenic tobacco
CHAPTER 3

MODIFICATION OF PLASTID TRANSFORMATION VECTORS AND CONSTRUCTION OF CHIMERIC GENE EXPRESSION CASSETTES

3.1. INTRODUCTION

3.2. MATERIALS AND METHODS

3.2.1. Bacterial strains and plasmids

3.2.2. Antibodies

3.2.3. PCR and sequencing

3.2.4. Oligonucleotides used for DNA sequencing and PCR

3.2.5. Oligonucleotides used to make synthetic polylinkers

3.2.6. Expression of chimeric genes in E.coli

3.2.7. Western blotting

3.3. RESULTS

3.3.1. Structural analysis of pZS197

3.3.2. Modification of pZS197 to give pCP197

3.3.2.1. Construction of pCP197

3.3.2.2. Modification of pSSH-B1 to give pSSH-B1a

3.3.3. Construction of modified T7 gene 10 expression cassettes for directing T7 RNAP-mediated gene expression in chloroplasts (pETCP5 and pETCP6)

3.3.4. Construction and cloning of chimeric genes into chloroplast transformation vectors

3.3.4.1. Introduction of genes encoding camel single chain antibodies into the chloroplast transformation vectors pCP197 and pSSH-B1a

3.3.4.2. Introduction of genes encoding human α-globin and β-globin into the chloroplast transformation vectors pCP197 and pSSH-B1a

3.3.4.3. Introduction of genes encoding a thermostable β-glucosidase and a β-glucanase into the chloroplast transformation vector pSSH-B1a
3.3.5. Expression of chimeric genes in *E. coli*  
3.3.5.1. Expression of camel antibody genes in *E. coli*  
3.3.5.2. Expression of hemoglobin in *E. coli*  
3.3.5.3. Expression of lamA and celB in *E. coli*  
3.3.5.4. T7 RNAP-dependant expression of GFP in *E. coli*  

3.4. DISCUSSION  
3.4.1. Construction of versatile derivatives of chloroplast transformation vectors  
3.4.2. Modification of the T7 G10 expression casette  
3.4.3. Expression of chimeric genes in *E. coli*
CHAPTER 4

TRANSIENT AND STABLE TRANSFORMATION OF TOBACCO CHLOROPLASTS.

4.1. INTRODUCTION

4.2. MATERIALS AND METHODS
   4.2.1. Plasmids
   4.2.2. Oligonucleotides used in the PCR analysis of Nt-197Hb
   4.2.3. Preparation of tungsten particles for biolistic transformation
   4.2.4. Coating washed tungsten particles with DNA
   4.2.5. Biolistic transformation of tobacco leaf tissue
   4.2.6. GUS and celB histochemical assays
   4.2.7. Regeneration of stable transformants

4.3 RESULTS
   4.3.1. Transient expression of GUS in tobacco leaves
   4.3.2. T7 RNAP-dependant transient expression of celB in tobacco leaves
   4.3.3. Stable transformation of tobacco chloroplast DNA with a dicistronic human hemoglobin operon
   4.3.4. Inheritance of spectinomycin resistance in the C3 line
   4.3.5. A cross between the C3 transplastomic line and a nuclear transgenic line expressing the ST14 fusion protein produces a novel yellow-bleached phenotype

4.4. DISCUSSION
   4.4.1. Detection of T7 RNAP-dependent celB expression in transiently transformed leaf cells
   4.4.2. Stable transformation of tobacco chloroplast DNA with the genes for adult human hemoglobin
GENERAL CONCLUSIONS AND FUTURE PROSPECTS

REFERENCES
CHAPTER 1

GENERAL INTRODUCTION

1.1. GENE EXPRESSION IN CHLOROPLASTS

1.1.1. Structure and function of the chloroplast genome

Chloroplasts develop from small, undifferentiated plastids called proplastids during leaf mesophyll cell differentiation. The primary role of the chloroplast is to supply the cell with fixed carbon and energy as a result of photosynthetic carbon assimilation. In addition, important steps in lipid biosynthesis and amino acid metabolism also occur in chloroplasts (Galili, 1995; Ohlrogge and Browse, 1995). Higher plant cells may contain up to 500 chloroplasts averaging 5-9 μm in length and each containing up to 100 copies of a circular double-stranded DNA genome. The complete nucleotide sequence of several higher plant plastid genomes (plastomes) has been determined including tobacco (Shinozaki et al., 1986), Arabidopsis (Sato et al., 1999), spinach (R. Mache, personal communication), rice (Hiratsuka et al., 1989) and pine (Wakasugi et al., 1994).

In general, higher plant plastomes range in size from 120-160 Kb and gene content, sequence and genomic organization are highly conserved. One of the most striking features of plastome organization is the presence of a pair of large inverted repeats (IRs) that divide the genome asymmetrically into a large single copy (LSC) of 80-100 kb and small single copy (SSC) region of 15-29 kb. The IRs, ranging in size from 5 to 76 kb, account for most of the size variation in plastid genomes and duplicate the rRNA operon and several other plastid genes. A notable exception to this type of organization is the plastid genome of pea and several other legume lineages in which the region corresponding to the IR is single copy.

The plastid genome of higher plants contains about 100 genes that are generally organised into polycistronically transcribed operons which contain functionally related genes. In general, the plastome encodes two functional classes of gene: (1) photosynthesis genes which encode components of the photosynthetic apparatus and (2) genetic system genes which encode components involved in transcription and translation. The tobacco plastid genome contains 40 photosynthesis genes. These encode 5 photosystem I subunits, 13 photosystem II subunits, 4 subunits of the electron transfer chain, 6 ATP synthase subunits, 11
NADH dehydrogenase subunits and 1 gene (rbcL) encoding the large subunit of ribulose bisphosphate carboxylase-oxygenase (Rubisco). The genetic system genes encode 4 rRNAs, 30 tRNAs, 21 ribosomal proteins, 4 subunits of RNA polymerase and 1 intron maturase. The tobacco plastome also contains clpP and accD genes which encode products that function in proteolysis and lipid biosynthesis, respectively, and several conserved orfs whose function is still unknown (Sugiura, 1992).

The most widely promoted scenario for the evolutionary origin of organelles is the endosymbiont hypothesis which proposes that mitochondria and chloroplasts originated from bacteria-like progenitors that were incorporated into a nucleus-containing host cell. Many lines of evidence, based on a comparison of the prokaryotic features of organelles and their genomes with those of extant eubacterial and cyanobacterial lineages support the hypothesis that mitochondria have evolved from the former while chloroplasts evolved from the latter. Indeed, on the basis that almost all chloroplast rRNA gene (rrn) clusters so far reported have the same basic structure as that found in Synechococcus 6301 it has been suggested that the ancestor of the chloroplast was related to Synchococcus 6301 (Kaneko et al., 1996). In order to reconcile the vast difference in gene content between the genomes of present-day chloroplasts and their cyanobacterial ancestor, the hypothesis further proposes that endosymbiosis resulted in loss of autonomy of the endosymbiont which in turn led to relocation of most of its genetic material to the nucleus of the host cell. Thus the vast majority of proteins in chloroplasts are encoded by nuclear genes and following translation in the cytoplasm, are targeted into the chloroplast via N-terminal transit peptide targeting signals.

Recently, attempts have been made to estimate the size of the chloroplast proteome based on comparative analyses of the protein coding capacity of the completely (or almost completely) sequenced genomes of yeast, Synechocystis and Arabidopsis (Abdallah et al., 2000), the identification of genes in the latter which encode proteins possessing a predicted chloroplast transit peptide and the identification of Arabidopsis homologues of Synechocystis genes. These analyses predicted that the chloroplast proteome of Arabidopsis comprises between 1900 and 2500 proteins of which 650-900 are encoded by genes that were originally transferred from the genome of the cyanobacterial endosymbiont to the nucleus.

It is not known with certainty what determined which genes of the original endosymbiont were retained in the plastid genome of present-day plants.
and which were relocated to the nucleus. However, it is interesting to note that in general the same set of genes are encoded by extant plastomes in many different lineages which suggests that selection rather than chance has dictated what genes were retained in the plastid. Also, by comparing plastomes it is apparent that those photosynthesis proteins whose genes are most resistant to transfer to the nucleus are located at or close to the functional core of the photosynthetic reaction centres (Race et al., 1999). Allen et al. (1993) have proposed that the source of this selection is the requirement for rapid regulation of expression of the core genes of the photosystems which requires that these genes be present in the plastid. The rationale behind this proposal is that the synthesis of components of the photosystem reaction centres must be capable of responding rapidly to changes in redox potential in order to limit the production of reactive oxygen species which damage the photosynthetic membrane. In addition, the genetic system genes are also retained so that this rapid regulation can be effected through rapid changes in transcription and translation.

1.1.2. Transcription in higher plant chloroplasts

1.1.2.1. Chloroplast RNA polymerase activities

It is now accepted that higher plant plastids contain at least two distinct RNA polymerase (RNAP) activities; (1) the plastid encoded plastid RNAP, PEP, which is primarily encoded by the plastome and (2) the nucleus encoded plastid RNAP, NEP, which is encoded in the nucleus and imported into plastids. The core subunits of PEP are encoded by the plastid rpoA, rpoB and rpoC1/C2 genes which encode subunits homologous to the α, β and β′ subunits of the E.coli RNAP. The cloning of Arabidopsis genes encoding plastid targeted σ-like factors (Isono et al., 1997; Tanaka et al., 1997) and the identification of σ-like factors in mustard chloroplasts (Tiller et al., 1991) indicates that the prokaryotic mechanism for promoter selection has been retained in higher plant plastids. In chloroplasts PEP is the dominant RNAP and is responsible for the majority of photosynthesis and genetic system gene transcription. The subunit composition of PEP changes at different developmental stages and in mustard chloroplasts PEP consists of upto 13 different polypeptides which include an RNA binding protein and proteins that enable the RNAP to function effectively in an environment containing the free-radicals generated as photosynthetic by-products (Pfannschmidt et al., 2000).

For many years there was speculation that plastids contained a second nuclear encoded plastid RNAP (NEP) activity. The evidence for this was based
on the biochemical isolation of a separate 110Kd single-subunit RNAP activity from spinach chloroplasts (Lerbs-Mache et al., 1993) and the observation that transcriptional activity was present in plastids from plants that do not contain the genes for the PEP (Morden et al., 1991) or are unable to synthesis PEP (Hess et al., 1993; Han et al., 1993). Allison and Maliga (1996) provided the first direct evidence for the existence of NEP by demonstrating that some plastid genes were still transcribed in transplastomic tobacco plants that had the PEP rpoB gene deleted (ΔrpoB plants). The ΔrpoB plants lacked three of the four core subunits of PEP because transcription of the rpoC1 and rpoC2 genes downstream of rpoB was also abolished. Serino and Maliga (1998) generated ΔrpoA tobacco plants that had the same transcriptional pattern as the ΔrpoB plants and thus demonstrated that none of the PEP core subunits formed part of NEP. The cloning of an Arabidopsis gene that encodes a chloroplast targeted protein that is 55% homologous to the mitochondrial RNAP (Hedtke et al., 1997) and the identification of a promoter motif present in both mitochondrial and NEP promoters (Liere and Maliga, 1999) is further evidence for the existence of NEP and suggests that it is a phage-type RNAP related to the mitochondrial RNAP. In addition, a maize nuclear gene which encodes a plastid-targeted T7 RNAP-like protein has recently been cloned and polymerase assays indicated that the protein functions as an RNAP (Chang et al., 1999).

Bligny et al. (2000) separated three different transcriptional activities from spinach chloroplasts and presented evidence that the plastid contains a second nucleus encoded polymerase which they called NEP-2. This enzyme was shown to be distinct from both PEP and the phage-type NEP (now designated NEP-1) on the basis of the following criteria: (i) in vitro transcription assays revealed that the NEP-2 activity did not recognise the rbcL PEP promoter but initiated transcription from the rrn non-PEP type PC promoter. It was further shown that the PC promoter was transcribed by a NEP activity because it was transcribed in spectinomycin grown spinach which lack PEP activity; (ii) NEP-2 could not be inhibited by tagetin which is a potent inhibitor of transcriptional elongation by PEP; (iii) NEP-2 could be distinguished from NEP-1 because antibodies raised against NEP-1 did not react with NEP-2-containing extracts; (iv) NEP-2 was found to produce longer transcripts on a T7 promoter template than purified NEP-1 extracts; (v) point mutations in the region of the PC promoter affected the NEP-1 and NEP-2 activities differently. Because of the distinctiveness of NEP-2 from the phage-type and prokaryotic-type RNAPs and the observation that a eukaryotic-like promoter element (TATA) is found in some plastid promoters (Link, 1984; Eisermann et al., 1990), Bligny et al. (2000)
speculated that NEP-2 may represent a eukaryotic-type RNAP that was co-opted for plastid transcription during the evolutionary integration of plastids into eukaryotic cells.

### 1.1.2.2. Chloroplast promoters

The prokaryotic nature of the PEP transcriptional apparatus made it possible to easily identify many PEP promoters on the basis of their similarity to the -10/-35 σ70-type promoters that are commonly found in *E.coli*. As in *E.coli* the consensus PEP promoter consists of two hexanucleotide sequences, -10 (TATAAT) and -35 (TTGACA), upstream of the transcriptional initiation site (TIS), separated by 16-18 nucleotides. *In vitro* transcription studies determined that these elements were responsible for transcription not only in *E. coli* but also in chloroplasts (Bradley and Gatenby, 1985). Some PEP promoters contain other elements, in addition to the -10 and -35 elements, that function in transcriptional modulation.

Two interesting PEP promoters, the *psbD* blue light-responsive promoter (BLRP) and the *psbA* promoter contain sequences other than the -10 and -35 elements which are important for transcription. Both of these genes encode photo-labile components of the photosystem II reaction centre and hence high levels of *psbD* and *psbA* expression is required in response to light. Sequence elements of the BLRP that are important for transcription have been identified by a combination of *in vivo* and *in vitro* studies. The BLRP from tobacco has been characterised in transformed chloroplasts by deletion analysis (Allison *et al.*, 1995) and the barley BLRP has been characterised in *in vitro* transcription assays (Kim *et al.*, 1999a). These studies have shown that the prokaryotic -10 element but not the -35 element is required for barley BLRP transcription *in vitro* (Kim *et al.*, 1999a) and that a conserved region approximately 20 bp long located immediately upstream of the -35 element is required for light responsive transcription from the tobacco BLRP *in vivo* (Allison *et al.*, 1995).

This conserved region known as the AGG-box and its cognate DNA binding protein complex, AGF, have been shown to be required for transcription from the barley BLRP *in vitro* (Kim and Mullet, 1995). An additional conserved region located further upstream of the AGG-box in BLRPs which is known as the PGT-box because of its similarity to the GT-motif found in light-regulated nuclear genes, has been shown to contribute to the strength of the tobacco *psbD* BLRP *in vivo* (Allison *et al.*, 1995). Furthermore, in barley, a protein complex
(designated PGTF complex) has been shown to specifically bind to the PGT element in a manner that is dependent on phosphorylation (Kim et al., 1999b).

The psbA promoter has also been characterised using deleted and mutated promoter elements in in vitro transcription assays (Satoh et al., 1999; Kim et al., 1999a). These studies found that unlike the BLRP, sequences important for psbA transcription are located within the -10/-35 core region. Satoh et al. (1999) found that PEP extracted from developing chloroplasts in the basal portion of wheat leaves required both the -35 and -10 elements for transcription. On the other hand Kim et al., (1999a) found using PEP extracts from whole barley seedlings that a TATA-box-like sequence between -10 and -35 and the -10 element itself but not the -35 element were required for transcription. However, Satoh et al., (1999) found that PEP extracted from mature chloroplasts isolated from wheat leaf tips did not require the -35 element nor the TATA-box-like sequence and that an extended version of the -10 element (TGT TATA CT, the -10 region is italicised) was sufficient for transcription. Similar promoters containing the extra TGT sequence that do not require -35 regions for initiation of transcription are found in E.coli (Kumar et al., 1993).

NEP promoters were identified by the mapping of transcripts initiating from non-σ70-type promoters that were active in the non-photosynthetic plastids in the tobacco BY2 cell line (Vera and Sugiura, 1995; Vera et al., 1996), tobacco mutants containing deleted PEP rpo genes (Allison et al., 1996; Hajdukiewicz et al., 1997; Serino and Maliga, 1998), tobacco seedlings grown in the presence of a PEP inhibitor (Kapoor et al., 1997) and the barley albostrians and maize iojap mutants that lack plastid ribosomes (Hubschmann and Borner, 1998; Silhavy and Maliga, 1998). Transcripts from these non-consensus promoters were usually undetectable in chloroplasts because many of these transcription units were preferentially transcribed from alternative PEP promoters and the level of transcription from these promoters was found to increase in the absence of PEP. Genuine transcription initiation sites (TISs) for NEP promoters were mapped by identifying the 5' nucleotide in transcripts by primer extension analysis and primary transcripts were distinguished from transcripts arising from RNA processing by their ability to be capped by guanylyltransferase in vitro.

Sequence alignments in the region of the TISs determined that in most cases NEP promoters in both monocots and dicots consisted of two blocks of conserved sequence that were separated by 10-20 nucleotides. The larger block (box I) consisted of approximately 15 nucleotides and was located mostly
upstream of the TIS in a manner reminiscent of phage RNAP promoters. The smaller sequence block (box II) was approximately 6 nucleotides long and was located upstream of box I at around position -35 relative to the TIS. All NEP promoters identified to date contain sequences homologous to box I except for \( P_{clpP-53} \) (\( clpP \) promoter initiating 53 nucleotides upstream of the coding sequence) in tobacco which contains no homology to either box I or box II sequence elements (Hajdukiewicz et al., 1997). Most contain both box I and box II consensus elements except for \( P_{accD-129} \), \( P_{clpP-511} \) and \( PrpoB-345 \) in tobacco (Hajdukiewicz et al., 1997; Serine and Maliga, 1998) and \( P_{clpP-111} \) and \( PrpoB-147 \) in maize (Silhavy and Mailga, 1998) which only contain box I consensus elements.

Liere and Maliga (1999) developed a NEP \textit{in vitro} transcription system using plastid extracts from \( \Delta rpo \) tobacco plants and functionally characterised the NEP promoter that exclusively transcribes the tobacco \( rpoB \) operon (\( Nt-PrpoB \)). They determined by deletion analysis that the minimum \( Nt-PrpoB \) consists of a 15-nucleotide segment extending from -14 to +1 relative to the TIS which confirmed that box I sequence elements identified by sequence alignments were functionally significant. The activity of \( Nt-PrpoB \) containing every possible single substitution in the 15 nucleotide region was determined in the \textit{in vitro} transcription assays. Point mutations at every position reduced transcription relative to wild-type \( Nt-PrpoB \) except for the -5 position which was neutral to substitution. A critical core promoter motif, CRT, between -8 and -6 that was the least tolerant to substitutions was identified. In all NEP promoter box I elements identified to date a YATA motif is conserved in the CRT region and is situated 3 to 7 nucleotides upstream of the TIS. The YATA motif is homologous to the functionally important YRTA motif found in mitochondrial promoters which supports the evidence that the NEP gene arose from the duplication of the mitochondrial RNAP gene (Hedtke et al., 1997). Surprisingly, the A at position -5 immediately downstream of CRT in \( Nt-PrpoB \) which is conserved in most NEP promoter YATA motifs was found to be neutral to substitutions in this study. Also, it is surprising that a T to G transversion at the TIS did not reduce promoter activity given that all NEP promoters identified to date (except \( Nt-P_{clpP-53} \) which has G as the TIS) have an A or T as the TIS.

Kapoor and Sugiura (1999) developed a similar \textit{in vitro} transcription system from the non-photosynthetic plastids of cultured tobacco BY2 cells. Using deletion analysis they determined that the \( Nt-P_{atpB-290} \) NEP promoter was contained in a 43 nucleotide region extending from -35 to +8 which included both
the box I and box II sequence elements that had been identified for Nt-PatpB-290 by sequence alignment (Kapoor et al., 1997). In vitro transcription assays with Nt-PatpB-290 containing multiple substitutions in the 43 nucleotide region confirmed that both the box I and box II elements were important for promoter strength and that substitutions in the important YATA motif identified by Liere and Maliga (1999) caused the greatest reduction in promoter activity.

Sriraman et al. (1998) characterised Nt-PclpP-53, the non-consensus constitutive NEP promoter (Hajdukiewicz et al., 1997), by generating transplastomic tobacco plants containing a series of Nt-PclpP-53 deletion derivatives fused to a reporter gene. They determined that Nt-PclpP-53 sequences from -5 to +25 were sufficient to support specific transcription initiation and that the corresponding region was conserved and functioned as a promoter in liverworts and conifers. However, the corresponding region in rice, maize and barley although conserved, does not function as a promoter and yet this region from rice was capable of initiating transcription in tobacco chloroplasts. The evidence suggests that the Nt-PclpP-53-type NEP promoter which is unlike any previously described NEP or phage RNAP promoter, requires an activating factor for transcription by NEP and that this factor is absent in monocots. Most NEP promoters are inactive in chloroplasts (Allison et al., 1996; Hajdukiewicz et al., 1997) and it may be the case that NEP promoters like Nt-PclpP-53 that are highly active in chloroplasts require activating factors. Another example is the CDF2 factor from spinach which probably activates NEP-mediated transcription of the 16S rRNA (rrn) operon in spinach chloroplasts (Iratini et al., 1997).

1.1.2.3. The role of PEP and NEP
In general, PEP promoters are found upstream of photosynthesis genes and gene clusters while most house keeping genes have both NEP and PEP promoters. Some transcription units like accD and the rpo operon are exclusively transcribed from NEP promoters. Given that genetic system genes are preferentially transcribed relative to photosynthetic genes early in chloroplast development (Rapp et al., 1992; DuBell and Mullet, 1995) it has been suggested that NEP and PEP act sequentially in a developmental cascade with PEP taking over from NEP in chloroplasts. The fact that the rpoB operon is exclusively transcribed by NEP in all plants analysed to date supports the cascade model. However, the identification of the PclpP-111 NEP promoter in maize (Silhavy and Maliga, 1999) that exclusively drives clpP transcription and is active in chloroplasts in vivo suggests that both polymerases are present in chloroplasts and that transcription is controlled by promoter-specific transcription factors.
1.1.2.4. Transcriptional regulation of plastid gene expression

Although the importance of post-transcriptional processes in the regulation of plastid gene expression is widely recognised (Sugita and Sugiura, 1996) recently it has become apparent that plastid transcription plays a major role in plastid gene expression. Depending on the light conditions and the developmental stage of the plant, plastid transcription rates have been found to vary by as much as 300 fold for certain genes and these rates are in general predictive of mRNA and protein levels (Baumgartner et al., 1993; Christopher et al., 1992; Rapp et al., 1992). The recent characterisation of the NEP transcription system and the fact that both PEP and NEP are active in chloroplasts (Sriraman et al., 1998) has also increased the appreciation of the extent of transcriptional regulation in chloroplasts. The identification of plastid transcription factors and complex plastid promoters like the psbD BLRP is further evidence that plastid transcription is extensively regulated in response to environmental and developmental signals.

CDF2 (referred to earlier) is a sequence-specific DNA binding protein which binds between two inactive $\sigma^{70}$-type promoters upstream of the rrn operon in spinach chloroplasts. In vitro transcription assays have shown that CDF2 suppresses PEP transcription from the flanking promoters (Iratini et al., 1994). However, rrn transcription in spinach chloroplasts was mapped to a non-$\sigma^{70}$-type promoter (PC) located between the two PEP promoters and therefore it was proposed that CDF2 specifically activates NEP transcription from PC. The fact that CDF2 and PC transcription are absent in spinach root amyloplasts supports the proposal that CDF2 specifically activates PC transcription in leaves (Iratini et al., 1997). In tobacco chloroplasts CDF2 is not present and rrn transcription is initiated from a PEP promoter (Allison et al., 1996) that corresponds to one of the two repressed promoters in spinach. Thus, the CDF2 transcription factor regulates rrn expression in spinach in a species-specific and an organ-specific manner.

In addition, the AGG and PGT motifs of the psbD BLRP specifically bind the AGF and PGTF protein complexes, respectively, and these factors probably activate transcription of psbD in response to light. It has been shown that the PGTF complex looses its affinity for the BLRP when subjected to ADP-dependent phosphorylation (Kim et al., 1999b). Given that ADP is more prevalent under dark conditions this may be the mechanism by which transcriptional activation of the BLRP is light-specific. Also, the TATA-box-like sequence in the psba promoter that has been identified as being important for
psbA transcription (Kim et al., 1999a) may also bind a unique transcription factor that specifically modulates the transcription of psbA in response to light.

An interesting feature of plastid transcriptional regulation is the switch in PEP promoter specificity from consensus $\sigma^{70}$-type promoters in developing chloroplasts to the exclusive transcription of PEP promoters that are transcriptionally activated by light in mature chloroplasts (Satoh et al., 1999). Run-on transcription assays and in vitro transcription assays were carried out using developing chloroplasts from the basal portion of wheat leaves and from mature chloroplasts in wheat leaf tips. Samples were taken from dark grown seedlings and from dark grown seedlings that had been illuminated for two hours. The activity of five PEP promoters including the psbD BLRP and the psbA promoter was determined. In developing chloroplasts all the promoters were transcribed in both the light and dark except for the psbD BLRP which was light-specific. However, in mature chloroplasts only the psbA promoter and the psbD BLRP were significantly transcribed under light conditions and no PEP transcription was detected in the dark. Thus, the PEP had switched from recognising $\sigma^{70}$-type promoters irrespective of light conditions to only recognising promoters that contain elements in addition to -10 and -35 that probably bind transcriptional activating factors under light conditions. These findings demonstrate that transcriptional regulation plays a major role in turning off the expression of most photosynthesis genes when the chloroplast has matured.

1.1.3. Post-transcriptional control of plastid gene expression

1.1.3.1. Plastid mRNA stability

mRNA stability plays an important role in the regulation of plastid gene expression by determining the availability of transcripts for translation. Initially, it was thought that the inverted repeat (IR) sequences found in the 3'-UTRs of most monocistronic and polycistronic mRNAs functioned as transcriptional terminators like their prokaryotic counterparts. However, IRs have been shown in tobacco plastid in vitro transcriptional systems (Stern and Gruissem, 1987) and in vivo in transplastomic Chlamydomonas (Rott et al., 1996) to be inefficient transcriptional terminators. It has now been determined that Chlamydomonas plastid transcriptional units are transcribed as longer precursors and that a precise endonucleolytic cleavage followed by processive 3' to 5' exonucleolytic digestion generates the mature transcript (Stern and Kindle, 1993). It is thought that the secondary stem-loop structure formed by the IR sequence protects the mRNA from further 3' to 5' exonuclease digestion and it has been demonstrated that
when 3' IR structures are deleted from pre-mRNAs, the mutant RNAs are rapidly degraded in vitro (Stern and Gruissem, 1987) and in vivo (Stern et al., 1991).

It is now evident that both the 5' and 3' UTRs of plastid mRNAs contribute to stability. An analysis of transcription and mRNA steady-state levels in tobacco chloroplasts transformed with the GUS reporter gene fused to deletion derivatives of the rbcL 5'-UTR and the rbcL promoter (Shiina et al. 1998) showed that an IR sequence in the rbcL 5'-UTR compensates for reduced transcription from the rbcL promoter in the dark by increasing transcript stability. Eibl et al. (1999) constructed chimeric reporter gene constructs that all contained the tobacco rrm operon promoter and all possible combinations of 5'- and 3'-UTRs from the tobacco psbA and rbcL genes. By measuring transcript steady-state levels they determined that both the 5' and 3'-UTRs make a contribution to transcript stability and that the 5'-UTR from both genes has the strongest effect.

Plastid mRNA stability is not constant and changes in response to environmental and developmental signals which are mediated by nucleus-encoded factors. Several nucleus encoded RNA-binding proteins (RNPs) that are required for mRNA stability have been identified. A 28Kd RNP has been isolated from spinach chloroplasts that co-purifies with the 3' end processing complex and is phosphorylated in vitro by the spinach casein kinase II which reduces its affinity for the 3'-UTR (Lisitsky and Schuster, 1995). It is possible that phosphorylation modulates 3' end processing and mRNA stability by controlling the affinity of the 28Kd RNP for the 3' UTR. Also, in Chlamydomonas, the mcdl chloroplast mutant has been shown to lack a nucleus-encoded factor that is required for petD mRNA stability by apparently protecting it from a processive 5' to 3' exonucleolytic activity (Drager et al., 1999).

1.1.3.2. RNA processing and editing
mRNA processing and editing play an important role in regulating the expression of at least some plastid genes by determining the efficiency of translation and in some cases the activity of the protein product. For example, a methyl jasmonate-induced change in the length of the 5'-UTR of the barley rbcL transcript disrupted translation (Reinbothe et al., 1993) which suggests that correct 5'-end formation, by appropriate post-transcriptional processing can be important for translation. Also, most chloroplast genes are transcribed as part of a polycistronic pre-mRNA that is processed to form monocistronic transcripts. In some cases, orfs located on the larger precursor transcripts are less efficiently translated in vivo than when present on the corresponding monocistronic transcripts (Barkan et al., 1994) and
in vitro assays have shown that the dicistronic psaC/ndhD transcript was untranslatable (Hirose and Sugiura, 1997). RNA editing on the other hand can also play an important role in translation efficiency. There are several examples in higher plant plastids where RNA editing is required for the creation of an initiation codon (Sugiura et al., 1998). For example, in maize the rpl2 transcript is edited to change an ACG codon to the AUG initiation codon required for translation (Hoch et al., 1991). In addition, RNA editing plays an important role in PEP activity by changing a codon in the rpoA transcript to a conserved codon which encodes an amino acid important for catalytic activity (Hirose et al., 1999).

1.1.4. Chloroplast translation

1.1.4.1. Components of the translational apparatus

The chloroplast translational machinery is encoded in the nuclear and plastid genomes and in many respects resembles the prokaryotic system. The chloroplast ribosome is prokaryotic-like with respect to its sedimentation coefficient and its sensitivity to antibiotics. Chloroplast rRNA is highly homologous to prokaryotic rRNA and many of the plastid ribosomal proteins (RPs) have prokaryotic homologs. Of the 60 or so chloroplast ribosomal proteins (RPs) approximately one third are plastid encoded and all rRNA is plastid encoded. In addition, the tobacco plastome encodes 30 tRNA species which is sufficient to support plastid gene expression if expanded wobble base-pairing occurs (Sugiura et al., 1998). Other prokaryotic-like components include aminoacyl-tRNA synthetases, initiation factors, elongation factors and termination factors. Of these only the IF-1 initiation factor is plastid encoded in higher plant species. There is evidence that some components of the eukaryotic translational apparatus may have been recruited into the plastid translational system. At least five of the nucleus-encoded plastid RPs are not homologous to any known E.coli RPs (Subramanian et al., 1993) and no homologous orfs have been identified in the genome sequences of E.coli or the cyanobacterium, Synechocystis 6803 (Sugiura et al., 1998). Also, a nucleus-encoded RNA binding protein that is homologous to cytoplasmic translation factors has been identified as a translational activator in the Chlamydomonas chloroplast (Yohn et al., 1998) and it is possible that the higher plant translational apparatus contains similar proteins.
1.1.4.2. Light-dependent regulation of translation in chloroplasts

Plastid transcripts are very stable compared to *E.coli* with half lives ranging from 6 hours for *psaA* to over 40 hours for *psbA* (Mullet and Klein, 1987). Thus, it makes sense that the expression of photosynthetic genes with long half-lives like *psbA* is regulated by light so that gene expression levels can respond to short term environmental changes. It has been demonstrated that light promotes translational initiation of *psbA* translation by increasing the abundance of initiation complexes bound to *psbA* mRNA (Kim and Mullet, 1994) and although, there is some *psbA* transcription in the dark and in amyloplasts the transcript is not incorporated into polysomes (Deng and Gruissem, 1988). These findings suggest that *psbA* is most actively synthesised in light grown chloroplasts but not in the dark nor in non-photosynthetic plastids.

Much research has been carried out on light-dependent translation and the evidence suggests that the 5'-UTRs of plastid mRNAs play an important role in this process. The 5'-UTR of the *psbA* mRNA was shown to confer light-induced translation on a reporter gene in transplastomic tobacco plants (Staub and Maliga, 1993). In addition, there is also evidence that 3'-UTRs complement their 5'-UTR counterparts in translational activation because Eibl *et al.* (1999) found in transplastomic plants that a reporter gene construct which contained both the 5'- and 3'-UTRs of the *psbA* mRNA was translated with twice the efficiency of a reporter construct containing the *psbA* 5'-UTR and the *rpl32* 3'-UTR despite the fact that the mRNA steady-state level for the latter transcript was three times that of the former. This is not surprising considering that in the cytoplasm 5'- and 3'-UTRs interact to form a circular mRNA that is required for efficient translation (reviewed in Gallie, 1998). In higher plants few of the transacting factors that mediate light-activated translation have been identified. However, in spinach chloroplasts a homolog of the prokaryotic S1 RP that functions in translation initiation was identified (Alexander *et al.*, 1998) and it was shown that when isolated from light-grown seedlings the spinach S1 protein was capable of binding to the *psbA* 5'-UTR but not when isolated from dark-grown spinach (Klaff and Gruissem, 1995).

In *Chlamydomonas* on the other hand many nuclear-encoded protein factors that are important in light-activated translation have been isolated and some of their genes have been sequenced and cloned. For example, a protein complex that binds to the *psbA* 5'-UTR in *Chlamydomonas* has been identified and the study of this complex has made a large contribution to the understanding of light-dependent translation in plastids. The complex consists of four protein subunits.
and its affinity for the psbA 5'-UTR is greatest in illuminated cells (Danon and Mayfield, 1991). Cross-linking studies have shown that a 47Kd RNA binding protein (RB47) is the protein in the complex that is in direct contact with the psbA 5'-UTR and the nuclear gene encoding this protein has been cloned (Yohn et al., 1998a). Sequence analysis revealed that RB47 was homologous to the poly(A) binding proteins (PABPs), a family of eukaryotic proteins that bind to the poly(A) tails of cytoplasmic mRNAs and function in translational initiation. There is evidence that the chloroplast localised PBAP (cPABP) also functions in translation initiation because a Chlamydomonas nuclear mutant that completely lacked cPABP had a 95% reduction in the amount of psbA mRNA that was associated with ribosomes (Yohn et al., 1998b). The finding that interaction between plastid 5'- and 3'-UTRs contributes to translational efficiency in tobacco (Eilb et al., 1999) suggests that protein homologs of cPABP may also be found in higher plant chloroplasts because PABPs have been identified as components of complexes that circularise cytoplasmic mRNAs (Wells et al., 1998).

Kim and Mayfield (1997) cloned the nuclear gene for a second component in the psbA mRNA binding complex, RB60 and sequence analysis and in vitro import assays determined that it encoded a chloroplast-localised homolog of protein-disulfide isomerase (cPDI). PDI is a multifunctional enzyme that is involved in protein folding and is typically found in the ER. The cPDI sequence contained conserved regions that encode the catalytic sites that regulate the formation, reduction and isomerisation of disulfide bonds in PDIs which suggests that it functions as a transducer of the chloroplast redox potential in a process that alters the binding properties of the entire protein complex. This idea was supported by a previous finding that oxidation prevented the cPABP-mediated binding of the protein complex to psbA mRNA in vitro which could be restored by reducing the complex (Danon and Mayfield, 1994a).

More recently, it has been demonstrated that changing the cysteines in the RNA binding domains of cPABP to serines by site-directed mutagenesis rendered the protein insensitive to redox regulation of its affinity for the psbA 5'-UTR (Fong et al., 2000). Thus, in vivo the cPDI probably catalyses the formation and reduction of disulfide bonds between these cytosines in response to chloroplast redox potential. Also, ADP-dependent phosphorylation of cPDI was shown in vitro to inhibit complex binding to psbA mRNA (Danon and Mayfield, 1994b). It was proposed that in vivo these mechanisms allow for a reversible switch regulating psbA expression because the translation activating complex will have high affinity for the psbA 5'-UTR during daylight hours when
photosynthetic activity is high and consequently levels of ATP relative to ADP are high and the chloroplast compartment is at its most reducing.

1.1.4.3. The role of the Shine-Dalgarno sequence in translational initiation

In *E. coli* almost all mRNAs contain Shine-Dalgarno (SD) sequences (typically GGAGG) that are located 7 +/-2 nucleotides upstream of the initiation codon and the spacing between the SD and initiation codon is critical to guide the 30S subunit of the ribosome to the initiation codon by RNA-RNA base-pairing between 16S rRNA and the SD sequence during initiation of translation (McCarthy and Brimacombe, 1994). *In vitro* binding and translation assays indicate that SD sequences present within 20 nucleotides of the initiation codon are functional in higher plant chloroplasts. Toeprint analyses have established that the 30S ribosomal subunit associates with the SD sequences just upstream of the initiation codons in barley *rbcL* and *psbA* mRNAs (Kim and Mullet, 1994). Mutation of SD-like sequences in the following tobacco chloroplast transcripts has revealed that they are required for *in vitro* translation; GGA at positions -10 to -6 in *rps14* (Hirose et al., 1998), GGAG at -18 to -15 in *atpE* and GGAGG at -10 to -6 in *rbcL* (Sugiura et al., 1998).

Svab and Maliga (1993) developed a chloroplast expression cassette that contained a hybrid 5'-UTR consisting of the first 30 nucleotides of the *rrn* 5'-UTR fused to the last 18 nucleotides (which contains the SD sequence) of the *rbcL* 5'-UTR downstream of the *rrn* promoter. Based on the mapped transcription initiation sites of *rrn* and *rbcL* in tobacco chloroplasts (Allison et al., 1996) this hybrid 5'-UTR is missing 80 nucleotides of the *rrn* 5'-UTR and 160 nucleotides of the *rbcL* 5'-UTR. Therefore, the fact that it is effective in translation *in vivo* (Svab and Maliga 1993; McBride et al., 1995) indicates that the *rbcL* SD sequence in the hybrid 5'-UTR is functional in translation. However, there is also evidence that sequences other than the SD are responsible for at least some translational initiation in higher plant chloroplasts. The evidence for this comes from the observation that 30 of the 79 protein coding genes (including 9 orfs) in the tobacco plastome contain no SD-like sequences within the 20 nucleotides upstream of the initiation codon (Sugiura et al., 1998). Also, it was found that the *psbA* transcript could not bind *E. coli* ribosomes *in vitro* whereas the *rbcL* transcript could (Kim and Mullet, 1994) which suggests that the SD-like sequences in the *psbA* 5'-UTR on their own are not sufficient for translation initiation *in vivo*. Hirose and Sugiura (1995) identified a *psbA* mRNA-specific AU-rich sequence (UAAAUAAA) located -21 to -14 (relative to the initiation
codon) which was required for translation in vitro. Thus, it is possible that this sequence binds a protein factor that is required for psbA translation.

It is also likely that an alternative mechanism to straight forward SD-mediated translational initiation is present in Chlamydomonas because true SD sequences are rarely found near the initiation codon in its chloroplast DNA. Bruick and Mayfield (1999) have proposed a model for translational initiation in Chlamydomonas chloroplasts whereby the 30S ribosomal subunit binds to a remote SD sequence far upstream of the initiation codon and remains in stand-by mode until a nuclear encoded message-specific translational activator binds downstream and guides the ribosomal subunit to the initiation codon. The support for this model comes from the finding that the mutagenesis of a remote SD-like sequence in psbA mRNA results in a significant reduction in translational activity in vivo (Mayfield et al., 1994) and the deletion of sequences that places this SD sequence closer to the initiation codon abolished translation whereas the same message was competent for translation in E.coli. (Bruick and Mayfield, unpublished). It is possible that a similar mechanism is responsible for guiding the higher plant ribosome to the psbA initiation codon.

Despite the fact that the plastid translational apparatus has been modified since endosymbiosis, prokaryotic translational signals are still very effective in promoting translation. For example, a gene under the control of the bacteriophage T7 gene 10 (T7G10) promoter and translational leader sequence are efficiently expressed when T7 RNAP is present in the chloroplast compartment (McBride et al., 1994). Indeed, the evidence suggests that not only are T7G10 translational signals functional but that they are very efficiently translated. Staub et al. (2000) generated transplastomic tobacco plants expressing a human hormone from chloroplast transgenes. They found that a transgene containing the psbA promoter fused to the T7G10 leader sequence was expressed many times more efficiently than an otherwise identical transgene that contained the psbA promoter and 5'-UTR even though mRNA levels for both were similar.
1.2. PLASTID TRANSFORMATION IN HIGHER PLANTS

Plastid transformation in higher plants is a multistep process involving (1) the delivery of the transforming DNA into a chloroplast in a plant leaf cell or protoplast, (2) integration of the transforming DNA into the plastome through homologous recombination between plastid DNA sequences in the transforming DNA (donor DNA) and the corresponding sequences in the plastome (recipient DNA) and (3) shoot regeneration on medium containing selection for the transformed plastome (transplastome) and (4) the eventual sorting out of transformed and untransformed plastids resulting in homoplasmy. Plastid-specific non-lethal selectable markers are used to select for the transplastome which may be encoded by the donor DNA itself in the form of mutant plastid ribosomal genes that replace the corresponding wild-type copies. Alternatively, the selectable marker used may be a bacterial antibiotic resistance gene that integrates into the plastome through two homologous recombination events between flanking donor DNA and the plastome. The selectable markers enable the plastid ribosome to function in the presence of antibiotics and therefore transplastomic plants are identified by the regeneration of green shoots from bleached wild-type tissue on shoot regenerating medium containing antibiotics. A leaf cell can contain up to 50,000 plastomes and therefore it can take several rounds of shoot regeneration in the presence of antibiotics to obtain uniformly transformed (homoplastic) plants.

1.2.1. Transformation procedures

Stable plastid transformation is now routinely carried out in tobacco using the biolistic and the polyethylene glycol (PEG) DNA delivery systems. The biolistic process was used by Svab et al. (1990) to generate the first stable transplastomic higher plant from tobacco leaf tissue. The biolistic process is technically simple and is the method of choice for transient gene expression analyses or stable genetic transformation in plastids. In the biolistic process microscopic tungsten or gold particles are coated with the transforming DNA and accelerated into leaf tissue by an explosive charge or a blast of high pressure helium gas. This is a physical process that allows the delivery of the transforming DNA directly into the plastids without killing the cell. An advantage of this process is that it can be used to deliver DNA into any tissue from any cultivar regardless of genotype.
Stable transplastomic tobacco plants have also been generated from tobacco mesophyll protoplasts that take-up the transforming DNA when treated with polyethylene glycol (PEG) (O’Neill et al., 1993; Golds et al., 1993). Although this system is cost effective and efficient in the recovery of transformants (Koop et al., 1996) the requirement for an efficient protoplast isolation and culturing procedure limits the range of species to which this technique can be applied. In addition this method is technically demanding and requires specialised tissue culture skills.

It is also of interest that it has been possible to directly inject DNA into tobacco leaf chloroplasts using an extremely fine-tipped syringe driven by the thermal expansion of galinstan, a liquid metal alloy (Knoblauch et al., 1999). Using this system it was possible to transiently express GFP in a single injected chloroplast and to observe GFP activity spreading to neighbouring chloroplasts. This system is technically demanding and may not be feasible for use in the production of stably transformed plants. However, it will be useful in determining whether DNA or protein molecules can be exchanged between chloroplasts and may find an application in the transformation of species that prove difficult to transform using currently available methods.

1.2.2. Selectable markers

The most commonly used selectable markers in the stable transformation of plastids confer resistance to both the spectinomycin and streptomycin antibiotics which are potent inhibitors of plastid protein synthesis in dicotyledonous plants. In the first transplastomic tobacco plant Svab et al. (1990) replaced a 3.7kb plastid DNA region of the tobacco plastome with a plastid fragment that contained a mutant 16S rRNA gene which encoded resistance to both spectinomycin and streptomycin and transplastomic lines were selected by shoot regeneration in the presence of spectinomycin. Kavanagh et al. (1999) have generated transplastomic tobacco by substituting the 16S rRNA gene and the 3’ portion of the rps12 gene with mutant counterparts from the Solanum nigrum plastome that encode spectinomycin and streptomycin resistance, respectively.

Svab and Maliga (1993) generated stable transplastomic tobacco plants by inserting into the plastome a chimeric bacterial gene, aadA, which encodes an aminoglycoside 3”-adenylyltransferase (AADA) which detoxifies both spectinomycin and streptomycin. The kanamycin resistance gene has also been used as a selectable marker in the transformation of tobacco plastids (Carrer et al., 1993). Recently a fluorescent antibiotic resistance marker which is at least as
efficient as aadA has been developed for plastid transformation (Khan and Maliga, 1999). It consists of GFP translationally fused to AADA and can be used to visually identify transformed and homoplasmic tissue.

The mechanism by which antibiotic resistance is conferred is different for mutant plastid ribosomal genes and aadA. The products of the mutant plastid 16S rRNA and rps12 genes are incorporated into the plastid ribosome and prevent the antibiotics from binding in a process referred to as binding-type antibiotic resistance. On the other hand the aadA gene encodes an enzyme which inactivates the antibiotic. Binding-type antibiotic resistance markers appear to be more stringent in the elimination of wild-type plastomes during shoot regeneration because typically only one round of regeneration is required to obtain homoplasmic plants (Kavanagh et al., 1999) whereas AADA selection requires at least two rounds of regeneration (Svab and Maliga, 1993). It has been proposed that binding-type antibiotic resistance is more stringent because in this case a plastid only becomes fully resistant when all plastome copies are mutant and resistant plastids give no cross protection to susceptible plastids in the same cell (Dix and Kavanagh, 1995). On the other hand the detoxifying activity of AADA produced within resistant plastids may make the entire cell resistant which reduces the selection pressure on the remaining population of susceptible plastids.

The range of selectable markers available for use in plastid transformation experiments is quite small and this may become a problem in the transfer of transplastomic technology to graminaceous species that are naturally resistant to both spectinomycin and streptomycin. Other potentially selectable markers identified in higher plants that could be used in plastid transformation are mutant 23S rRNA and psbA genes that confer resistance to the antibiotic lincomycin (Cseplo and Maliga, 1984) and the triazine herbicides (Goloubinoff et al., 1984), respectively. It may also be possible to identify plastid mutants in higher plants similar to the mutant 16S rRNA and atpB genes identified in Chlamydomonas chloroplast DNA which confer resistance to erythromycin (Newman et al., 1990) and tentoxin (Avni et al., 1992), respectively.

A particularly interesting candidate for use as an alternative selectable marker gene is the mutant plastid psbA gene which confers resistance to the herbicide triazine. If this selectable marker could be used successfully in the selection of transplastomic plants it would reduce the need to use antibiotic resistance genes which continues to be of concern to the public. However, Dix and Kavanagh (1995) have suggested that this selectable marker may not be
sufficiently strong for the initial selection of transplastomes given that *psbA* expression is dependant on photosynthesis. Recently, transplastomic potato plants have been regenerated from leaf explants using a chimeric *aadA* gene for selection that contained the *psbA* gene expression signals (Sidorov et al., 1999). This demonstrated that the photosynthetic *psbA* expression signals are suitable for the recovery and rooting of transformants. Furthermore, the *16S rRNA* promoter which was previously thought to be active only in chloroplasts was sufficient to drive *aadA* expression in proplastids to select transplastomic rice plants from embryogenic cells (Khan and Maliga, 1999). Thus, the mutant *psbA* selectable marker could prove to be an effective selectable marker for plastid transformation in dicots and cereal crops.

1.2.3. Transformation vectors
Plastid transformation vectors typically consist of a standard *E. coli* cloning vector containing a plastid DNA fragment which targets the region of the plastome that is to be transformed. Because plastid transformation proceeds through homologous recombination, only plastid DNA sequences and heterologous sequences flanked by plastid DNA will be incorporated into the plastome. The amount of plastid DNA that is included in plastid transformation vectors generally varies from 3 Kb (Svab and Maliga, 1993) to 8 Kb (Kavanagh et al., 1999). Most plastid transformation vectors utilise the *aadA* gene as the selectable marker because it has been reported to be required for high efficiency transformation (Svab and Maliga, 1993) and because by selecting appropriate flanking sequences from the plastome, *aadA* based vectors can potentially be used to target any region of the plastome for transgene integration. In contrast, vectors in which selection is based on mutant plastid ribosomal genes, can target only the corresponding region of the plastome. This apparent limitation may not in effect limit their usefulness because it has been shown that genes present on two unlinked DNA molecules can be simultaneously integrated into the plastome and “co-transformed” homoplasmic plants can be recovered based on selection for only one of the transforming DNAs (Carrer and Maliga, 1995).

1.2.4. Transformation efficiency
Stable plastid transformation is now routine in tobacco and it has been reported that transformants can be recovered at an efficiency of one transformant per leaf bombarded with pZS197, a transformation vector utilising *aadA* as the selectable marker (Svab and Maliga, 1993). Svab and Maliga (1993) reported that pZS197 was 100 times more efficient in the recovery of transplastomic tobacco than a transformation vector that contained a mutant *16S rRNA* gene as the selectable
marker. They proposed that the increase in transformation efficiency was due to
the improved recovery of transplastomic lines due to the superiority of \textit{aadA}
based selection. However, Koop \textit{et al.} (1996) determined using the PEG-
mediated plastid transformation system that usually 20 to 40 transformants are
recovered for every $10^6$ treated protoplasts using \textit{aadA} as the selectable marker.
Kavanagh \textit{et al.} (1999) determined using the same transformation system that
approximately 10 transformants were recovered for every $10^6$ treated protoplasts
using mutant plastid ribosomal genes as the basis of selection. Therefore, it seems
likely that the selectable marker used does not have a major bearing on the overall
efficiency of transformation.

Other factors that may influence the efficiency of plastid transformation
are the plastome site chosen for transformation, the amount of homologous
plastid DNA included in the transformation vector for targeting and the plastid
regulatory sequences chosen to express the selectable marker in vectors utilising
\textit{aadA} selection. Sidorov \textit{et al.} (1999) reported that the efficiency of recovery of
transplastomic potato plants was similar when \textit{aadA} was targeted to the single
copy region or the inverted repeat region of the plastome. Koop \textit{et al.} (1996)
found that transformation efficiencies were comparable using two \textit{aadA} based
vectors despite differences in integration sites, sizes of homologous flanking
regions and plastid regulatory sequences used for \textit{aadA} expression. This indicates
that transformation efficiencies are not greatly influenced by the above mentioned
differences between transformation vectors.

The prospect of developing a 'universal vector' for plastid
transformation in many different species would be an important development that
would facilitate the transfer of plastid transformation technology to other species.
However, sequence divergence between the donor and target plastid sequences
may have a bearing on the feasibility of this approach if homeologous
recombination significantly reduces the overall efficiency of transformation.
Kavanagh \textit{et al.} (1999) demonstrated that a 7.8kb plastid DNA fragment from
\textit{S. nigrum} that was 2.4% divergent from the corresponding tobacco plastome
region could integrate into the tobacco plastome by homeologous recombination
and that transplastomic lines could be recovered at a frequency not much less than
that achieved by Koop \textit{et al.} (1996) using donor DNA completely homologous to
the target DNA. Thus, homeologous recombination does not appear to
significantly reduce transformation efficiency.
1.2.5. Transformation in species other than tobacco

Although stable plastid transformation is now routine in tobacco it has only been reported in two other species, *Arabidopsis* (Sikdar et al., 1998) and potato (Sidorov et al., 1999), in the ten years since the generation of the first transplastomic plant. One of the reasons for this is that the standard plastid transformation protocols utilise leaf tissue or protoplasts as the target tissue and in other species these cells are not as totipotent as they are in tobacco. Thus, the recovery of transformants is hampered by the poor shoot regeneration response. However, Sidorov et al. (1999) successfully adapted the tobacco plastid transformation protocol to potato by mechanically wounding the potato leaf tissue to improve the shoot regeneration response sufficiently to recover transplastomic potato plants. Even so, the frequency of recovery was still an order of magnitude less than that reported for tobacco. Thus, it may be necessary to use a more totipotent cell type for generating transplastomic plants in other species. For example, Hall et al. (1996) found that the stomatal guard cells were the most totipotent cells in sugar beet leaves.

In the Gramineae, only cells derived from embryogenic tissue are totipotent and therefore this tissue is most often used in transformation protocols. There was some doubt as to whether the biolistic process would be effective in the delivery of DNA into the proplastids in this tissue type because they are about the same size as the projectiles commonly used in the biolistic transformation of plant cells (Bilang and Potrykus, 1998). However, Khan and Maliga (1999) have generated heteroplasmic transplastomic lines from rice embryogenic cell suspensions at a respectable frequency using the biolistic process. Given that the biolistic process can be utilised in plastid transformation in embryogenic cells it may be worth investigating the use of this tissue type as an alternative to leaf cells in the transformation of dicotyledonous species.

1.2.6. Genetic improvement of crop plants

1.2.6.1. Advantages of plastid transformation over nuclear transformation

Although *Agrobacterium*-mediated nuclear transformation in many crop species is now routine and several transgenic crops are now grown on a commercial scale, transformation of the plastome has several advantages over nuclear transformation in the genetic enhancement of crop plants. (1) Transplastomic plants can contain up to 50,000 transgenes per cell which dramatically increases expression levels compared to nuclear transgenes (Staub *et al.*, 2000; McBride *et al.*, 1995). For
example, transplastomic plants have been utilised in the production of recombinant *Bacillus thuringiensis* crystal toxin protein to between 3 and 5% (McBride *et al.*, 1995) and recombinant human somatotropin to greater than 7% (Staub *et al.*, 2000) of the total soluble protein (tsp) in leaf tissue. Whereas, the maximum expression level of crystal toxin genes in the nucleus is about 0.6% of tsp despite gene modification to increase expression level (Sutton *et al.*, 1992) and Staub *et al.* (2000) only managed to express somatotropin from a nuclear transgene to 0.025% tsp. High levels of expression may be required to prevent the emergence of crystal toxin resistance among insects and to make the production of recombinant proteins in plants economically feasible.

(2) Plastid transformation is mediated by homologous recombination which means that all transgenic lines are identical and therefore the number of lines that need to be screened is small. In contrast, nuclear transformation involves random insertion which gives rise to position effects and results in varying expression levels between transgenic lines. In addition, nuclear transgenes may be subject to gene silencing effects. (3) Polycistronic operon expression would make the introduction of multigenic pathways (eg. the nitrogen fixing pathway) more feasible in transgenic plants. High levels of gene expression from polycistronic units is possible in chloroplasts even without processing to give monocistronic transcripts (Staub and Maliga, 1995). However, in the nucleus the presence of an upstream cistron can drastically reduce the expression of a downstream cistron (Angenon *et al.*, 1989) which suggests that polycistronic operon expression is not feasible in nuclear transformants. (4) In most higher plants, plastids are maternally inherited which means that the likelihood that chloroplast transgenes might be transmitted via pollen to wild relatives is small (Scott and Wilkinson, 1999).

1.2.6.2. Feasibility of producing recombinant proteins in chloroplasts

In addition to the obvious advantages conferred by high expression levels there are other chloroplast features which indicate that it is a good candidate for recombinant protein production. For example, in bacterial systems the removal of the N-terminal methionine is important in the production of fully functional recombinant human hemoglobin (Hoffman, *et al.*, 1990). In order to achieve this efficiently in bacteria co-expression of a methionine amino peptidase (MAP) is required (Shen *et al.*, 1997). However, Staub *et al.* (2000) found that recombinant human somatotrophin produced in chloroplasts had the N-terminal methionine removed with greater than 95% efficiency which suggests that the
chloroplast contains an efficient MAP activity. In addition, somatotrophin requires controlled disulfide bond formation for biological activity and no known plastid-encoded protein contains disulfide bonds. However, analysis of the recombinant somatotropin produced in chloroplasts revealed that it contained the correct disulfide bonds which indicates that although the chloroplast stroma is a strongly reducing environment, the production of appropriate disulfide bonds in recombinant proteins is not compromised.

1.3. BACTERIOPHAGE T7 RNA POLYMERASE BASED EXPRESSION SYSTEMS

The RNAP encoded by gene 1 of bacteriophage T7 specifically recognises several highly conserved 23 nucleotide promoters in the phage genome. In addition, T7 RNAP is very active and can elongate transcripts about five times faster than E.coli RNAP (Chamberlin and Ring, 1973; Golomb and Chamberlin, 1974). For these reasons T7 RNAP has been successfully utilised for controlled and high level expression of foreign genes in E.coli (Studier and Moffatt, 1986). Indeed, it is so effective in E.coli that in some cases the host RNAP cannot compete and T7 RNAP becomes responsible for almost all cellular transcription (Studier and Moffat, 1986). This finding highlights the importance of controlling T7 RNAP expression in E.coli expression systems by placing it under the control of an inducible promoter in order to ensure that cells can proliferate and mature before inducing the expression of a transgene.

T7 RNAP has also been successfully utilised in mammalian systems to direct the expression of a reporter gene carried into the cell in a modified adenovirus vector (Tomanin et al., 1997). In this system human cell lines were transfected with modified adenovirus vectors containing sequences encoding a nuclear targeted T7 RNAP and a reporter gene under the control of the T7 promoter. Significant levels of reporter gene expression was detected when cells were coinfected with both vectors.

A T7 RNAP-based chloroplast expression system was developed in tobacco by McBride et al. (1994) in which a constitutively expressed chloroplast-targeted T7 RNAP transactivated the expression of a plastid transgene (GUS) to a high level. More recently an inducible chloroplast-based T7 RNAP expression system was developed by placing the chloroplast targeted T7 RNAP under the control of the inducible tobacco promoter from the PR-1a gene (Heifetz et al.,
unpublished, referred to in Heifetz, 2000). They found that GUS expression from a plastid bourne transgene was increased 5000-fold seven days after the induction of the PR-la gene promoter by a single foliar application of BTH, a benzothiadiazole derivative. This demonstrates that this system is effective in producing high levels of a foreign protein in a controlled manner and in a short period of time following induction.

It is interesting to note that there are currently no reports of T7 RNAP-mediated expression of nuclear transgenes in plants despite the fact that T7 RNAP is an ideal candidate for directing specific and high levels of expression in other systems. However, given that high levels of mRNA accumulation may trigger gene silencing in plants (reviewed in Waterhouse et al., 1999) it is possible that the extreme activity of T7 RNAP may be the reason that nucleus-based T7 RNAP expression systems have not been successfully developed. Furthermore, it has been reported that T7 RNAP is unable to processively transcribe higher eukaryotic chromatin in transgenic Drosophila (McCall and Bender, 1996) or when added to isolated mammalian nuclei (Jenuwein, et al., 1993). If the fundamental structure of eukaryotic chromatin inhibits transcript elongation by T7 RNAP then it would be a significant barrier to the development of a nucleus-based T7 RNAP expression system in plants.

Indeed localising the T7 RNAP in chloroplasts can present some problems of its own. McBride et al. (1994) found that pollen from tobacco producing high levels of T7 RNAP activity was infertile in crosses with chloroplast transformed lines. Also the plants expressing high levels of T7 RNAP activity grew abnormally (personal communication). It is conceivable that T7 RNAP in the plastid could direct the production of anti-sense mRNA which could block translation or T7 RNAP could bind to plastid DNA and interfere with processes like replication or transcription. But the likelihood of T7 RNAP specifically interacting with the plastome seems small given that the T7 RNAP promoter is a 23-bp sequence and the tobacco plastome is only 155kb. However, natural variants of the T7 promoter do occur and there are reports that this promoter can tolerate single nucleotide substitutions at different sites in in vitro transcription assays (Diaz, et al., 1993; Chapman and Burgess, 1987).
CHAPTER 2

PRODUCTION OF TRANSGENIC TOBACCO PLANTS EXPRESSING GENE FUSIONS ENCODING A CHLOROPLAST-TARGETED T7 RNA POLYMERASE

2.1. INTRODUCTION

Ribulose-1,5-bisphosphate carboxylase (Rubisco) is the chloroplast enzyme active in photosynthetic CO₂ fixation. The holoenzyme consists of eight large subunits (LSU) encoded by the plastidial rbcL gene and eight small subunits (SSU) encoded by the nuclear rbcS genes. Rubisco is probably the most abundant protein in leaves where it may account for more than 50% of the total soluble protein (Kung, 1976). Tobacco SSU is synthesised on cytoplasmic ribosomes as a 20 Kd precursor protein which is translocated into the chloroplast stroma by means of an N-terminal targeting sequence known as a transit peptide (Highfield and Ellis, 1978). The transit peptide is cleaved by the stroma-processing peptidase to generate the 14 Kd mature SSU protein (mSSU) which associates with LSU in the Rubisco complex.

RbcS genes are amongst the most highly expressed genes in leaves and their expression is strongly light-regulated (Dean, 1989). The rbcS promoter has been shown to be more active in tobacco leaf tissue than the CaMV 35S promoter (Jefferson et al., 1987) and is therefore an attractive candidate for driving high-level expression of foreign proteins in leaves. Also, because the rbcS promoter is predominantly active in green tissues, it should reduce the possibility of a transgene product causing deleterious effects in reproductive tissues or during embryo development. RbcS genes have been cloned and sequenced for many plant species including the tobacco Ntss23 gene (Mazur and Chui, 1985) used here to direct light-regulated expression of chloroplast-targeted T7 RNA polymerase fusion proteins in the leaves of transgenic tobacco plants.

T7 RNA polymerase (T7 RNAP), encoded by gene I of bacteriophage T7 (hereafter referred to as the T7 RNAP gene) specifically recognises several highly conserved 23 bp promoters in the phage genome. Because of its high
promoter specificity, T7 RNAP-based expression systems have been developed for the selective and controlled expression of foreign genes in bacteria (Studier and Moffatt, 1986) and for transient expression in eukaryotic cells (Tomanin et al., 1997). An analogous T7 RNAP-based plastid-localized expression system was developed by McBride et al. (1994) in which a chloroplast-targeted SSU-T7 RNAP fusion protein expressed from a nuclear transgene in tobacco was used to transactivate transcription of a second transgene located on the plastid genome.

The SSU transit peptide coding region derived from a pea rbcS gene has been used to direct attached proteins into chloroplasts in transgenic tobacco either on its own (Van den Broeck et al., 1985) or when 5 N-terminal amino acid residues of mSSU are included in the fusion protein (Kanevski and Maliga, 1994). Also, sequences from an Arabidopsis rbcS gene coding for the transit peptide and the N-terminal 24 residues of mSSU directed a truncated crystal toxin protein into chloroplasts in transgenic tobacco (Wong et al., 1992). Although the SSU transit peptide alone is sufficient to target foreign proteins into chloroplasts in transgenic tobacco, it is thought that N-terminal residues of mSSU may play an important role in the targeting efficiency (Wasmann et al., 1986).

Wasmann et al. (1986) found from in vitro chloroplast import studies that the SSU transit peptide promoted more efficient chloroplast targeting of neomycin phosphotransferase (NPT-II) when the first 23 mSSU residues were included in the fusion protein. But when expressed in transgenic tobacco, less than 10% of the SSU-NPT-II fusion protein with the mSSU sequence was localized in chloroplasts (Kuntz et al., 1986) whereas most of the SSU-NPT-II fusion protein lacking the mSSU sequence was chloroplast localized (Van den Broeck et al., 1985 and Kuntz et al., 1986). Therefore the role that N-terminal residues from mSSU play in targeting fusion proteins to the chloroplast is uncertain.

Our aim, in this chapter, was to produce fertile transgenic plants expressing a chloroplast-targeted T7 RNAP as the first step in the development of an expression system similar to that described by McBride et al. (1994) (discussed above). The transgenic plants generated by McBride et al. (1994) expressed an SSU-T7 RNAP fusion protein comprising the SSU transit peptide and 12 N-terminal amino acids of mSSU fused to amino acid residue 2 of T7 RNAP under the transcriptional control of a “double” CaMV 35S promoter. Although T7 RNAP-dependent expression of plastid-localized transgenes was demonstrated, no direct biochemical data were presented concerning the extent of
plastid localization of the T7 RNAP fusion protein. Furthermore, pollen from tobacco producing high levels of T7 RNAP activity was infertile in crosses with chloroplast transformed lines (McBride et al., 1994) and plants expressing high levels of T7 RNAP activity grew abnormally (personal communication).

The gene fusions we describe differ from those described by McBride et al. (1994) with respect to (i) the amino acid sequence at the SSU-T7 RNAP fusion junction and (ii) the use of rbcS promoter sequences instead of the CaMV 35S promoter to drive gene expression.

We also set out to investigate whether the efficiency of targeting of an SSU-T7 RNAP fusion protein to chloroplasts was affected by the number of N-terminal residues of mSSU located between the transit peptide and the T7 RNAP domains. Therefore we constructed three chimeric genes in which tobacco rbcS sequences (Ntss23; Mazur and Chui, 1985) encoding the SSU transit peptide and either the first 8, 14 or 25 amino acids of mSSU were fused in-frame with the coding sequence of T7 RNAP. (Hereafter, we refer to these chimeric genes as st transgenes (or specifically st8, st14 and st25) and to their encoded proteins as ST fusion proteins (or specifically STS, STM, ST25)). Each st transgene was placed under the transcriptional control of the rbcS promoter sequences derived from Ntss23. The GUS reporter gene, under the control of the constitutive CaMV 35S promoter, was included in the binary vectors linked to the st genes to function as a marker of transformation and to follow the segregation of st genes in progeny plants.

Transgenic plants expressing ST fusion proteins were analysed as follows:

1) ST fusion protein expression levels were determined by western blot analysis of total leaf protein homogenates using an anti-T7 RNAP polyclonal antibody.
2) The T7 RNAP activity of ST fusion proteins was determined by T7 RNAP activity assays of total leaf protein homogenates.
3) Leaf specific expression was demonstrated by analysis of leaf and root tissue.
4) ST fusion protein stability was assessed by western blot analysis and determination of T7 RNAP activity in total leaf protein extracts from transgenic seedlings placed in the dark over a 16 day period.
5) Inheritance patterns of st transgenes in progeny plants were determined and homozygous lines were identified by western blotting and testing for the inheritance of the linked nptII and GUS marker genes.
6) The sub-cellular localization of ST fusion protein was determined by western blot analysis and T7 RNAP activity assays on total leaf protein homogenates and gradient-purified intact chloroplasts.

7) A northern blot analysis using chloroplast gene-specific probes was undertaken to compare transcription in chloroplasts of non-transgenic and transgenic plants expressing ST fusion proteins.
2.2. MATERIALS AND METHODS

2.2.1. Bacterial strains and plasmids

*Escherichia coli* DH5α (Hanahan, 1983) was used as the host for cloning experiments and for the amplification of plasmid DNA for purification. *Escherichia coli* HB101 (Boyer and Roulland-Dussoix, 1969) containing pRK2013 (Ditta et al., 1980) was used as the helper strain in triparental mating. *Agrobacterium tumefaciens* LBA4404 carrying the non-oncogenic helper Ti plasmid pAL4404 was used as the host in triparental mating (Hoekema et al., 1983).

pCAR2 (unpublished plasmid) a pUC19-based plasmid which contains a 2.3 kb Hind III-Xbal DNA fragment comprising the tobacco *rbcS* gene Ntss23 (as described in Figure 3 of Mazur and Chui, 1985) was used as the source of *rbcS* sequences.

pAR3132 (Dunn et al., 1988), a pBR322-based plasmid was the source of codons 11-883 of T7 RNAP used in the construction of pBinTS25 and in the making of a probe for northern analysis.

pRok2 (Hilder et al., 1987) was the pBIN19-based binary vector used in the construction of pBinTS25.

pRok1/7 EE (unpublished plasmid) a pUC7-based vector which contains the GUS expression cassette from pROK1 (Baulcombe et al., 1986) was the source of the *CaMV 35S* promoter fragment and pCTAK1 (unpublished plasmid) was the source of the GUS : NOS terminator fragment used in the construction of pBinTGS25.

2.2.2. Cloning Enzymes and Molecular cloning techniques

Restriction enzymes were obtained from New England Biolabs, Boehringer Mannheim, Promega and GibcoBRL. T4 DNA ligase and calf intestinal alkaline phosphatase were purchased from Boehringer Mannheim. Nucleic acid manipulations were performed as described by Sambrook et al. (1989) unless otherwise stated.

2.2.3. *E.coli* growth media

*E.coli* was grown on Luria-Bertani (LB) broth (10g Bacto-tryptone, 5g Bacto-yeast extract, and 5g NaCL per litre, pH 7.0) and supplemented with 1.5% agar
for solid medium which was supplemented with Ampicillin (100mg/L) or Kanamycin (50mg/L) to select for plasmid containing cells.

2.2.4. Transformation of *E.coli*

*E.coli* competent cells were prepared and transformed using a variation of the method developed by Mandel and Higra (1970). 1ml of an overnight culture was inoculated into 40ml of liquid LB medium in a 250ml flask and incubated at 37°C with shaking at 200 rpm to the exponential stage. The cells were pelleted by centrifugation at 4000 rpm for 10 min at 4°C and resuspended in 20ml of 0.1M CaCl₂. Cells were pelleted by centrifugation for 5 min and resuspended in 2ml of 0.1M CaCl₂, 15% glycerol and stored at -70°C. Plasmid DNA and ligation reactions were mixed with 20µl and 100-200µl of competent cells respectively and incubated on ice for 30 min. Heat shock was carried out at 42°C for 1 min and on ice for 30 seconds. Four volumes of LB broth was added and the mixture was incubated at 37°C for 1 hour with 100 rpm shaking followed by plating on LB agar plates containing the appropriate antibiotic.

2.2.5. Purification of plasmid DNA (boiling method)

This method was derived from the method of Holmes and Quigley (1981). Transformed *E.coli* single colonies were grown in 10ml LB broth at 37°C overnight, harvested by centrifugation at 3500 rpm for 10 min at 4°C and resuspended in 60µl of 50mM Tris-HCl, pH 8.0, 10mM EDTA, pH 8.0, 25% sucrose. 20µl of Lysozyme (10mg/ml in 50mM Tris-HCl, pH 8.0) was added with mixing and after 5 min 500µl of 50mM EDTA, pH 8.0, 50mM Tris-HCl, pH 8.0, 5% sucrose, 5% Triton X-100 was added. Boiling for 1 min followed by centrifugation at 15000 rpm for 15 min precipitated and sedimented chromosomal DNA and other cell debris. The supernatent was treated with 5µl of RNase (10mg/ml) at 42°C for 15 min. The solution was phenol extracted once and plasmid DNA was precipitated by adding one volume of isopropanol. The plasmid DNA was recovered by centrifugation at 15000 rpm for 10 minutes followed by washing of the pellet with 70% ethanol. Binary and pUC based plasmids were resuspended in 10µl and 40µl of sterile deionised water respectively.

2.2.6. Oligonucleotides and sequencing

The following oligonucleotides were used in sequencing reactions to confirm accurate construction of *st* gene fusions.

1) *rbcS* promoter oligo : 5'-ATATAGGGTGCGGTGGC-3' which corresponds to positions 957-974 in the SSU gene Ntss23 (Mazur and Chui, 1985).
2) rbcS transit peptide oligo (Tp3): 5′-GCATGAATTCATCGATACAATGGCTTCCTCA G-3′. The last 16 nucleotides correspond to positions 1045-1060 of Ntss23 of which the last 13 nucleotides occur in transit peptide coding sequence. The Sanger dideoxy method (Sanger et al., 1977) was used for all sequencing reactions and PCR cycle-sequencing reactions were carried out using the fluorescent dye terminator kit supplied by Amersham. Sequencing reactions were performed according to manufacturers instructions and were analysed on an Applied Biosystems model 373A automated sequencer.

2.2.7. Triparental mating
Binary vectors containing the various gene constructs were transformed into A. tumefaciens LBA4404 using the triparental method described by Bevan (1984). Putative transformants were checked for the presence of the constructs by purifying plasmid DNA from A. tumefaciens cultures grown in LB containing streptomycin 500mg/L and kanamycin 50mg/L and transforming the binary plasmid into E.coli DH5α in order to obtain sufficient quantities of plasmid DNA for restriction analysis.

2.2.8. In vitro plant material
Nicotiana tabacum (tobacco) cv Samsun NN was maintained under sterile conditions in MS medium (Murashige and Skoog, 1962) in a growth room under a 16 hour light/8 hour dark cycle at 23-25°C. Sterilisation of tobacco seeds for plating was performed by soaking briefly with absolute ethanol and soaking for 10 min in a 20% bleach solution. After soaking the seeds were washed three times with sterile distilled H2O.

2.2.9. Transformation of tobacco
A. tumefaciens mediated transformation of tobacco nuclear DNA was performed by the method described by Horsch et al. (1985). A. tumefaciens containing the various st gene constructs was grown overnight in LB broth in the presence of streptomycin 500mg/L and kanamycin 50mg/L. Cells were pelleted by centrifugation at 4000 rpm for 10 min at 4°C and resuspended in MS liquid medium (Murashige and Skoog, 1962). The suspension was incubated for 30 min in a sterile environment with freshly prepared leaf discs from in vitro tobacco plants. The leaf discs were then transferred onto shoot regeneration NBM medium (MS containing 0.8% agar, benzylaminopurine (BAP) 1μg/ml, α-naphthalene acetic acid (NAA) 0.1μg/ml and B5 vitamins; inositol 100μg/ml, thiamine 10μg/ml, nicotinic acid 1μg/ml and pyridoxine-HCl 1μg/ml). After two days of cocultivation the leaf discs were transferred to NBM medium containing
carbenicillin 200mg/L and kanamycin 50mg/L. Individual shoots that appeared on selective medium after 3-4 weeks were cut and transferred to MS medium containing the same selection. Shoots that rooted in the presence of selection were planted in the greenhouse for analysis.

2.2.10. Preparation of tobacco protein extracts
In general crude leaf, root and seedling extracts for western blots and GUS assays were prepared by grinding tissue with a mortar and pestle on ice in 50mM NaH₂PO₄, pH 7.0, 10mM β-mercaptoethanol, 0.1% Triton X-100 at a 5 : 1 buffer : tissue ratio followed by a brief spin in a microfuge to remove debris. Total leaf protein extracts that were prepared for T7 transcription assays alone or for transcription assays and western blotting, were prepared in the same way but using a Tris-based buffer (50mM Tris-HCl, pH 7.0, 5mM EDTA, 10mM β-mercaptoethanol, 1mM PMSF, 0.1% Triton X-100). In some cases total leaf protein extracts that were resolved by SDS-PAGE alongside purified intact chloroplasts were extracted in 330mM Sorbitol, 50mM Hepes-KOH, pH 7.7 (SH) buffer supplemented with 10mM B-mercaptoethanol, 1mM PMSF, 0.1% Triton X-100.

2.2.11. Preparation of Percoll gradients
Intact chloroplasts were purified from crude chloroplast pellets by sedimentation in Percoll gradients as described by Bartlett et al. (1982). PBF stock Percoll for gradients was prepared by adding 30mg/ml polyethyleneglycol (4000), 10mg/ml bovine serum albumin, 10mg/ml Ficoll to Percoll (colloidal PVP coated silica from Sigma) and mixing at room temperature for 2 hours. 40% and 80% PBF Percoll stocks were prepared in 330mM Sorbitol, 50mM Hepes-KOH, pH 7.7, 1mM MgCl₂, 1mg/ml L-Ascorbic acid, 250μg/ml Glutathione. Gradients were prepared by carefully layering 5ml of 40% Percoll on top of 5ml of 80% Percoll in a 15ml corex tube and stored at 4°C until use.

2.2.12. Preparation of intact chloroplasts
The method used for the purification of intact chloroplasts from tobacco leaves was based on the method developed by Bartlett et al. (1982). Because tobacco chloroplasts are highly sensitive to mechanical grinding intact chloroplasts were carefully prepared from high quality leaf tissue taken from actively growing greenhouse plants that had not been exposed to heat or water stresses. All procedures were carried out on ice using instruments and buffers that had been prechilled. A healthy leaf approximately 15-20cm in length was detached immediately before preparation, washed with deionised H₂O and veins were
Typically about 10g of leaf tissue was homogenised by four 2 second bursts in a Moulinex blender in 200ml of ice cold GR buffer (330mM Sorbitol, 50mM Hepes-KOH, pH 7.7, 1mM MgCl₂, 1mM MnCl₂, 1mg/ml L-Ascorbic acid, 0.5mg/ml bovine serum albumin). The homogenate was filtered through four layers of Miracloth and the filtrate was centrifuged at 1500 rpm for 10 min at 4°C in a Sorvall RC-5B centrifuge using a prechilled GSA rotor. The crude chloroplast pellet was gently resuspended over the course of a 30 min period on ice in 4ml of GR buffer. Intact chloroplasts were purified by sedimentation in Percoll density gradients at 4000 rpm for 10 min at 4°C in a swing out rotor in an IEC Centra CL3R refrigerated centrifuge. Broken chloroplasts floated on top of the 40% Percoll layer and were removed with a Pasteur pipette before carefully removing the intact chloroplasts with a separate pipette from the 40%/80% interface. Intact chloroplasts were washed twice by adding 5 volumes of SH buffer, gently inverting the corex tube and sedimentation in the IEC Centra CLR3 at 1500 rpm for 10 min at 4°C. Washed intact chloroplasts were resuspended in 1ml of SH buffer. The typical yield from 10g of starting material was 1mg of chlorophyll.

2.2.13. Thermolysin treatment of intact chloroplasts
Thermolysin protease type X from Bacillus thermoproteolyticus rokko was purchased from Sigma as a 20% sodium/calcium salt and was dissolved in SH buffer to give a 4mg/ml solution which was stored at -20°C. 250µg of protease/mg of chlorophyll was added to the intact chloroplast suspension and the eppendorf was inverted twice and incubated on ice for 30 min. The eppendorf was inverted half way through the treatment to resuspend settling chloroplasts. The protease was inactivated by adding EDTA or EGTA to a final concentration of 10mM or 20mM respectively.

2.2.14. Determination of protein and chlorophyll concentrations
The concentration of protein in leaf, root and seedling protein extracts was determined using the method of Bradford (1976). The chlorophyll concentration in leaf protein extracts and intact and broken chloroplast preparations was determined using the method of Arnon (1949).

2.2.15. Western blotting
SDS-PAGE and western blotting was carried out essentially as described by Sambrook et al. (1989). Leaf, root and seedling protein extracts and intact and broken chloroplast fractions were prepared for SDS-PAGE as described by Laemmli (1970). Protein extracts were resolved in 6.75% separating gels in the
BioRAD SDS-PAGE apparatus run at 140 volts until the dye front eluted into the running buffer. Protein gels were prepared and stained with Coomassie Brilliant Blue R250. Duplicate protein extracts for western blotting were resolved in polyacrylamide gels at the same time and were transferred to nitrocellulose membranes (Schleicher & Schuell) in a BioRad Transblot™ apparatus at 150mA overnight. The membrane was blocked by treating with nonfat dried milk to prevent nonspecific binding of immunoglobulins and incubated with a 1 : 1500 dilution of rabbit anti-T7 RNAP polyclonal antibodies (a gift from Paul Fisher Department of Pharmacological Sciences, Health Sciences Center, State University of New York at Stony Brook, Stony Brook, NY 11794). Alkaline phosphatase coupled anti-rabbit antibodies were used in the secondary immunological reaction and the localisation of alkaline phosphatase was visualised by treatment with 5-bromo-4-chloro-3-indolyl phosphate and nitro blue tetrazolium.

2.2.16. Densitometry of protein gels and western blots
Relative amounts of Rubisco LSU and ST fusion protein in total leaf protein extracts and intact chloroplast fractions in protein gels and western blots were determined by capturing the image of gels and blots with a camera using UVPGrab Software and determining relative signal intensities using UVP GelWorks for Windows software.

2.2.17. T7 RNAP transcription assays
Purified native T7 RNA polymerase enzyme and assay buffer were supplied by Boehringer Mannheim. Aliquots (5µl) of total leaf, total seedling and purified chloroplast protein fractions were incubated in a 30µl final volume transcription assay mixture (33µg/ml phage T7 genome, 1X T7 RNAP assay buffer, 0.2mM rA/C/GTP, 0.1mM rUTP, 33µCi/ml 32p UTP (Amersham)) at 37°C for 30min. Reactions were placed on ice and 3µl of carrier solution (5mg/ml bovine serum albumin, 10mg/ml tRNA) was added. 500µl of 10% trichloroacetic acid was added to the reaction with mixing and the tube was placed on ice for 30 min. Precipitated transcripts were pelleted by centrifugation at 15000 rpm for 15 min at 4°C. Unincorporated 32P UTP was removed with the supernatant and safely discarded. Count per minute (CPM) values were determined for the radioactive pellet using the PACKARD TRI-CARB® 1500 liquid scintillation analyser or using a sensitive Geiger probe.
2.2.18. Fluorometric GUS assay

Determination of GUS specific activity by fluorometric assay was performed on total leaf protein extracts as described by Jefferson et al. (1987). GUS expression in total leaf protein extracts was visualised by exposing GUS assays to long wave ultraviolet light on a transilluminator.

2.2.19. Northern analysis

The double stranded DNA probe used for the detection of st gene transcription was obtained by gel purification of the 830 bp Hpal/Ndel T7 RNAP gene fragment from pAR3132. All chloroplast probes were obtained by PCR amplification of approximately 500bp of tobacco plastid DNA (ptDNA) from the transcriptional units of the following plastid genes: rpoC, psbD, rps18 (probe includes the 3'UTR of rpl33), clpP, rps12, 16S rRNA, ndhA and ndhH. Below is shown the oligos used in the PCRs and the position of the 5' nucleotide (nt) of each oligo in tobacco ptDNA (Shinozaki et al., 1986). C denotes that the identifying nucleotide is complementary to the nucleotide at that position in ptDNA.

<table>
<thead>
<tr>
<th>Gene</th>
<th>5' nt position in ptDNA</th>
<th>oligos for PCR amplification</th>
</tr>
</thead>
<tbody>
<tr>
<td>rpoC</td>
<td>22,028</td>
<td>5'-GGTACATGAACAGCCATTTG-3'</td>
</tr>
<tr>
<td></td>
<td>22,549(C)</td>
<td>5'-ATTAGTACAAGAAGCCGTGG-3'</td>
</tr>
<tr>
<td>psbD</td>
<td>34,463</td>
<td>5'-TGACTATAGCCCTTGGTAAG-3'</td>
</tr>
<tr>
<td></td>
<td>35,072(C)</td>
<td>5'-CCGGCAACTCCCATCATATG-3'</td>
</tr>
<tr>
<td>rps18</td>
<td>70,339</td>
<td>5'-TGTGTTCTACCTTCTCAAAGG-3'</td>
</tr>
<tr>
<td></td>
<td>70,998(C)</td>
<td>5'-AATGAACTCCGGGAAGGTAGAGTAG-3'</td>
</tr>
<tr>
<td>clpP</td>
<td>73,086</td>
<td>5'-GTTAAGAAGCTCTCCGGCATG-3'</td>
</tr>
<tr>
<td></td>
<td>73,665(C)</td>
<td>5'-CGCATGTACGGTTCCTAAAGG-3'</td>
</tr>
<tr>
<td>rps12</td>
<td>100,088/142,424(C)</td>
<td>5'-GGTGGATCTCGAAAGATATG-3'</td>
</tr>
<tr>
<td></td>
<td>100,647/141,901(C)</td>
<td>5'-GTAAAGGATCGTCAACAAGG-3'</td>
</tr>
<tr>
<td>ndhA</td>
<td>122,101</td>
<td>5'-CGCTTCCACTATATCACTG-3'</td>
</tr>
<tr>
<td></td>
<td>122,800(C)</td>
<td>5'-ATCAAGAAGATCTCACCAGT-3'</td>
</tr>
<tr>
<td>ndhH</td>
<td>123,937</td>
<td>5'-GCAAATTGATAAAACCTGGTG-3'</td>
</tr>
<tr>
<td></td>
<td>124,420(C)</td>
<td>5'-CGAGTTGAAGGCGTAGATTATG-3'</td>
</tr>
</tbody>
</table>

16S rRNA sequences not available

The PCR amplifications were carried out essentially as described by Sambrook et al. (1989) except that in the case of ndhA 2.5mM MgCl2 was used in.
the reaction to increase yield. The thermal cycling parameters were 94°C for 1 min, 55°C for 2 min, 72°C for 3 min for 30 cycles with a 5 min extended polymerisation step at the end of the reaction.

All buffers used for northern blot analysis were treated with DEPC or were made with DEPC-treated sterile deionised H$_2$O. Total RNA from tobacco was extracted as described by Kristel et al. (1996) and was prepared for electrophoresis by heating in sample buffer (50% v/v formamide, 15% v/v formaldehyde (37% stock), 00.1% v/v bromphenol blue, 1X running buffer) for 15min at 65°C and cooling on ice for 1 min before loading. The samples were electrophoresed at 60 volts for 4 hours in 1% agarose/formaldehyde gels (running buffer: 2.5M Boric acid, 20mM Tri-Sodium Citrate, pH was adjusted to 7.5 with NaOH). RNAs were directly transferred onto Hybond-N filters (Amersham) as described by Sambrook et al. (1989). Transferred RNA was fixed to the filter by exposure to ultraviolet light for 2 min on a UV transilluminator. DNA probes were labelled with $^{32}$P dCTP (Amersham) by the random primer labelling method (Feinberg and Vogelstein, 1983). Hybridisation of the probe to the filter and stringent washing was carried out as described for Southern blotting by Maniatis et al. (1982). Washed filters were placed in sealed plastic bags and autoradiographed at -70°C using AGFA CURIX BLUE HC-S PLUS medical X-ray films.
2.3. RESULTS

2.3.1. Construction of chimeric genes encoding ST fusion proteins

2.3.1.1. Construction of pBinTS25 and pBinTGS25

The plasmid pBinTS25 and its derivative pBinTGS25 contain promoter and coding sequences derived from a tobacco rbcS gene (Ntss23; Mazur and Chui, 1985) in a translational fusion with the gene coding for T7 RNAP. A schematic representation of the steps involved in the construction of these plasmids is shown in Figure 2.1.

In the construction of pBinTS25, pCAR2 (unpublished) was the source of the rbcS sequences and pAR3132 (Dunn et al., 1988) provided sequences encoding codons 11-883 of phage T7 RNAP. Figure 2.2 shows details of the nucleotide and predicted amino acid sequence in the region coding for the rbcS transit peptide in pCAR2. RbcS sequences comprising approximately 1 kb of the promoter region and adjacent sequences encoding the transit peptide and the N-terminal 25 amino acids of the mSSU protein were introduced upstream of the T7 RNAP gene by the ligation of a 1.4 kb HindIII/Mfel fragment from pCAR2 into HindIII/EcoRI-digested pAR3132 to give pSSU-T7. Correct ligation resulted in the in-frame fusion of codon 25 of the mature rbcS orf to codon 11 of T7 RNAP. This was confirmed using the oligonucleotide primer Tp3 to sequence across the gene fusion junction.

A 4.2 Kb HindIII/BamHI from pSSU-T7 was then cloned into HindIII/BamHI-digested pRok2, a derivative of pBin19 (Jefferson et al., 1987 and Bevan et al., 1984). Correct ligation resulted in the replacement of the CaMV 35S promoter in pRok2 with the st gene fusion upstream of the NOS terminator. In addition, the SV40 early polyadenylation signal located downstream of the T7 RNAP gene in pAR3132 was retained in all our constructs. Confirmation of the structure of the resulting plasmid, pBinTS25 was obtained by restriction digestion (Figure 2.3).

pBinTGS25 was constructed to contain a GUS expression cassette in addition to the st gene fusion. The GUS reporter gene cassette was introduced at the EcoRI site immediately downstream of the st expression cassette in pBinTS25. To achieve this, the 900bp EcoRI/BamHI CaMV 35S promoter
fragment from pRok1 and the 2.2kb BamHI/EcoRI GUS-NOS fragment from pTAK1 were isolated by gel purification. Both fragments were cloned into the EcoRI site downstream of the NOS terminator in pBinTS25 in a triple ligation reaction.

Because the newly introduced GUS expression cassette could insert in either of two orientations, restriction digestions were performed to identify clones that contained the st and GUS genes in the same transcriptional orientation. We selected this orientation to ensure that antisense T7 RNAP transcripts could not be generated from the CaMV 35S promoter driving GUS expression. Restriction digestion of putative pBinTGS25 clones was performed to identify the clones containing the GUS cassette in the desired orientation (Figure 2.4). EcoRI digestion excised a 3Kb fragment (Figure 2.4, lane 2) approximately equivalent in size to the newly introduced GUS cassette. BamHI digestion yielded a 1.2Kb NOS terminator/CaMV 35S promoter fragment (Figure 2.4, lane 7) which confirmed that the NOS terminator from the st chimeric gene was located immediately adjacent to the CaMV promoter from the GUS gene (Figure 2.1).

2.3.1.2. Construction of pBinTGS8 and pBinTGS14

In order to investigate the efficiency of targeting of ST fusion proteins into chloroplasts, two additional constructs, pBinTGS8 and pBinTGS14, were made which encoded fusion proteins containing different numbers of N-terminal amino acids of the mSSU.

The construction pathways for pBinTGS8 and pBinTGS14 (Figure 2.5) were identical except that different oligonucleotide linkers (A or B) were cloned into pCAR2 in the first step in each construction. In the construction of pBinTGS8 the MscI/EcoRI fragment from pCAR2 (Figure 2.2) was replaced with the synthetic linker A formed by the annealing of two single-stranded oligonucleotides (Figure 2.6A) to give pCAR2A. The A linker, containing sequences encoding amino acids 5 - 8 of mSSU fused to codons 6 to 10 of T7 RNAP, was ligated into MscI/EcoRI digested pCAR2. This resulted in the destruction of the MscI site. The 1.4Kb HindIII/EcoRI modified rbcS fragment was excised from pCAR2A and cloned into the HindIII/EcoRI-digested pAR3132 to give pSSUAT7. Finally the 2.9Kb HindIII/HpaI rbcS-T7 RNAP fragment from pSSUAT7 was cloned into the HindIII/HpaI digested pBinTGS25 to give pBinTGS8.
The construction of pBinTGS14 was identical to pBinTGS8 except that the synthetic linker B (Figure 2.6B) was cloned into the MscI/EcoRI-digested pCAR2 to give pCAR2B which was then used in place of pCAR2A during subsequent construction steps (Figure 2.5).

At the level of restriction digest analysis, the difference between pBinTGS25 and pBinTGS8 and pBinTGS14 was the presence of an EcoRI site at the rbcS-T7 RNAP junction in the latter 2 plasmids (Figure 2.5) which was destroyed in the construction of pBinTS25/TGS25 (Figure 2.1). HindIII/EcoRI digests were performed on pBinTGS8 and pBinTGS14 to confirm the presence of the 1.4Kb HindIII/EcoRI rbcS fragment due to the extra EcoRI site (Figure 2.7). pBinTGS8 was constructed to contain codon 8 of the mature rbcS orf fused to codon 6 of T7 RNAP. However, codons 4 and 5 of T7 RNAP are also included because they are identical to codons 7 and 8 of the mature rbcS orf (Figure 2.6). pBinTGS14 contained codon 14 of the mature rbcS orf (which is the same as codon 3 of T7 RNAP) fused to codon 4 of T7 RNAP. The presence of the synthetic linkers A and B in the correct reading frame in pBinTGS8 and 14 respectively was confirmed by sequencing across the rbcS-T7 RNAP junction in each vector.
Figure 2.1. Construction of tobacco transformation vector pBinTGS25.
Figure 2.2. Nucleotide and amino acid sequence in the SSU transit peptide region. Shows nucleotides 973 to 1452 of the tobacco SSU gene (Ntss23) which contains the transit peptide region in red. This is followed by the sequence encoding the first 47 amino acids of the mature SSU which contains a 93 base intron between codons 2 and 3. The locations of the MfeI restriction site used in the construction of pBinTS25 and the MscI and EcoRI sites used in the construction of pBinTGS8 and pBinTGS14 are shown in red.
**Figure 2.3.** Restriction digestion analysis of pBinTS25. Digested plasmid DNA was analysed by electrophoresis in a 0.8% agarose gel. Lane 1, Kb ladder; lane 2, HindIII/BamHI digested pRok2 showing excision of the 900bp CaMV promoter; lane 3, undigested pBinTS25; lane 4, HindIII/BamHI digest showing excision of the 4.2Kb SSU-T7 fragment from pBinTS25.

**Figure 2.4.** Restriction digestion analysis of pBinTGS25. Lanes 1 and 2, EcoRI/HindIII and EcoR1 digestions of mavi-preped DNA of pBinTGS25 showing excision of 4.2Kb HindIII/EcoR1 SSU-T7 fragment and of the 3Kb EcoR1 GUS fragment; lane 3, Kb ladder; lanes 4 and 5, undigested and HindIII linearised mavi-preped pBinTGS25; lane 6, Kb ladder; lane 7, BamHI excision of the 1.2Kb NOS-CaMV fragment from mavi-preped pBinTGS25.
Figure 2.5. Construction of tobacco transformation vectors pBinTGS8 and pBinTGS14.
Figure 2.6. A. Nucleotide and amino acid sequence in the region of the SSU-T7 RNAP junction in pBinTGSS. Sequence for the synthetic linker A is shown in red. *Codon 10 of T7 RNAP is a GAC to GAA (Asp to Glu) mutation introduced in the construction of pAR3132 (Dunn et al., 1988). B. Shows the annealed linker B that was used to construct pBinTGS14 and the amino acid sequence detail in this region.
Figure 2.7. Restriction digestion to confirm the structure of pBinTGS8 and pBinTGS14. Lane 1, undigested pBinTGS8; lanes 2 and 3, HindIII/EcoRI excision of the 1.4Kb SSU fragment from pBinTGS8 and pBinTGS14, respectively; lane 4, HindIII/EcoRI digestion of pBinTGS25; lane 5, Kb ladder.
2.3.2. Analysis of transgenic tobacco plants expressing ST fusion proteins

2.3.2.1. Generation of transgenic tobacco plants

Nicotiana tabacum cv Samsun NN leaf sections were transformed with Agrobacterium tumefaciens LBA4404 containing binary vectors pBinTS25, pBinTGS25, pBinTGS8 and pBinTGS14 by the triparental mating method (Bevan, 1984). Transgenic shoots were selected and rooted in kanamycin-containing medium. Each construct yielded transgenic shoots with equal efficiency and many lines were planted in the greenhouse and allowed to set seed.

2.3.2.2. Western blot detection of ST fusion protein expression

In order to determine if the tobacco plants that were kanamycin resistant were expressing ST fusion proteins, total soluble leaf protein homogenates were prepared from the leaves of putative transgenic plants growing in the greenhouse. The total leaf protein extracts were resolved by SDS-PAGE and subjected to western blot analysis. Blots were probed with rabbit anti-T7 RNAP polyclonal antibodies. Figure 2.8 shows the detection of a strong signal for the ST25 fusion protein in five out of the six Nt-TGS25 transgenic lines analysed (lanes 1, 2, 4, 6 and 7). A much fainter signal was detected in Nt-TGS25-27 leaf protein extracts (lane 5). The protein was not detected in wild-type tobacco (Nt) leaf protein extracts (lanes 8 and 9) and had the same mobility as 0.8 units of purified native T7 RNAP (Boehringer Mannheim) that had been added to an Nt leaf protein extract (lane 3). The ST25 fusion protein appeared as a doublet in the leaf protein extracts of four of the transgenic lines (lanes 2, 4, 6 and 7).

Figure 2.9A shows the western detection of ST fusion protein in total leaf protein extracts from several Nt-TS25 and Nt-TGS25 transgenic lines. ST expression levels varied from plant to plant. It was expressed at the highest level in Nt-TGS25-19, 23, 24, 28 and Nt-TS25-37 (lanes 4, 7, 8, 10 and 12, respectively). Expression was lower in Nt-TGS25-14, 20 and 25 (lanes 3, 5 and 9, respectively) and barely detectable in Nt-TGS25-13, Nt-TS25-3 and Nt-TS25-40 (lanes 1, 2 and 11, respectively). Ten units of purified native T7 RNAP was loaded as a positive control (lane 6) and the signal detected migrated with approximately the same molecular mass as the ST fusion protein detected in the transgenic lines.
Figure 2.8. Immunoblotting of the ST25 fusion protein in plants transformed with pBinTGS25. Each lane contains approximately 50μg of total leaf protein extract probed with anti-T7 RNAP antibodies. Lanes 1, 2, 4, 5, 6 and 7 contain samples from Nt-TGS25-35, 33, 29, 27, 23 and 15, respectively; lane 3 contains 0.8 units of native T7 RNAP added to a total leaf protein extract prepared from wild-type tobacco (Nt); lanes 8 and 9 contain total leaf protein from Nt plants.

Figure 2.9 A and B. Immunoblotting of ST fusion proteins in transgenic tobacco. Each lane contains approximately 50μg of total leaf protein extract probed with antibodies. A. Western blotting of plants transformed with pBinTGS25 and pBinTS25. Lanes 2, 11 and 12, Nt-TS25-3, 40 and 37, respectively; lane 6, 10 units of native T7 RNAP; lanes 1, 3, 4, 5, 7, 8, 9 and 10, Nt-TGS25-13, 14, 19, 20, 23, 24, 25 and 28, respectively. B. Western blotting of plants transformed with pBinTGS8, pBinTGS14 and pBinTGS25. Lanes 1 and 2, Nt-TGS25-15 and 21; lane 3, 10 units of T7 RNAP; lanes 4 and 5, Nt-TGS14-1 and 2, respectively; lanes 6-9, Nt-TGS8-8, 9, 10 and 15, respectively.
Figure 2.9B shows the western detection of ST fusion proteins in total leaf protein extracts from Nt-TGS8, Nt-TGS14 and Nt-TGS25 transgenic lines. Expression level also varies in Nt-TGS8 and Nt-TGS14 lines. High levels of expression were detected in Nt-TGS25-15 and 21, Nt-TGS14-1 and Nt-TGS8-10 (lanes 1, 2, 4 and 8, respectively). Low levels of expression were detected in Nt-TGS14-2, Nt-TGS8-8, 9 and 15 (lanes 5, 6, 7 and 9, respectively). Again 10 units of purified native T7 RNAP was loaded as a positive control (lane 3).

The appearance of the ST25 fusion protein as a doublet in Nt-TGS25-23 leaf protein extracts extracts in Figure 2.8 and 2.9 is most probably an artefact of SDS-PAGE because it appears as a single band in Nt-TGS25-23 F1 progeny in Figure 2.15. Also in Figure 2.9A the native T7 RNAP appears as a doublet and in most western blots the ST fusion protein is detected as a single band.

2.3.2.3. T7 RNAP activity determinations in transgenic tobacco plants

In order to determine if the ST fusion proteins that were expressed in transgenic tobacco had T7 RNAP activity, in vitro transcription assays were performed on total leaf protein extracts from young plants of the Nt-TGS8, Nt-TGS14 and Nt-TGS25 transgenic lines. An initial set of experiments using extracts from non-transgenic and transgenic plants confirmed the template specificity of the T7 RNAP assay i.e. only background levels of $^{32}$P-UTP incorporation into acid precipitable counts were observed if the T7 DNA template was omitted from the in vitro transcription assay or if it was substituted with a phage lambda DNA template (data not shown). These experiments also showed that T7 DNA was not transcribed to any significant extent by the endogenous RNAPs present either in total tissue extracts or in purified chloroplasts (data not shown).

The bar chart in Figure 2.10 shows the T7 RNAP specific activity values determined for a selection of Nt-TGS25 lines that had tested positive by western blotting with anti-T7 RNAP antibodies. T7 RNAP specific activity was determined in counts per minute per microgram (CPM/μg) of total leaf protein assayed. Background activity was determined by assaying total leaf protein extracts from wild-type tobacco. The left most bar in the chart represents the CPM value determined for 1 unit of native T7 RNAP assayed as part of a wild-type tobacco leaf protein extract. All of the Nt-TGS25 lines analysed contain significant T7 RNAP activity above background and the specific activity of T7 RNAP varied from plant to plant.
The bar chart in Figure 2.11 shows the determination of T7 RNAP specific activities in the Nt-TGS8 and Nt-TGS14 lines that were tested for ST fusion protein expression by western blotting in Figure 2.9B. Nt-TGS25-23 and Nt-TGS27 were included as examples of high and low expressors respectively (Figure 2.8). The specific activity values were calculated in the same way as for the previous assay. It is clear from the bar chart that TGS25-23, TGS14-1 and TGS8-10 contain high levels of T7 RNAP activity relative to the other plants assayed. However TGS14-2, TGS8-8 and 9 contain T7 RNAP activity that is barely above background. In general, the levels of T7 RNAP activity detected in individual transgenic plants correlated positively with the levels of immunodetectable T7 RNAP protein (Figure 2.9B).

Table 2.1 contains estimations of the specific activities in unit/µg of total leaf protein for the Nt-TGS25 lines shown in Figure 2.10. The unit values were determined by comparison of CPM/µg values for the transgenic lines with the CPM/µg value determined for 1 unit of native T7 RNAP in a wild-type tobacco leaf protein extract. It was not possible to perform unit/µg estimations for the transgenic lines assayed in Figure 2.11 because the native T7 RNAP was not added to a wild-type tobacco leaf protein extract in this set of assays.

2.3.2.3. Determination of GUS activity in transgenic lines
Relative GUS activities were determined in total leaf protein extracts prepared from transgenic lines that had the GUS gene linked to the st gene by qualitative estimation of MU fluorescence intensity in GUS assays using a UV transilluminator. Table 2.2 shows the relative GUS activities present in the Nt-TGS8, Nt-TGS14 and Nt-TGS25 transgenic lines tested. Total leaf protein extracts that did not contain any visible GUS activity after incubation at 37°C for 1 hour were incubated overnight.

2.3.2.4. Organ specificity of T7 RNAP expression
Western blotting of protein extracts from leaf and root tissue was performed using T7 RNAP specific antibodies (Figure 2.12). Root and leaf extracts were prepared in GUS extraction buffer from a single plant of the TGS14-7 homozygous line. Samples were prepared for SDS-PAGE and loaded on the basis of equal protein concentration. As in previous westerns a distinct band with a molecular mass of approximately 100 Kd corresponding in size to native T7 RNAP was detected in the leaf protein extract. No corresponding band was detected in the root tissue sample.
Figure 2.10. T7 RNAP activity values determined from total soluble leaf protein extracts prepared from Nt-TGS25-15, 23, 24, 28, 29, 33 and 35 transgenic lines. Specific activity was determined in counts per minute per microgram of leaf protein assayed (CPM/ug). 1U = the CPM value determined for 1 unit of native T7 RNAP assayed in a total leaf protein extract from wild-type tobacco. Nt = CPM/ug value determined for a wild-type tobacco total leaf protein extract.
Figure 2.11. T7 RNAP activity values determined from total soluble leaf protein extracts prepared from Nt-TGS8-8, 9, 10 and 15, Nt-TGS14-1 and 2 and Nt-TGS25-23 and 27 transgenic lines. Activities are given in counts per minute per microgram of leaf protein assayed (CPM/ug). 1U = the CPM/unit value determined for native T7 RNAP. Nt = the CPM/ug value determined for wild-type tobacco total leaf protein extract.
Table 2.1. T7 RNAP specific activities in unit/µg of total leaf protein determined from the CPM/µg values for the Nt-TGS25-23 transgenic lines and native T7 RNAP in Figure 2.10.

<table>
<thead>
<tr>
<th>Transgenic lines</th>
<th>T7 RNAP specific activity (unit/µg of total leaf protein)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nt-TGS25-15</td>
<td>1.4</td>
</tr>
<tr>
<td>Nt-TGS25-23</td>
<td>0.6</td>
</tr>
<tr>
<td>Nt-TGS25-24</td>
<td>0.4</td>
</tr>
<tr>
<td>Nt-TGS25-28</td>
<td>1.1</td>
</tr>
<tr>
<td>Nt-TGS25-29</td>
<td>1.1</td>
</tr>
<tr>
<td>Nt-TGS25-33</td>
<td>1.3</td>
</tr>
<tr>
<td>Nt-TGS25-35</td>
<td>0.8</td>
</tr>
</tbody>
</table>

Table 2.2. GUS activities in Nt-TGS8, Nt-TGS14 and Nt-TGS25 transgenic lines determined by crude GUS assays. Activity ranges from very strong (++++) to strong (+++) to weak (+) to undetectable (-).

<table>
<thead>
<tr>
<th>Plant Nt-</th>
<th>GUS activity</th>
<th>Plant Nt-</th>
<th>GUS activity</th>
<th>Plant Nt-</th>
<th>GUS activity</th>
</tr>
</thead>
<tbody>
<tr>
<td>TGS8-8</td>
<td>+</td>
<td>TGS14-5</td>
<td>+</td>
<td>TGS25-23</td>
<td>+++</td>
</tr>
<tr>
<td>TGS8-9</td>
<td>+</td>
<td>TGS14-7</td>
<td>+</td>
<td>TGS25-24</td>
<td>+++</td>
</tr>
<tr>
<td>TGS8-10</td>
<td>+++</td>
<td>TGS25-13</td>
<td>-</td>
<td>TGS25-27</td>
<td>++</td>
</tr>
<tr>
<td>TGS8-12</td>
<td>+++</td>
<td>TGS25-14</td>
<td>-</td>
<td>TGS25-28</td>
<td>++</td>
</tr>
<tr>
<td>TGS8-13</td>
<td>+++</td>
<td>TGS25-15</td>
<td>+++</td>
<td>TGS25-29</td>
<td>+++</td>
</tr>
<tr>
<td>TGS8-14</td>
<td>++</td>
<td>TGS25-19</td>
<td>+++</td>
<td>TGS25-33</td>
<td>+++</td>
</tr>
<tr>
<td>TGS8-15</td>
<td>+</td>
<td>TGS25-20</td>
<td>+</td>
<td>TGS25-35</td>
<td>+++</td>
</tr>
<tr>
<td>TGS14-4</td>
<td>+++</td>
<td>TGS25-21</td>
<td>++</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
2.3.2.5. Stability of the ST fusion protein in tobacco

In order to provide an estimate of the stability of ST fusion proteins in vivo over time, we took advantage of the fact that their expression in our transgenic lines was directed by the light-regulated rbcS promoter. Transcription directed by the rbcS promoter is rapidly down-regulated when plants are transferred to continuous darkness (Giuliano et al., 1988; Fritz et al., 1991). On the assumption that this would lead to a cessation in further synthesis of ST proteins, we monitored the levels of ST protein and T7 RNAP activity remaining over time. In order to ensure that the st transgene in the line used in this analysis (Nt-TGS25-23) was exclusively under the control of the rbcS promoter T7 RNAP activities were determined in total leaf and root protein extracts (result not shown).

Seeds from the Nt-TGS25-23 F1 homozygous line were sterilised using the ethanol/bleach method, plated on MS salts medium containing 3% sucrose and 0.8% agar in petri dishes and grown in the light in a growth room for 25 days. Individual petri dishes were then wrapped in aluminium foil to exclude light. Seeds were plated at different times so that all seedlings would spend the same amount of time (25 days) growing under normal light conditions but be kept in the dark (D) for different time periods (3 - 16 days). In this way all seedling treatments were analysed on the same day to determine ST levels and T7 RNAP activity. Ten seedlings from a single petri dish were processed together for each time point.

Figure 2.13 shows a western blot analysis of total protein extracts prepared from the TGS25-23 F1 seedlings at D0, D1, upto D16. Seedlings at all time points contained detectable levels of ST protein. However, over the course of the 16 day period in the dark there is a noticeable decline in ST levels whereas the concentration of Rubisco LSU which appeared as a 57 Kd background band remained constant over the 16 day period.

T7 RNAP activity assays were also carried out on each total seedling protein extract from D0 to D16. Figure 2.14 shows a bar chart representing the T7 RNAP activities present in Nt-TGS25-23 F1 seedlings over the 16 day period in the dark. T7 RNAP activity was calculated in cpm/μg of total protein assayed. These data confirm what was found by western blot analysis. T7 RNAP activity persists in the seedlings at a high level over the 16 day period. After 16 days in the dark approximately 50% of the activity present on day 0 can still be detected.
Figure 2.12. Immunoblott developed by anti-T7 RNAP on crude leaf and root protein extracts prepared from Nt-TGS14-7 resolved on a 6.75% SDS-PAGE separating gel. Lanes 1 and 2, approximately 20μg of protein extract from leaf and root, respectively; lane 3, high molecular weight marker visualised with Ponceau S.

Figure 2.13. Immunoblotting of ST25 in total protein extracts from Nt-TGS25-23 F₁ seedlings. Each lane contains approximately 20μg of total protein. Lanes 1, 20 units of native T7 RNAP; lane 2, Nt; lanes 3 to 9, protein extracts from Nt-TGS25-23 F₁ seedlings at Day 0 (D₀), D₁, D₃, D₅, D₇, D₉, and D₁₆, respectively.
Figure 2.14. T7 RNAP activities present in Nt-TGS25-23 F1 seedlings over the course of a 16 day period in the dark. CPM values were measured using liquid scintillation counting. The left most bar represents the CPM value determined for 10 units of native T7 RNAP. Nt = wild-type tobacco.
2.3.2.6. Inheritance of transgenes

2.3.2.6.1. Inheritance of the \textit{st} genes in the F\textsubscript{1} progeny of selfed primary transformants

Inheritance of \textit{st} genes in the F\textsubscript{1} progeny of selfed primary transformants was demonstrated by western blot detection of ST proteins in total soluble leaf protein extracts. Figure 2.15 shows a western blot analysis of total soluble leaf protein samples from ten Nt-TGS25 F\textsubscript{1} seedlings. ST protein expression was inherited by five of the ten seedlings analysed.

Figure 2.16 shows the ST fusion protein levels in total leaf proteins from three Nt-TGS14-4 F\textsubscript{1} plants, three Nt-TGS8-12 F\textsubscript{1} plants, two Nt-TGS14-7 F\textsubscript{1} plants and two Nt-TGS8-13 F\textsubscript{1} plants. F\textsubscript{1} plants selected for analysis were GUS positive and were grown in the green house in separate pots under the same conditions. It is clear from the western blot that different levels of ST expression was inherited in the progeny of Nt-TGS8-12 (lanes 4, 5 and 6) and Nt-TGS14-7 (lanes 7 and 8) transgenic lines.

2.3.2.6.2. Inheritance of kanamycin resistance in the F\textsubscript{1} progeny of selfed primary (F\textsubscript{0}) transformants

The concentration of kanamycin required to prevent growth of non-transgenic seedlings was determined by germinating wild-type \textit{N. tabacum} var. Samsun seeds on MS medium containing 0, 100, 200, 300, and 500 mg/L kanamycin. All seedlings grew normally except those germinated on 500 mg/L kanamycin (kan500). Seedlings growing in the absence of kanamycin were green and developed to the 4-6 leaf stage. Seedlings germinated in the presence of kan500 were smaller, bleached and did not develop beyond the two leaf stage.

In order to determine the transgene copy number and to identify homozygous lines the inheritance pattern of the kanamycin resistance marker was determined by plating seeds from selfed transgenic plants on MS medium containing Kan500. Table 2.3 shows the data from these seed plating experiments which includes; the number of seeds plated, the number that germinated, and the number that were kanamycin resistant for each plant tested. The percentage of germinated seedlings that were kanamycin resistant was calculated in each case. Seed progeny of selfed F\textsubscript{0} plants with percentage kanamycin resistance greater than 95\% and less than 100\% indicate that the parent line had a transgene copy number greater than 1. Those with approximately 75\% kanamycin resistance
Figure 2.15. Immunoblot showing segregation of ST25 expression in the F1 progeny of Nt-TGS25-23. Lanes 1 to 10, 50μg each of total leaf protein from progeny plants.

Figure 2.16. ST fusion protein levels in the leaves of the seed progeny of Nt-TGS14-4 F1 plants, Nt-TGS8-12 F1 plants, Nt-TGS7-14 F1 plants and Nt-TGS8-13 F1 plants. All total soluble leaf protein samples contain approximately 25μg of total leaf protein. Lanes 1, 2 and 3, Nt-TGS14-4 F1, 4, 5 and 6, Nt-TGS8-12 F1, 5, 9 and 2 plants; lanes 7 and 3, Nt-TGS14-7 F1 and 2 plants; lanes 9 and 10, Nt-TGS8-13 F1, 5 and 7 plants.
Table 2.3. Inheritance of kanamycin resistance in the progeny of primary transformants (F₀) and F₁ generation. Percentage resistance is the percentage of germinated seedlings that were kanamycin resistant.

<table>
<thead>
<tr>
<th>Parent</th>
<th>Number of seeds</th>
<th>Number germinated</th>
<th>Kanamycin resistant</th>
<th>Percentage resistance</th>
</tr>
</thead>
<tbody>
<tr>
<td>TGS8-12 F₀</td>
<td>100</td>
<td>87</td>
<td>85</td>
<td>97.7%</td>
</tr>
<tr>
<td>TGS8-13 F₀</td>
<td>82</td>
<td>63</td>
<td>60</td>
<td>95.3%</td>
</tr>
<tr>
<td>TGS14-4 F₀</td>
<td>157</td>
<td>157</td>
<td>150</td>
<td>95.5%</td>
</tr>
<tr>
<td>TGS14-7 F₀</td>
<td>93</td>
<td>93</td>
<td>66</td>
<td>71%</td>
</tr>
<tr>
<td>TGS25-23 F₀</td>
<td>111</td>
<td>110</td>
<td>79</td>
<td>72%</td>
</tr>
<tr>
<td>TGS25-23 F₁E</td>
<td>133</td>
<td>130</td>
<td>99</td>
<td>76%</td>
</tr>
<tr>
<td>TGS25-23 F₁G</td>
<td>121</td>
<td>120</td>
<td>31</td>
<td>74%</td>
</tr>
<tr>
<td>TGS25-23 F₁I</td>
<td>60</td>
<td>58</td>
<td>58</td>
<td>100%</td>
</tr>
<tr>
<td>TGS25-28 F₁C</td>
<td>150</td>
<td>123</td>
<td>123</td>
<td>100%</td>
</tr>
</tbody>
</table>

Table 2.4. GUS specific activities in the progeny of Nt-TGS8 and Nt-TGS14 transformants. GUS specific activity was determined in nmoles MU/mg/min using fluorimeter readings.

<table>
<thead>
<tr>
<th>Leaf protein extract</th>
<th>nmoles MU/mg/min</th>
<th>Leaf protein extract</th>
<th>nmoles MU/mg/min</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nt Samsun</td>
<td>0.2</td>
<td>TGS14-4 F₁3</td>
<td>2.3</td>
</tr>
<tr>
<td>TGS8-12 F₁2</td>
<td>0.9</td>
<td>TGS14-4 F₁6</td>
<td>1.7</td>
</tr>
<tr>
<td>TGS8-12 F₁5</td>
<td>3.8</td>
<td>TGS14-4 F₁8</td>
<td>5.4</td>
</tr>
<tr>
<td>TGS8-12 F₁9</td>
<td>3.1</td>
<td>TGS14-4 F₁9</td>
<td>3.9</td>
</tr>
<tr>
<td>TGS8-12 F₁10</td>
<td>0.6</td>
<td>TGS14-7 F₁2</td>
<td>1.2</td>
</tr>
<tr>
<td>TGS8-13 F₁5</td>
<td>3.2</td>
<td>TGS14-7 F₁7</td>
<td>2.8</td>
</tr>
<tr>
<td>TGS8-13 F₁7</td>
<td>1.6</td>
<td>TGS14-7 F₁9</td>
<td>1.2</td>
</tr>
<tr>
<td>TGS8-13 F₁8</td>
<td>3.3</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
indicate that the parent line contained a single transgene and hence the 3:1 inheritance pattern when selfed. Homozygous lines were identified by 100% kanamycin resistance in the seed progeny of the F1 line when the F0 generation resulted from a single transgene insertion event.

2.3.2.6.3. Inheritance of GUS activity in the seed progeny of selfed primary (F0) transformants

GUS activities, measured using a fluorimeter, were determined from total soluble leaf protein extracts prepared from the progeny of primary transformants. All plants assayed were growing in separate pots under the same conditions in the greenhouse. GUS activities were determined for a number of the progeny of Nt-TGS8-12, Nt-TGS8-13, Nt-TGS14-4 and Nt-TGS14-7 transgenic lines. All plants assayed had previously tested positive for GUS expression in crude assays and the level of immunodetectable ST fusion protein had been previously determined for a number of these plants (Figure 2.16). Table 2.4 shows the GUS activities present in the F1 plants assayed. Different expression levels of GUS activity were inherited in the F0 seed progeny. The levels of GUS activity correlate with the level of ST fusion protein expression as determined for a subset of these plants (Figure 2.16). The exception is Nt-TGS14-4 F1 8 which had the highest GUS activity but had a lower ST expression level than Nt-TGS14-4 F1 3.

2.3.2.7. Sub-cellular location of ST fusion proteins

2.3.2.7.1. Immunodetection of ST fusion proteins in untreated and thermolysin treated gradient-purified intact chloroplasts

ST fusion protein levels and T7 RNAP activity was determined using Percoll gradient-purified intact chloroplasts that were prepared from the leaf tissue of healthy transgenic plants. Figures 2.17B, 2.18B, and 2.19B show the detection of ST in total leaf protein extracts and in untreated, thermolysin- and thermolysin/Triton-X100- treated intact chloroplasts for a number of Nt-TGS8, Nt-TGS14 and Nt-TS25/TGS25 transgenic plants. ST levels in total leaf protein and intact chloroplast samples were quantified by densitometry of the approximately 100 kD ST signal detected by western blots analysis (Figures 2.17B, 2.18B and 2.19B). The densitometry readings were adjusted to reflect ST levels in LSU-equivalent samples of total leaf protein and intact chloroplast (Figures 2.17C, 2.18C and 2.19C). The adjustment factor required to make protein extracts prepared from intact chloroplasts equivalent to their corresponding total leaf protein extracts was determined from densitometry.
readings of LSU levels in Coomassie Blue-stained protein gels (Figures 2.17A, 2.18A and 2.19A).

Figure 2.17 shows the protein gel and western blot that were analysed by densitometry to determine the levels of ST fusion protein in LSU-equivalent total leaf protein and intact chloroplast samples from Nt-TGS8-12, Nt-TGS14-4 and Nt-TGS25-23 transgenic plants. Figure 2.17C shows that the levels of ST associated with intact chloroplasts was typically higher (by a factor of 1.75 - 2; see Table 2.6) than that detected in the corresponding LSU-equivalent total leaf protein extracts. A number of minor species with higher and lower molecular weights than the ST8 protein were detected in the total leaf protein extract from Nt-TGS8-12 (Figure 2.17B, lane 1). A subset of the minor bands were also detected in the corresponding intact chloroplast preparation (Figure 2.17B, lane 5). The lower molecular weight species probably represent degradation products.

Figure 2.18 shows the protein gel and western blot that were analysed by densitometry to determine the levels of ST protein in LSU-equivalent total leaf protein extracts and in protein extracts from untreated- and thermolysin- treated intact chloroplast samples from Nt-TGS8-14, Nt-TGS14-8 and Nt-TG25-39 transgenic plants. Again the untreated intact chloroplasts contained more ST protein than the corresponding LSU-equivalent total leaf protein extract. The level of ST associated with intact chloroplasts was significantly reduced (particularly in intact chloroplasts from Nt-TGS8-14 and Nt-TGS14-8) following treatment with thermolysin. In contrast, thermolysin treatment of intact chloroplasts had little effect on the levels of LSU. Quantitative estimation of the levels of ST associated with thermolysin-treated intact chloroplasts (Figure 2.18C) was made by adjusting densitometry values by the same factor that was used to relate the levels of ST protein in untreated intact chloroplasts to those found in total leaf protein on the basis of LSU-equivalence.

Figure 2.19 shows repeated determinations of the levels of ST in total leaf protein extracts and in purified intact chloroplast samples of Nt-TGS14-4 and Nt-TGS25-23. In contrast to Figure 2.17 the levels of fusion protein in untreated intact chloroplasts was lower than that in total leaf protein extracts and in this case it was not significantly reduced by thermolysin treatment. Intact chloroplasts treated with thermolysin in the presence of 0.1% Triton X-100 contained no detectable fusion protein (lanes 4 and 8). Also, the protein gel clearly shows the protection of Rubisco LSU in thermolysin-treated intact chloroplasts and the
Figure 2.17 A, B and C. ST fusion protein levels in total leaf protein extracts and their corresponding intact chloroplast preparations from Nt-TGS8-12, Nt-TGS14-4 and Nt-TGS25-23 transgenic plants. A. 6.75% polyacrylamide gel. Lanes 1, 2 and 3, total tissue extracts from Nt-TGS8-12, Nt-TGS14-4 and Nt-TGS25-23, respectively; lane 4, protein marker; lanes 5, 6 and 7, purified intact chloroplasts prepared from Nt-TGS8-12, Nt-TGS14-4 and Nt-TGS25-23, respectively. The chlorophyll content in each total leaf protein and untreated intact chloroplast sample is shown and the LSU ratios determined by densitometry for each total leaf protein extract and its corresponding intact chloroplast sample are also shown. B, Immunoblot of duplicate samples probed with anti-T7 RNA polymerase. Lane identifications are as described for A. C, ST fusion protein densitometry readings adjusted to reflect ST fusion protein levels in LSU equivalent total leaf proteins (T) and intact chloroplasts (I) from Nt-TGS8-12 (red), Nt-TGS14-4 (yellow) and Nt-TGS25-23 (blue).
Figure 2.18 A, B and C. ST fusion protein levels in total leaf protein extracts, purified intact chloroplasts and thermolysin treated purified intact chloroplasts. A. SDS-PAGE. Lanes 1, 2 and 3, total leaf protein extract, intact purified chloroplasts and thermolysin treated intact chloroplasts, respectively, from Nt-TGS8-14; lanes 5, 6 and 7, same samples from Nt-TGS14-8; lanes 8, 9 and 10, same samples from Nt- TS25-39. Lane 4, 10 units of native T7 RNAP. The chlorophyll content in each total leaf protein and untreated intact chloroplast sample is shown and the LSU ratios determined by densitometry for each total leaf protein extract and its corresponding intact chloroplast samples are also shown. B. Immunoblot of duplicate SDS-PAGE. Lane identifications are as described for A. C. ST fusion protein densitometry readings adjusted to reflect ST fusion protein levels in LSU equivalent total leaf proteins (T), intact chloroplasts (I) and thermolysin treated intact chloroplasts (IT) from Nt-TGS8-14 (red), Nt-TGS14-8 (yellow) and Nt-TS25-39 (blue).
Figure 2.19 A, B and C. ST fusion protein levels in total leaf protein extracts, purified intact chloroplasts and thermolysin treated purified intact chloroplasts. A. SDS-PAGE. Lanes 1 to 4, total leaf protein extract, intact purified chloroplasts, thermolysin treated intact chloroplasts and thermolysin/triton treated intact chloroplasts, respectively, from Nt-TGS14-4; lanes 5 to 8, same samples from Nt-TGS25-23. The chlorophyll content in each total leaf protein and untreated intact chloroplast sample is shown and the LSU ratios determined by densitometry are also shown. B. Immunoblot of duplicate SDS-PAGE. Lane identifications are as described for A. C. ST fusion protein densitometry readings adjusted to reflect ST fusion protein levels in LSU equivalent total leaf proteins (T), intact chloroplasts (I) and thermolysin treated intact chloroplasts (IT) from Nt-TGS14-4 (yellow) and Nt-TGS25-23 (blue).
degradation of LSU in thermolysin/Triton X-100- treated intact chloroplasts (lanes 3 and 4, and 7 and 8).

2.3.2.7.2. Detection of T7 RNAP activity in untreated and in thermolysin-treated gradient-purified intact chloroplasts

Figure 2.20 shows the determination of T7 RNAP activity in total leaf protein extracts and in extracts prepared from untreated, thermolysin- and thermolysin/Triton X-100-treated intact chloroplasts for a number of Nt-TGS8, Nt-TGS14 and Nt-TS25/TGS25 transgenic plants. Initially, activities were calculated in cpm/μg of chlorophyll. These values were then reconverted to cpm on the basis of LSU-equivalence of total leaf protein and purified chloroplast preparations. The conversion factor was determined by densitometric quantitation of the LSU band in SDS-PAGEs containing total leaf protein and intact chloroplast samples.

Figure 2.20A shows the T7 RNAP activities in the same protein preparations that were analysed for ST fusion protein levels in Figure 2.18. In Figure 2.20B the total leaf protein and intact chloroplast preparations were made from the same plants that were assayed in Figure 2.19. All transgenic plants assayed contained T7 RNAP activity in thermolysin-treated and untreated intact chloroplasts. In most cases the T7 RNAP activity was reduced in thermolysin-treated intact chloroplasts and virtually no T7 RNAP activity was detected in intact chloroplasts treated with thermolysin/Triton X-100. When intact chloroplast preparations and their corresponding LSU-equivalent total leaf protein extracts were compared, the disparity in the levels of T7 RNAP activity detected was much greater than the disparity in the levels of ST fusion protein detected by western blotting (compare Figures 2.18, 2.19 and 2.20).

Table 2.5 shows the average ratio of ST fusion proteins in thermolysin-treated intact chloroplasts to that found in total leaf protein extracts (IT/T) for ST8, ST14 and ST25 calculated from the levels of fusion protein determined by densitometry in Figures 2.18C and 2.19C. The Table shows that on average the amount of fusion protein associated with thermolysin treated intact chloroplasts was 0.95, 0.88 and 0.58 times that detected in LSU equivalent total leaf protein extracts from plants expressing ST8, ST14 and ST25, respectively. Table 2.6 shows the average ratio of ST fusion protein in untreated intact chloroplasts to that in the total leaf protein extract (I/T) for ST8, ST14 and ST25 calculated from the ST fusion protein levels in LSU equivalent samples (Figures 2.17C, 2.18C and 2.19C). This shows that the average amount of fusion protein associated with
intact chloroplasts was 1.84, 1.52 and 1.26 times that detected in LSU equivalent total leaf protein extracts from plants expressing ST8, ST14 and ST25, respectively. Table 2.7 shows the average I/T ratios for ST fusion protein activity values in Figure 2.20. This shows that on average the amount of T7 RNAP activity associated with intact chloroplasts was 5.6, 6.22 and 5.1 times that detected in LSU equivalent total leaf protein extracts from plants expressing ST8, ST14 and ST25, respectively. Table 2.8 shows the average IT/T ratios for the ST fusion protein activity values in Figure 2.20. This shows that on average the amount of T7 RNAP activity associated with intact chloroplasts was 5.6, 3.3 and 4.1 times that detected in LSU equivalent total leaf protein extracts from plants expressing ST8, ST14 and ST25, respectively.

2.3.2.7.3. Chlorophyll content of total leaf protein and intact chloroplast preparations

The chlorophyll content of total leaf protein extracts and intact chloroplast preparations that were LSU-equivalent (as determined by densitometry of Comassie-stained SDS-PAGs) was not equal. Table 2.9 shows the set of ratios of chlorophyll content in intact chloroplasts to that in total leaf protein extracts that were analysed in Figures 2.17, 2.18 and 2.19. The ratio value ranges from 1.3 to 7. In other words, in LSU-equivalent total leaf protein and intact chloroplast samples the intact chloroplast preparation contained between 1.3 and 7 times the quantity of chlorophyll that was present in the total leaf protein extract.

The protein gel in Figure 2.21 shows gradient-purified intact chloroplasts that contained virtually no Rubisco LSU. The total leaf protein extract and intact chloroplasts were prepared in the usual way from the same leaf of the Nt-TGS25-28 F1 C transgenic plant. Three aliquots of the total leaf protein extract were loaded in lanes 1, 2 and 3 containing 5, 2 and 1 μg of chlorophyll respectively. An aliquot of intact chloroplasts containing 6 mg of chlorophyll was loaded in lane 4. The LSU level in the intact chloroplast sample was much lower than that present even in the total leaf protein sample that contained 1 mg of chlorophyll.

2.3.2.7.4. Rubisco LSU and ST fusion protein levels in intact and broken chloroplasts

Rubisco LSU and ST fusion protein levels were determined in intact and broken chloroplast suspensions by SDS-PAGE and western blotting. Intact and broken chloroplasts were prepared for the Nt-TGS8-12 primary transformant in Percoll step gradients and were washed twice in SH buffer. Figure 2.22A shows a total
Figure 2.20 A and B. T7 RNAP activity determinations for total leaf protein extracts, intact chloroplast and thermolysin treated intact chloroplast preparations that were analysed for ST fusion protein levels in Figures 2.18 and 2.19. For each plant analysed the activity values in CPM/µg of chlorophyll were adjusted to give activity values from samples that were LSU equivalent before thermolysin treatment. The adjustment factor was determined by measuring chlorophyll concentrations and by quantitating LSU levels in SDS-PAGE using densitometry. A. T7 RNAP activities in total leaf protein (T), untreated intact chloroplasts (I) and thermolysin treated intact chloroplasts (IT) from Nt-TGS8-14, Nt-TGS14-8 and Nt-TS25-39. B. T7 RNAP activity in T, I and ITT from Nt-TGS14-4 and Nt-TS25-39. Activity in intact chloroplasts treated with triton and thermolysin was equal to background activity in total leaf protein from wild-type tobacco and is not shown in the graphs.
Table 2.5. The ratio of ST fusion protein in thermolysin treated intact chloroplasts to that in the total leaf protein extract (IT/T) for ST8, ST14 and ST25 calculated from the adjusted densitometry readings in Figures 2.18C and 2.19C.

<table>
<thead>
<tr>
<th>Transgenic line Nt-</th>
<th>ST8 IT/T</th>
<th>ST14 IT/T</th>
<th>ST25 IT/T</th>
</tr>
</thead>
<tbody>
<tr>
<td>TGS8-14 (Figure 2.18)</td>
<td>0.95</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TGS14-8 (Figure 2.18)</td>
<td></td>
<td>1.13</td>
<td></td>
</tr>
<tr>
<td>TS25-39 (Figure 2.18)</td>
<td></td>
<td></td>
<td>0.75</td>
</tr>
<tr>
<td>TGS14-4 (Figure 2.19)</td>
<td></td>
<td>0.64</td>
<td></td>
</tr>
<tr>
<td>TGS25-23 (Figure 2.19)</td>
<td></td>
<td>0.42</td>
<td></td>
</tr>
<tr>
<td><strong>Average ratio</strong></td>
<td><strong>0.95</strong></td>
<td><strong>0.88</strong></td>
<td><strong>0.58</strong></td>
</tr>
</tbody>
</table>

Table 2.6. The ratio of ST fusion protein in untreated intact chloroplasts to that in the total leaf protein extract (I/T) for ST8, ST14 and ST25 calculated from the adjusted densitometry readings in Figures 2.17C, 2.18C and 2.19C.

<table>
<thead>
<tr>
<th>Transgenic line Nt-</th>
<th>ST8 I/T</th>
<th>ST14 I/T</th>
<th>ST25 I/T</th>
</tr>
</thead>
<tbody>
<tr>
<td>TGS8-12 (Figure 2.17)</td>
<td>1.75</td>
<td>1.74</td>
<td>2.07</td>
</tr>
<tr>
<td>TGS14-4 (Figure 2.17)</td>
<td>1.93</td>
<td>2.29</td>
<td>1.16</td>
</tr>
<tr>
<td>TGS25-23 (Figure 2.17)</td>
<td></td>
<td>0.54</td>
<td>0.54</td>
</tr>
<tr>
<td>TGS8-14 (Figure 2.18)</td>
<td>1.93</td>
<td>2.29</td>
<td></td>
</tr>
<tr>
<td>TGS14-8 (Figure 2.18)</td>
<td></td>
<td>0.54</td>
<td>0.54</td>
</tr>
<tr>
<td>TS25-39 (Figure 2.18)</td>
<td>1.93</td>
<td>2.29</td>
<td>1.16</td>
</tr>
<tr>
<td>TGS14-4 (Figure 2.19)</td>
<td>1.52</td>
<td>1.52</td>
<td>1.26</td>
</tr>
<tr>
<td>TGS25-23 (Figure 2.19)</td>
<td></td>
<td>0.54</td>
<td>0.54</td>
</tr>
<tr>
<td><strong>Average ratio</strong></td>
<td><strong>1.84</strong></td>
<td><strong>1.52</strong></td>
<td><strong>1.26</strong></td>
</tr>
</tbody>
</table>
Table 2.7. The ratio of ST fusion protein activity in untreated intact chloroplasts to that in the total leaf protein extract (I/T) for ST8, ST14 and ST25 calculated from the T7 RNAP activities in Figure 2.20.

<table>
<thead>
<tr>
<th>Transgenic line Nt-</th>
<th>ST8 I/T</th>
<th>ST14 I/T</th>
<th>ST25 I/T</th>
</tr>
</thead>
<tbody>
<tr>
<td>TGS8-14 (Figure 2.20)</td>
<td>5.6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TGS14-8 (Figure 2.20)</td>
<td></td>
<td>5.44</td>
<td>7.00</td>
</tr>
<tr>
<td>TS25-29 (Figure 2.20)</td>
<td></td>
<td></td>
<td>3.2</td>
</tr>
<tr>
<td>TGS14-4 (Figure 2.20)</td>
<td></td>
<td>7.00</td>
<td></td>
</tr>
<tr>
<td>TGS25-23 (Figure 2.20)</td>
<td></td>
<td></td>
<td>3.2</td>
</tr>
<tr>
<td><strong>Average ratio</strong></td>
<td>5.6</td>
<td>6.22</td>
<td>5.1</td>
</tr>
</tbody>
</table>

Table 2.8. The ratio of ST fusion protein activity in thermolysin treated intact chloroplasts to that in the total leaf protein extract (IT/T) for ST8, ST14 and ST25 calculated from the T7 RNAP activities in Figure 2.20.

<table>
<thead>
<tr>
<th>Transgenic line Nt-</th>
<th>ST8 IT/T</th>
<th>ST14 IT/T</th>
<th>ST25 IT/T</th>
</tr>
</thead>
<tbody>
<tr>
<td>TGS8-14 (Figure 2.20)</td>
<td>5.6</td>
<td>3.4</td>
<td>5.1</td>
</tr>
<tr>
<td>TGS14-8 (Figure 2.20)</td>
<td></td>
<td>3.2</td>
<td></td>
</tr>
<tr>
<td>TS25-29 (Figure 2.20)</td>
<td></td>
<td>3.2</td>
<td></td>
</tr>
<tr>
<td>TGS14-4 (Figure 2.20)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TGS25-23 (Figure 2.20)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Average ratio</strong></td>
<td>5.6</td>
<td>3.3</td>
<td>4.1</td>
</tr>
</tbody>
</table>
Table 2.9. Intact chloroplast/total leaf protein (I/T) chlorophyll ratio determined for LSU equivalent total leaf protein and intact chloroplast preparations assayed in Figures 2.17, 2.18 and 2.19.

<table>
<thead>
<tr>
<th>Transgenic line Nt-</th>
<th>I/T Chlorophyll ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>TGS8-12 (Figure 2.17)</td>
<td>3</td>
</tr>
<tr>
<td>TGS8-14 (Figure 2.18)</td>
<td>7</td>
</tr>
<tr>
<td>TGS14-4 (Figure 2.17)</td>
<td>1.3</td>
</tr>
<tr>
<td>TGS14-4 (Figure 2.19)</td>
<td>2.9</td>
</tr>
<tr>
<td>TGS14-8 (Figure 2.18)</td>
<td>3.4</td>
</tr>
<tr>
<td>TGS25-23 (Figure 2.17)</td>
<td>2.1</td>
</tr>
<tr>
<td>TGS25-23 (Figure 2.19)</td>
<td>2.4</td>
</tr>
<tr>
<td>TS25-39 (Figure 2.18)</td>
<td>1.7</td>
</tr>
</tbody>
</table>
leaf protein extract from Nt-TGS8-12 alongside samples of intact and broken chloroplast preparations (lanes 1, 2 and 3, respectively). The intact and broken chloroplasts contained equal concentrations of chlorophyll. The broken chloroplasts contained very little Rubisco LSU compared to the intact chloroplast sample.

Figure 2.22B shows the western blot detection of the ST8 fusion protein in the total leaf protein extract and the intact and broken chloroplast fractions that were shown in Figure 2.22A. The ST8 fusion protein was detected in all three tissue preparations. A significant ST8 signal was detected in the washed broken chloroplasts and densitometry readings indicated that as much as 20% of the fusion protein detected in the intact chloroplasts was also associated with washed broken chloroplasts.

2.3.2.8. A comparative northern blot analysis of chloroplast transcription in non-transgenic and transgenic tobacco expressing ST fusion proteins.

Following the determination that plants expressing ST fusion proteins contain a significant proportion of their total cellular T7 RNAP activity in chloroplasts, we decided to investigate whether its presence had any impact on chloroplast transcription. In order to achieve this we designed oligonucleotides for the amplification of plastid DNA fragments to use as radio-labelled probes in a northern blot analysis of plastid transcripts in transgenic and wild-type tobacco plants. To increase the likelihood of identifying transcriptional differences due to the presence of T7 RNAP activity we decided to analyse the transcription of plastid genes that occurred in the region of T7 promoter-like sequences found in the plastid genome of tobacco

2.3.2.8.1. Identification of T7 promoter-like sequences in the plastid genome of tobacco.

Given that the consensus T7 promoter is a 23bp sequence it is highly unlikely that a sequence with significant homology would be present in the 155kb tobacco plastid genome. Indeed, it is unlikely that plastid DNA would even contain a single copy of any given 9 nucleotide sequence since such a sequence would occur by chance once in a 260,000bp random nucleotide sequence. Therefore, we decided to search the published tobacco plastid genome sequence (Shinozaki et al., 1986) for sequences that showed homology to the 13bp polymerase binding domain element of the T7 promoter (nucleotides -17 to -5). We focussed on the binding domain because many of the substitutions in this domain that do not
Figure 2.21. Polyacrylamide gel showing the loss of LSU from a purified intact chloroplast preparation from Nt-TGS25-28 F1 C. Lanes 1, 2 and 3, total leaf protein samples containing 5, 2 and 1 µg of chlorophyll, respectively; lane 4, blank; lane 5, intact chloroplast sample containing 6 µg of chlorophyll.

Figure 2.22 A and B. Ribulose LSU and ST fusion protein levels in total protein, intact chloroplast and broken chloroplast preparations from Nt-TGS6-12. A. 6.75% polyacrylamide gel. Lane 1, total protein extract containing 4 µg of chlorophyll; lanes 2 and 3, intact and broken chloroplast preparations containing 15 µg of chlorophyll each. B. Nitrocellulose of duplicate samples probed with anti-T7 RNAP antibodies. Lane identifications are as for A.
abolish promoter activity have been identified empirically by *in vitro* T7 RNAP transcription assays (Chapman and Burgess, 1987; Diaz *et al.*, 1993). Also, it has been determined *in vitro* that the minimal binding domain (-15 to -5) on its own is sufficient for tight binding of T7 RNAP (Ujvari and Martin, 1997). Therefore if such domains exist in the plastid genome, they may function as efficient T7 RNAP binding sites even if binding doesn’t result in transcription of the template.

Using the FINDPATTERNS program (GCG Software suite; University of Wisconsin) and allowing for mismatches, we identified four T7 promoter-like sequences (P1-4) in plastid DNA that contained significant homology to the binding domain of the T7 promoter (Table 2.10). P1-4 contain 4, 5, 3 and 4 mismatches from the consensus T7 promoter in the binding domain region (-17 to -5), respectively. In respect of the individual mismatches that occur in the binding domain regions of P1-4 between -15 and -5, each has been shown to be permissible in *in vitro* T7 RNAP transcription assays (Chapman and Burgess, 1987; Diaz *et al.*, 1993). However, each of these permissible substitutions reduces promoter activity to between 20 and 70% that of wild-type and combinations of more than one in a single promoter may totally abolish activity. All other mismatches in the binding domain regions of P1-4 (except the C at -16 in P3) are present in the class II natural variants of the T7 promoter (Dunn and Studier, 1983). In the initiation domain regions of P1 and P3 there is 70% homology to the 23bp consensus T7 promoter if bases that occur in class II promoter variants are also included in the comparison.

Figure 2.23 shows the locations of P1-4 in linear maps of the relevant segments of the tobacco plastid genome and the predicted direction of any transcription that may occur from them. P1, P2 and P4 are located within the *rpoB, clpP* and *ndhH* genes, respectively. *RpoC1* is transcribed as part of the *rpoB/C1/C2* operon and therefore its transcription was analysed in order to determine whether the possible binding of T7 RNAP to P1 in *rpoB* might interfere with normal transcription. *PsbD* is located approximately 10kb downstream of P1 and therefore the *psbD* probe was employed as a control probe. The *clpP* probe was designed to test for any interference with *clpP* transcription caused by the possible binding of T7 RNAP to P2 and to test for any new transcripts that might be synthesised from P2. P3 is located in an intergenic region between *orf70B* and 3'-*rps12* and the transcription of the 3'-*rps12* gene downstream of P3 and the 16S rDNA gene upstream was analysed. P4 is located within the *ndhH* gene and therefore *ndhH* and *ndhA* transcription in the *ndhHAJ* operon was analysed.
For northern blot analysis, total RNA was prepared from the leaves of healthy Nt-TGS25-23 and wild-type tobacco plants that were growing in the greenhouse. Figure 2.24 shows the transcripts detected by autoradiography for each of the plastid genes analysed. Below each autoradiograph is a photograph of the corresponding ethidium bromide-stained RNA gel which shows the concentration of total RNA loaded in each lane. As expected multiple transcripts were detected for most of the plastid genes analysed. In most cases the northern blots showed that the concentration of transcripts in Nt-TGS25-23 leaves was higher than wild-type but this is most likely a loading effect. However, there appeared to be no differences between wild-type and Nt-TGS25-23 transgenic plants either in the number, size or relative concentration of transcripts detected by the various plastid gene probes.
Table 2.10. T7 promoter-like sequences (P1-4) in tobacco ptDNA with good homology to the binding domain (-17 to -5) of the T7 promoter. The sequence corresponding to the initiation domain is also shown. Deviations in P1-4 from the consensus T7 promoter sequence are underlined and the location of P1-4 in ptDNA is given by the position of the 5' nucleotide (Shinozaki et al., 1986).

<table>
<thead>
<tr>
<th>Promoter</th>
<th>Sequence</th>
<th>5' nucleotide position in ptDNA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Consensus T7 promoter</td>
<td>-17 -5 +6 TAATACGACTCAC TATAGGGAGA</td>
<td></td>
</tr>
<tr>
<td>P1</td>
<td>TATTAAGAACTCAT TAAAGCTCGA</td>
<td>25860</td>
</tr>
<tr>
<td>P2</td>
<td>CTATCAGACTCAG ATGGGTAAAT</td>
<td>72873</td>
</tr>
<tr>
<td>P3</td>
<td>TCATCAGACTCAC TAAAGTCGGA</td>
<td>101844(C)/140685</td>
</tr>
<tr>
<td>P4</td>
<td>CTATCCGACTCAA CTCCACACATA</td>
<td>124690</td>
</tr>
</tbody>
</table>
Figure 2.23. Linear maps of segments of the tobacco plastid genome showing the location of P1-4 and the direction of any T7 RNAP-mediated transcription that may occur from them. The genes on the upperside of the line are transcribed from left to right, and those on the lower side, from right to left. P identifies the regions of the plastid genes that were probed in the northern analysis. The positions in ptDNA of P1-4 (5' nucleotide) and the extremities of the fragments are shown.
Figure 2.24. Northern analysis of the transcription of several chloroplast genes in wild-type (wt) tobacco and N-TCS25-23 leaves. Below each northern is the corresponding ethidium stained RNA gel showing the total leaf RNA loaded in each lane.
2.4. DISCUSSION

Three chimeric genes encoding ST fusion protein were constructed in binary vectors under the control of the SSU promoter and were introduced into the nuclear genome of tobacco using the Agrobacterium-mediated transformation method (Horsch et al., 1985).

2.4.1. ST fusion protein expression and activity in transgenic tobacco

For each of the three constructs, transgenic plants were obtained which were found to express either high, low or practically undetectable levels of ST fusion protein in leaf tissue. High level ST fusion protein expressing lines also contained high levels of T7 RNAP activity. Variation in transgene expression levels between lines transformed with identical transgenes has been previously described (Dean et al., 1987) and is thought to be caused by the genomic DNA site of insertion or by gene silencing effects mediated by multiple copies of the transgene. Therefore it is not surprising that variation in ST fusion protein expression was observed or that plants expressing high levels of ST fusion protein and T7 RNAP activity also expressed similarly high levels of the linked GUS gene in plants that contained both.

T7 RNAP activity was detected in total leaf protein extracts from plants transformed with the three constructs and in a subset of assays activity was determined in unit/μg of total leaf protein by comparison with native T7 RNAP added to a total leaf protein extract from wild-type tobacco. The chloroplast targeted T7 RNAP expressed from the CaMV 35S promoter in the transgenic lines produced by McBride et al. (1994) expressed T7 RNAP activity from between 0.01 and 2.25 unit/μg of total leaf protein. In our transgenic lines the T7 RNAP specific activities in unit/μg were determined from the leaves of young plants and the maximum activity detected was 1.4 unit/μg for Nt-TGS25-15. Jefferson et al. (1987) found that GUS expression directed by the same SSU promoter used here, was many times higher in mature leaves than in young leaves and that the CaMV 35S promoter was many times less active in mature leaves than the SSU promoter. Therefore it is likely that the levels of T7 RNAP activity in mature leaves of our highest expressing lines would be significantly higher than what we observed in young leaf tissue.
2.4.2. Leaf tissue specificity of the SSU promoter

We detected ST fusion protein expression and activity in the leaves of transgenic tobacco generated using the three constructs pBinTGS8, 14 and 25. Previous studies have shown that expression of reporter genes from the SSU promoter in transgenic tobacco is greatest in mature leaves with progressively less expression in young leaves, stems and roots (Simpson et al., 1986; Jefferson et al., 1987) with roots containing expression 20 times less than mature leaves (Jefferson et al., 1987). The leaf specificity of the SSU promoter expression in our transgenic lines was confirmed by a western blot analysis of root tissue protein and leaf tissue protein extracts from the Nt-TGS14-7 transgenic line. High levels of ST fusion protein expression was detected in leaf tissue while the root tissue contained no detectable expression.

In the T7 RNAP expressing lines produced by McBride et al. (1994) expression was under the control of the CaMV 35S promoter which has been shown to be active in all plant organs (Odell et al., 1985). They found that pollen from tobacco producing high levels of T7 RNAP activity was infertile in crosses with chloroplast transformed lines. Expression of the ST fusion protein in the reproductive organs may have been the reason that high T7 RNAP expressors were infertile. All of our primary transformants were highly fertile when selfed.

2.4.3. Stability of ST fusion protein in transgenic tobacco

SSU mRNA steady-state levels fall to almost undetectable levels in plants that have been placed in continuous darkness for a 24 hour period (Giuliano et al., 1988; Fritz et al., 1991). Therefore it is reasonable to assume that placing ST expressing plants in the dark will shut off transcription of the transgene within a few days and ST synthesis will cease. We reasoned that by monitoring the residual levels of ST protein and activity over time, we could arrive at a crude estimate of its stability or half-life in planta.

Using this approach the stability of the ST25 fusion protein and its activity was investigated in the seed progeny of Nt-TGS25-23 F1I over the course of 16 days. The activity values in particular showed considerable variation from one time point to the next particularly between day 1 and day 3 where the activity in day 3 seedlings was 47% the activity at day 0 and the activity at day 5 jumps up to 90% of day 0 activity. These variations can be accounted for when one considers that each of the total protein extracts was made from different seedling pools that were unlikely to contain the same ST fusion protein level at Day 0. However the data are sufficient to conclude that the ST25 fusion protein is
very stable in transgenic tobacco. Indeed, both the western blot data and the T7 RNAP activity data indicate that ST25 fusion protein levels and activity decline gradually over time and that even after 16 days in the dark the T7 RNAP activity is 50% of what it was at day 0 i.e. a half-life of at least 16 days. Assuming that the half-life of ST25 in the light is similar to that observed in the dark, this suggests (i) that the levels of ST fusion proteins in mature leaves will probably be considerably higher than the levels found in very young leaves and (ii) if the st transgene is placed under the control of an inducible promoter, a considerable length of time will elapse before the maximal levels of ST fusion protein are reached.

The long half-life of the ST fusion protein may be due to an intrinsic resistance to degradation by proteases or to its localization within the chloroplast or a combination of both. Most chloroplast proteins are very stable once they have been properly assembled (Adam, 1996). For example it has been observed that there is little or no degradation of Rubisco in barley leaves placed in darkness for 7 days (Huffaker and Peterson, 1974). On the other hand photo-labile chloroplast proteins like D1 encoded by the psbA gene turn over rapidly due to photooxidative damage (reviewed in Barber and Anderson, 1992).

2.4.3. Transgene inheritance pattern and identification of homozygous lines

It was clearly shown that ST fusion protein expression is inherited in the seed progeny of F₀ and F₁ plants and that the level of ST fusion protein expression varied in the progeny of selfed F₀ lines. The variation in ST fusion protein expression was also reflected in the GUS expression levels in the same progeny plants. This variation can be explained by segregation of the transgene in the progeny of transgenic lines that were not homozygous for the transgene.

The inheritance pattern of kanamycin resistance in the F₀ and F₁ progeny of a transgenic line can be used to identify lines homozygous at the kanamycin and linked loci provided that the initial transformant was generated by a single transgene insertion event. For example, kanamycin resistance was segregating 3 : 1 in the Nt-TGS25-23 F₀ seed progeny and 100% of the Nt-TGS25-23 F₁ I seed progeny were resistant to kanamycin. These findings suggest that the Nt-TGS25-23 F₁ I line is homozygous for the kanamycin resistance gene and the other linked transgenes because (i) the inheritance pattern in the F₀ seed progeny suggest that the primary transformant was generated by a single transgene insertion event and (ii) all the progeny of Nt-TGS25-23 F₁ I
inherited the transgene. The seed progeny of primary transformants that were greater than 95% and less than 100% kanamycin resistant probably represent transgenic lines that were generated by two or more transgene insertion events because the independent assortment of multiple transgenes reduces the number of segregants that contain no kanamycin resistance gene.

Enzyme activity levels could also be used to give an indication of transgene copy number. For example, the four progeny plants of the Nt-TGS14-4 F₀ line that were tested for GUS activity each inherited different GUS activity levels which suggests the presence of more than one transgene in the parent line. This is supported by the fact that there was a low incidence of kanamycin sensitive Nt-TGS14-4 F₁ segregants.

The GUS activity expressed in three progeny plants of the Nt-TGS14-7 F₀ line was present at two discrete levels which is consistent the pattern expected in the progeny of a transformant generated by a single transgene insertion event where the transgenic locus can be either heterozygous or homozygous for the transgene. This is supported by the fact that the progeny of the Nt-TGS14-7 F₀ line were found to segregate 3 : 1 for kanamycin resistance. Therefore it is likely that the Nt-TGS14-7 F₁ line is homozygous for the linked kanamycin, GUS and st transgenes because it contains the highest level of GUS activity and ST fusion protein expression.

2.4.4. Sub-cellular location of ST fusion proteins

2.4.4.1. Detection of ST fusion protein and T7 RNAP activity in untreated and thermolysin-treated intact chloroplasts

The western blot and T7 RNAP activity analyses show that ST fusion protein was associated with and located within gradient purified washed intact chloroplasts from tobacco plants expressing the ST8, ST14 and ST25 fusion proteins. ST fusion protein and T7 RNAP activity were detected in intact chloroplasts that had been treated with thermolysin. In most cases treatment of intact chloroplasts resulted in a reduction in the amount of ST fusion protein and T7 RNAP activity that was associated with untreated intact chloroplasts. This was probably due to the degradation of ST fusion protein that was associated either with contaminating broken chloroplast membranes or with the outer membrane of intact chloroplasts. Also it is likely that spontaneous lysis of chloroplasts occurred to some extent during the thermolysin treatment. Nevertheless, thermolysin treatment confirmed that a significant proportion of the ST fusion protein
associated with intact chloroplasts was protected from degradation because of its location within the chloroplast. When Triton X-100 treated intact chloroplasts were treated with thermolysin the ST fusion protein was degraded because the Triton X-100 disrupts the chloroplast membrane and allows protease access to the stromal contents. Consequently no ST fusion protein was detected in thermolysin/Triton X-100 treated intact chloroplasts and T7 RNAP activity was dramatically reduced.

2.4.4.2. Efficiency of targeting of ST fusion proteins into chloroplasts

In order to determine the proportion of total cellular ST fusion protein that was located within the chloroplast it was necessary to determine ST fusion protein levels and T7 RNAP activity levels in total leaf protein extracts and intact chloroplast samples that contain equivalent numbers of chloroplasts. Initially these investigations were based on the assumption that Percoll gradient-purified "intact" chloroplasts were in fact mostly intact and consequently that chlorophyll content should be a reliable measure of the chloroplast content of total leaf protein extracts and purified chloroplast preparations. However, when total leaf protein and intact chloroplast samples that were equal in chlorophyll content were resolved by SDS-PAGE, the samples were often found not to be equivalent on the basis of chloroplast stromal protein content. This was most obvious in the case of Rubisco LSU, a stromal protein marker which was typically more abundant in total leaf protein extracts than in chlorophyll-equivalent samples of purified chloroplasts.

Indeed we found that the ratio of chlorophyll content in intact chloroplasts to that in Rubisco LSU equivalent total leaf protein extracts varied to a large degree from one intact chloroplast preparation to the next. In the most extreme case, 7 times the amount of intact chloroplasts (by chlorophyll content) were required to be LSU-equivalent to the total leaf protein extract. Although it is likely that gradient-purified intact chloroplasts are contaminated with some broken chloroplasts, in order to fully explain the extent of the difference and the degree of variation it is necessary to propose that mechanically prepared intact chloroplasts leak some of their stromal contents and that the extent of leakage varies from preparation to preparation.

At this point we decided that the best way to determine the proportion of ST fusion protein located within the chloroplast was to analyse total leaf protein and intact chloroplast samples that were equivalent on the basis of their LSU
content (rather than chlorophyll content) because the ST fusion protein should be targeted to the chloroplast stroma. We set out to do this by adjusting densitometry readings of ST fusion protein levels taken from western blots and activity values from T7 RNAP activity assays to give values based on LSU-equivalence.

However, it was apparent that the untreated intact chloroplast samples had a greater amount of associated ST fusion protein than was present in the corresponding LSU-equivalent total leaf protein extract. Therefore it was not possible to determine from these data the proportion of total cellular ST fusion protein that was chloroplast localised. Also, the higher levels of T7 RNAP activity found in intact chloroplasts compared with that found in the corresponding LSU-equivalent total leaf protein extracts was greater than the discrepancy in ST fusion protein levels detected by western blot analysis of the same samples. One possible explanation for this might be that total leaf protein extracts contain the total cellular pool of rNTPs (plastid and cytosolic) whereas the intact chloroplasts contain only the plastid rNTP pool. Thus in the T7 RNAP assays the total leaf protein extracts contribute a greater quantity of cold rNTPs which compete with the $^{32}$P UTP for incorporation into transcripts. This dilution effect might be sufficient to account for the finding that activity in total leaf protein extracts is lower than expected from ST content determinations. An alternative possibility is the presence of an inhibitor in total cellular protein extracts that is absent in purified chloroplast preparations.

Our analysis of ST targeting provided some evidence that the ST fusion proteins may have an affinity for broken chloroplast membranes. Although only one preparation was examined, we found that as much as 20% of the ST fusion protein associated with intact chloroplasts was also detected in a washed broken chloroplast fraction that contained almost no LSU. This suggests that a significant proportion of the ST fusion protein associated with intact chloroplasts may be adhering to the outer membrane or is alternatively associated with the thylakoid membranes. During the leaf homogenisation step used to prepare intact chloroplasts, the majority of chloroplasts burst and the released stromal proteins may co-purify with intact chloroplasts by binding to the outer membrane (Silva-Filho et al., 1996). In the majority of cases a significant proportion of the ST fusion protein and T7 RNAP activity that was associated with intact chloroplasts was sensitive to thermolysin treatment. Therefore it seems more likely that the broken chloroplast associated ST fusion protein was adhering to the outer membrane.
It was however possible to tentatively determine which fusion protein was most efficiently targeted to the chloroplast from the average ratios of ST fusion protein detected in untreated and thermolysin-treated intact chloroplasts to that present in LSU equivalent total leaf protein extracts (IT/T and I/T) for ST8, ST14 and ST25. If each fusion protein was targeted to the chloroplast with the same efficiency then the average IT/T and I/T ratios would be expected to be the same for the three fusion proteins. However, the ratios determined from the western blot data suggest that ST8 was the most efficiently targeted fusion protein followed by ST14 and ST25. Overall the data from the T7 RNAP activity determinations also suggest that increasing mSSU residues in the ST fusion protein reduces targeting efficiency. In addition by assuming that thermolysin treated intact chloroplasts best represent the amount of total cellular ST fusion protein that is located within chloroplasts it can be concluded from the western blot data that ST8, ST14 and ST25 are targeted to the chloroplast with 95, 88 and 58% efficiency, respectively. Although in this case it is important to note that the 95% figure was derived from one piece of data only.

Our data indicate that in the case of the ST fusion protein, increasing the number of N-terminal residues of mSSU reduces the targeting efficiency in a progressive manner. This has also been reported in the case of fusions between GUS and the transit peptide of the chlorophyll a/b binding protein (CAB) (Kavanagh et al., 1988). In this case, increasing the number of mature CAB residues up to 24 progressively increased targeting efficiency but further increases up to 53 or 126 residues decreased targeting efficiency. It has also been reported that an SSU transit peptide fusion protein with 23 mSSU residues was less efficiently targeted to chloroplasts in transgenic tobacco than the same fusion protein lacking mSSU residues (Kuntz et al., 1986). It is possible however that the efficiency with which the SSU transit peptide targets a given fusion protein into chloroplasts may depend more on the properties of the specific “foreign” fusion partner than on the number of mSSU residues per se.

2.4.5. Northern analysis of transcription in transgenic tobacco
Phage T7 promoters are rarely encountered in DNA unrelated to T7 DNA (Dunn and Studier, 1983). This is not surprising considering that T7 RNAP specifically recognises a 23 nucleotide promoter and that all phage T7 promoters are minor variants of the 23 nucleotide consensus. It is therefore highly unlikely, given its size, that the plastid genome would contain sequences capable of functioning as T7 promoters.
A number of in vitro studies have led to the proposal that the T7 promoter consists of two domains: a binding domain in the region of positions -16 and -5 with respect to the TIS and an initiation domain from about -5 to +5 (Chapman and Burgess, 1987; Diaz et al., 1993). Recently the minimal binding element that is both necessary and sufficient for tight binding has been determined to extend from -15 and include -5 (Ujvari and Martin, 1997). We could not identify any perfect matches for either of the domains (or variants that exist in the phage T7 genome) in the tobacco chloroplast DNA sequence. The best match found was a variant of the minimal binding domain of the phage T7 promoter found at position 4.7 in the phage DNA. It contained the first 10 of the 11 nucleotides in the domain but nucleotide 11 is reported to be necessary for tight binding in vitro (Ujvari and Martin, 1997). Thus it is not surprising that the T7 RNAP activity in our transgenic tobacco plants does not significantly alter chloroplast transcription.

We have found that there appears to be no differences in the transcription of eight plastid genes in wild-type tobacco leaves and in the leaves of the Nt-TGS25-23 plant which contains T7 RNAP activity within the chloroplast compartment. We selected the rpoC1, psbD, rps18, clpP, 3'-rps12, 16S rRNA, ndhH and ndhA plastid genes for analysis because they are located in the vicinity of the T7 promoter-like sequences that we identified in ptDNA. These promoter-like sequences were particularly homologous to the T7 promoter in the binding domain region and three of them were located within the transcriptional units of three of the above named plastid genes. Thus, we thought that the most likely activity of T7 RNAP in the chloroplast compartment would be DNA binding and we performed northern analyses to test for the consequential loss or reduction in transcription of plastid genes that might occur as a result of this binding.

However, we detected no loss of or reduction in transcription in the clpP gene nor in the rpoB/C1/C2 or the ndhHAJ operons all of which contained T7 promoter-like elements within their transcriptional units. Also, the clpP and the rps18 probes hybridised downstream of and in close proximity to T7 promoter-like elements but no new transcripts were detected in the northern analysis of Nt-TGS25-23 RNA with these probes. Therefore, we conclude that T7 RNAP does not interact with the selected T7 promoter-like elements in a manner that can significantly alter normal plastid transcription nor does there appear to be any new transcripts initiated from these sequences.
CHAPTER 3
MODIFICATION OF PLASTID TRANSFORMATION VECTORS AND CONSTRUCTION OF CHIMERIC GENE EXPRESSION CASSETTES

3.1. INTRODUCTION

Stable transformation of the plastid genome (plastome) in higher plants was first achieved in tobacco (Svab et al., 1990) by introducing mutated plastid DNA sequences into plastids using the particle bombardment method. The mutated plastid DNA integrated into the plastome through homologous recombination and the transformed plastome (transplastome) was selected for by virtue of the fact that the mutations in the transforming plastid DNA made the plastid ribosome insensitive to spectinomycin and streptomycin. Uniformly transformed (homoplasmic) lines were obtained by regenerating shoots in the presence of spectinomycin. Currently, the most commonly used method for plastid transformation in higher plants is based on the integration into plastid DNA of the bacterial gene encoding aminoglycoside 3'-adenyltransferase (aadA) through homologous recombination between plastid DNA flanking aadA in the transformation vector and the corresponding plastid DNA sequences in the plastome. The aadA gene confers resistance to spectinomycin and streptomycin and transgenic plants that are homoplasmic for the transplastome are selected by repeated rounds of shoot regeneration in the presence of spectinomycin. This method was first used by Svab and Maliga (1993) in the transformation of the plastome of tobacco using the pZS197 chloroplast transformation vector.

In this chapter we first set out to modify the chloroplast transformation vectors pZS197 (Svab and Maliga, 1993) and pSSH-B1 (an unpublished derivative of pSSH1 that contains a smaller plastid DNA insert) in order to facilitate the introduction of gene expression cassettes. The strategy employed was to insert a synthetic polylinker that would result in cloning sites being introduced to flank the aadA cassette in pZS197 and into a transcriptionally silent region of pSSH-B1. A difficulty with inserting expression cassettes into vectors in general is the availability of unique restriction sites that cut efficiently and this problem becomes particularly apparent when constructing multiple gene expression cassettes. Therefore in the pZS197 polylinker we included the sites for
two restriction enzymes, Pmel and Sgfl, that have eight-nucleotide recognition sequences that occur much more infrequently than the commonly used six cutters. The Pmel and Sgfl restriction enzymes have the ability to cut plasmid DNA effectively (Lonsdale et al., 1995).

Cassettes for high-level expression of a foreign gene in chloroplasts typically require a plastid promoter and both 5' and 3' untranslated regions (UTRs) derived from plastid genes. The 5' and 3' UTRs of plastid transcripts act as regulatory elements in the post-transcriptional control of gene expression (Eibl et al., 1999). In particular the 3' UTR has been associated with transcript stability and therefore is also known as the stability element. Regulatory elements from the 16S rDNA, psbA, rbcL, rps16 and rpl32 plastid genes have been used in chimeric gene construction for expression in chloroplasts (Eibl et al., 1999; Zoubenko et al., 1994; Svab and Maliga, 1993). It is desirable to use regulatory regions from different plastid genes in an expression cassette in order to reduce genomic instabilities that can have a negative impact on transformation efficiencies (Eibl et al., 1999). Modified chloroplast expression signals have also been used in chloroplast expression cassettes, for example the aadA expression cassette in pZS197 (Svab and Maliga, 1993) contains the 16S rDNA gene promoter (Prm) and the plastid rbcL gene leader sequence. Prm has also been shown to be very effective in chloroplasts when fused to the 5' UTR from the bacteriophage T7 gene 10 (T7G10) (Staub et al., 2000).

To engineer T7 RNAP-dependent expression of genes in chloroplasts we utilized the T7G10 expression cassette from pET3a (Rosenberg et al., 1987). The T7G10 promoter and leader sequence have already been shown to be effective in chloroplasts containing T7 RNAP activity (McBride et al., 1994). We modified the T7G10 expression cassette to contain the 3' UTR element from the plastid rps16 gene upstream of the T7G10 terminator and we introduced flanking restriction sites suitable for cloning expression cassettes into the modified pZS197 and pSSH-B1 transformation vectors. We chose the rps16 3' UTR because it had previously been used for chimeric gene expression in chloroplasts (Zoubenko et al., 1994). Also the risk of undesirable genomic rearrangements through homologous recombination should be reduced because the rps16 3' UTR is not included in the aadA expression cassette and the rps16 gene is not located near the plastid region targeted by pZS197.

We selected several genes as candidates for T7-mediated high-level expression in transformed chloroplasts based on the following considerations:
(i) the encoded protein should not disrupt chloroplast function (ii) the potential application of the encoded protein in the biotechnology industry and (iii) the encoded protein should have previously been shown to be expressed in an active form in *E. coli*. This latter criterion was considered to be particularly important because the transcriptional and translational machinery in chloroplasts is very similar to that found in *E. coli*.

Using the criteria discussed above, we selected genes coding for three distinct protein classes of interest to the biotechnology industry: (a) thermostable glycosyl hydrolases (b) single chain antibodies and (c) adult human hemoglobin. These are discussed in turn:

*a. Genes encoding thermostable glycosyl hydrolases*

We obtained two genes, *celB* and *lamA* from the hyperthermophile *Pyrococcus furiosus* which had been cloned and characterized by Professor Willem de Vos’s group at Wageningen University. The *celB* gene codes for one of the most thermostable β-glucosidases reported to date, an enzyme with a half-life of 85 hours at 100°C (Voorhorst *et al.*, 1995). In addition it shows significant β-galactosidase activity and can therefore be assayed using the classical X-Gal histochemical substrate. In this capacity and because of its thermostability, it can be used as a reporter enzyme in organisms which contain high levels of endogenous but not thermostable β-galactosidase activity (see Chapter 4). Furthermore, since the enzyme does not show hydrolytic activity on starch (personal communication) its expression in chloroplasts should not be deleterious. Antibodies for western blot detection of the *celB* protein were also available.

The *lamA* gene encodes the most thermostable β-glucanase reported to date with a temperature optimum of 100-105°C and like *celB*, it does not hydrolyse starch (Gueguen *et al.*, 1997). The thermostable enzymes encoded by *celB* and *lamA* are of biotechnological interest because of their ability to hydrolyse various classes of carbohydrates at very high temperatures (reviewed in Bauer *et al.*, 1998).

*b. Genes encoding single chain antibodies*

In most antibodies, the antigen binding domain is formed by the interaction between two variable domains which are part of separate heavy and light polypeptide chains linked via disulphide bridges in the mature antibody molecule. Thus the production of recombinant antibody fragments with full antigen binding capacity requires the expression of the two domains as separate polypeptides or
the linking of the domains using a peptide linker. Although the introduction of a peptide linker would simplify antibody expression it can create problems like reduced affinity, aggregation or proteolytic cleavage (Whitlow et al., 1993). However, the Camelidae produce a large proportion of their functional immunoglobulins as homodimers of heavy chains which lack light chains (Hamers-Casterman et al., 1993). Therefore, we obtained the genes for two single chain camel antibodies (cAbs) against tetanus toxin (cAb-TT2) and chicken egg lysozyme (cAb-Lys3) from S. Muylldermans's group at the vrije Universiteit Brussel.

Both cAb-TT2 and cAb-Lys3 have a predicted molecular mass of about 16 Kd, have been produced in a soluble form in E. coli and have been shown to interact specifically and with high affinity with their respective antigens (Arbabi Ghahroudi et al., 1997). These cAbs are ideal for testing in chloroplast expression systems because (i) they are small single chain polypeptides, (ii) they are translationally fused to the hemagglutinin protein of human influenza virus (HA tag) which can be used for affinity purification, (iii) antibodies against the HA tag are commercially available, (iv) their expression in chloroplasts should not be deleterious to plant growth. In addition, the relatively low-cost production of very large quantities (kilogram) of antibodies in plants would make the many potential biotechnological applications of recombinant antibodies more feasible. These include preventive immunotherapy by topical, oral or intravenous delivery of antibodies, the inactivation of bacterial toxins (eg the cholera and tetanus toxins) and a host of other potential applications in diagnostics, cancer therapy, etc.

c. cDNAs encoding the α and β-globin strands of adult human hemoglobin
We obtained two unmodified cDNAs encoding the α and β-globin strands of adult human hemoglobin from Professor Chien Ho of Carnegie Mellon University. These cDNAs have been used to express fully functional recombinant hemoglobin (based on spectral and oxygen binding studies) to high levels in E. coli (Shen et al., 1997). It would be interesting to determine the feasibility of producing large quantities of functional hemoglobin in chloroplasts. This would present an alternative production strategy for blood substitute products in a production system that would be free of contamination with human pathogens like HIV.

We describe the introduction of the various genes described above into the modified T7G10 expression cassette, the cloning of the gene expression
cassettes into chloroplast transformation vectors and the investigation of the functional integrity of some of the chimeric gene cassettes in an *E. coli* expression system.

### MATERIALS AND METHODS

#### 3.2.1. Bacterial strains and plasmids

*Escherichia coli* DH5α (Plasmids, 1983) was used as the host for cloning experiments and for the amplification of plasmid DNA for purification. *E. coli* JM110 DH1 (Grossberg and Davis, 1982; Studer et al., 1990) was used as host for expression of the fusion antibody fragment and hemagglutinin.

pZS107 (Svedberg and Mattig, 1995) was the parent plasmid used in the construction of pCP107.

pPSH-B3 (unpublished plasmid) is the modified version of pPS11 (Grosset et al., 1994) that was used in the construction of pLS1-B1a.

pET5a (Rosenthal et al., 1997) was used as the source for the *E. coli* gene 10 expression signal plus the termination that were used in the construction of pETCP5 and pETCP6.

pUC9 (Vamathe-Perrone, 1985) was used in the construction of pNTCP5 and pNTCP6.

DNA plasmids pUB-SET-2 and pUBN54-Leu8 (Chamanchi et al., 1997) were the source of the functional (ZAB-FT2) and neo-resistance (ZAB-FT3) canary antibody gene fragments, respectively.

pHP7 (Shen et al., 1997) is the source of the cDNA sequence of human and *E. coli* globin strands of *Escherichia coli* hemoglobin.

pLAWS33 and pLAWS31 are unpublished plasmids, provided by Professor Otto Oost, Dept. Microbiology, Wageningen Agricultural University, Suchthuizenweg 4, 6703 HW Wageningen, The Netherlands, which were source of the *P. aeruginosa* glutamyl tRNA*sec* gene encoding a thermostable glutamyl tRNA*sec* synthetase and *E. coli* gene encoding a thermostable p-galactosidase.

#### 3.2.2. Antibodies

Mouse monoclonal antibodies (clone 12C10) to the hemagglutinin protein of human influenza virus were supplied by Recombinant Mucinogen Technologies, Madison, Wisconsin, monoclonal antibodies (clones 114, 110G) to human hemagglutinin were supplied by AutogenBioclas.
3.2. MATERIALS AND METHODS

3.2.1. Bacterial strains and plasmids

*Escherichia coli* DH5α (Hanahan, 1983) was used as the host for cloning experiments and for the amplification of plasmid DNA for purification. *E. coli* BL21 DE3 (Grodberg and Dunn, 1988; Studier *et al.*, 1990) was used as host for expression of the camel antibody fragment and hemoglobin.

pZS197 (Svab and Maliga, 1993) was the parent plasmid used in the construction of pCP197.

pSSH-B1 (unpublished plasmid) is the modified version of pSSH1 (Kavanagh *et al.*, 1994) that was used in the construction of pSSH-B1a.

pET3a (Rosenberg *et al.*, 1987) was used as the source of the T7 gene 10 expression signals including the terminator that were used in the construction of pETCP5 and pETCP6.

pUC9 (Yanisch-Perron, *et al.*, 1985) was used in the construction of pETCP5 and pETCP6.

DNA phagemids pHEN4-TT2 and pHEN4-Lys3 (Ghahroudi *et al.*, 1997) were the source of the anti-tetanus (cAb-TT2) and anti-lysozyme (cAb-Lys3) camel antibody gene fragments, respectively.

pHE7 (Shen *et al.*, 1997) was the source of of the cDNA encoding the α and β-globin strands of adult human hemoglobin.

pLUW532 and pLUW511 (unpublished plasmids, provided by John van der Oost, Dept. Microbiology, Wageningen Agricultural University, H. van Suchtelenweg 4, NL-6703 CT Wageningen, The Netherlands) were the source of the *Pyrococcus furiosus* lamA gene encoding a thermostable endo-β-1,3-glucanase and celB gene encoding a thermostable β-glucosidase, respectively.

3.2.2. Antibodies

Mouse monoclonal antibody (clone 12CA5) to the hemagglutinin protein of human influenza virus was supplied by Boehringer Mannheim. Mouse monoclonal antibody (clone B11.10G) to human hemoglobin was supplied by AutogenBioclear.
3.2.3. PCR and sequencing

PCR amplification of cAb gene fragments, hemoglobin, lamA and celB genes and PCR amplification of chloroplast DNA fragments from the Nt-197Hb line were performed as described in Section 2.2.19. All sequencing reactions were performed as described in Section 2.2.6.

3.2.4. Oligonucleotides used for DNA sequencing and PCR

In cases where the oligonucleotide sequence was based on tobacco ptDNA (Shinozaki et al., 1986), the position of the 5' nucleotide in plastid genome is given in brackets and (C) denotes that the 5' nucleotide is complementary to the nucleotide at that position in the plastid genome.

Oligonucleotides used for sequencing pZS197 and related constructs:

(1) Prrn-Rev (102638) (C): 5'-GAAATATAGCCATCCCTGCCC-3'
(2) psbA3'UTR-For: sequence not available
(3) psbA3'UTR-Rev (446): 5'-GAAAGGAGCAATAGCACCCTCTTTG-3'
(4) orf512-Rev (59372) (C): 5'-ATACTTGGCTTGATTCAGGG-3'

Oligonucleotides used for sequencing pSSH-B1a and related constructs:

(1) Δ9-For (101521(C)/141008): 5'-TTTTCTTTGGGGAGGTATCGGGAAG-3'
(2) pSSH-Rev (101150/141369(C)): 5'-CTTGATCAATCCCTTTGCC-3'

Oligonucleotides used for sequencing pETCP5-derived expression cassettes:

(1) Hbα-For: 5'-GTTCAACCCTGCAGGTGCACG-3'
(2) Hbα-Rev: 5'-ACATCCTCTCCAGGGCCTCC-3'
(3) celB-Rev: 5'-TGATGCTATGTTCTCCTTGTCG-3'
(4) lamA-Rev: 5'-CCCGGGATTCCATAAGCTATTC-3'
(5) GFP-Rev: 5'-CCCTACCCCTCTCCACTGACAG-3'
(6) rps16-Rev (4937) sequence given below

Oligonucleotides used in the amplification of the rps16 3' UTR and cAb and hemoglobin genes. Underlined regions are completely homologous to ptDNA.

(1) rps16-For (5092) (C): 5'-CGGGATCTCCAACCGGAATTCATAAGAGG-3'
(2) rps16-Rev (4937): 5'-TGAAGATCTGGAATTCAATGGAAGCAATG-3'
(3) cAb-For: 5'-CGGAATTCATATGGCCCAGGTGCAGCTG-3'
(4) cAb-Rev: 5'-CGGGATCCTAGCCGGAACCGTAGTCCGG-3'
(5) Hbα2-For: 5'-GAGAATTCCATATGGTGCTGTCTGCC-3'
(6) Hbα2-Rev: 5'-CGCGGATCTGTAACATTAACCGGTATTTGGAGGTC-3'
3.2.5. Oligonucleotides used to make synthetic polylinkers

The 197 synthetic polylinker used in the construction of pCP197 was formed by annealing the following oligonucleotide pair;

197a: 5'-GATCACCTAGGCGATCGCAGTCGACAGATCTTAAGCTTATCGATGTTTAAAC-3'
197b: 5'-GATCGTTTAAACATCGATAAGCTTAAGCTGACTGCGATCAGCCTAGGT-3'

The SSH-B1a synthetic polylinker used in the construction of pSSH-B1a was formed by the annealing of the following oligonucleotide pair;

SSH-B1: 5'-TTAAGCTTCATCTAGACGTCGACGTTTAAACT-3'
SSH-B2: 5'-TTAAGGTTTAACATCGATAAGCTTATCGACTGCGATCGCCTAGGT-3'

The CP5 and CP6 synthetic polylinkers used in the construction of pETCP5 and pETCP6, respectively were formed by the annealing of CP5a with CP5b and CP6a with CP6b;

CP5a: 5'-AGCTCCTTAAGCGCTTATCGATATCGATGTTTAAAC-3'
CP5b: 5'-AATTGTTrAAACATCGATATCAGATCTAAGCTTAAGC-3'
CP6a: 5'-AGCTTCCTAGGAGATCTGATATCGTCGAC-3'
CP6b: 5'-AATTGTGACGATATCAGATCTCCTAGG-3'

3.2.6. Expression of chimeric genes in E.coli

Unless otherwise stated, the expression of recombinant proteins in E.coli BL21 DE3 was performed in the following manner. Overnight cultures were grown from single colonies in liquid LB medium supplemented with ampicillin 100mg/L, glucose 2% at 37°C with shaking at 200rpm. The next day a 1/10 dilution was made in fresh LB medium supplemented with ampicillin and glucose and the cells were grown at 37°C to the mid-exponential growth phase and induced by adding IPTG to a final concentration of 0.4mM. Induced cultures were then grown for a further 16 hours. Aliquots were taken immediately before induction (uninduced) and at various time points after induction and prepared for analysis by SDS-PAGE.
3.2.7. Western blotting

SDS-PAGE and western blotting were carried out essentially as described in Section 2.2.15. except that *E. coli* protein extracts from cultures expressing cAb fragments and hemoglobin were resolved on 15% polyacrylamide gels and nitrocellulose filters were probed with anti-HA at a concentration of 5μg/ml in the detection of cAb and anti-hemoglobin at a concentration 5μg/ml in the detection of hemoglobin. Alkaline phosphatase coupled anti-mouse antibodies were used in the secondary immunological reaction.
3.3. RESULTS

3.3.1. Structural analysis of pZS197

We aimed to construct a more user-friendly, modular version of pZS197 (Svab and Maliga, 1993) that would (i) facilitate the introduction of expression cassettes containing chimeric genes and their associated regulatory sequences, at either side of the aadA selectable marker and (ii) facilitate replacement of the rbcL and orf512 (aacD) plastid targeting sequences flanking the aadA cassette with other plastid DNA sequences. In order to achieve this objective, we first set out to produce a more detailed restriction map of pZS197 (Figure 3.1) by (i) analysis of the component sequences that make up pZS197 to identify all the predicted restriction enzyme cleavage sites, (ii) restriction digestion analysis of pZS197 to verify presence/absence of expected sites and (iii) sequencing of various junctions between the component fragments in pZS197 i.e. the junctions between vector (pBluescript) and plastid DNA (ptDNA), ptDNA and the aadA expression cassette and between the aadA coding sequence and the psbA gene 3' regulatory region.

The sequence of the SacII-EcoRV tobacco ptDNA fragment (nucleotides 57,750 and 60,595; Shinozaki et al., 1986) used in the construction of pZS197 was extracted from GenBank and the positions of restriction enzyme cleavage sites were mapped within the sequence using the GeneJockey sequence analysis software. Similarly, expected cleavage sites in the chimeric aadA cassette were determined by analysis of the Prm promoter region and the psbA3' UTR which consist of ptDNA nucleotides from 102,561-102,678 and 533-141, respectively. The synthetic leader sequence downstream of the Prm promoter and the sequence of the aadA coding region (GenBank, accession number AF061065) was also included in the analysis. The restriction map of pBluescript was obtained from the Stratagene catalog.

The agarose gels in Figure 3.2 show a sample of the restriction digestions in which pZS197 was incubated with one or more restriction enzymes. In some cases the plasmid DNA appears unchanged from control pZS197 DNA indicating that no sites for the restriction enzyme in question were present. In other cases pZS197 was linearised indicating the presence of one cleavage site (or two very closely spaced sites) and in cases where restriction fragments were excised from the plasmid more than one restriction site was present. Table 3.1
Figure 3.1. Restriction map of pZS197.
Table 3.1. Restriction enzyme cleavage sites that are not present, are unique or occur more than once in pZS197 as determined by restriction digestion analysis, database sequence analysis and sequencing. The number of sites present is indicated in brackets where determined.

<table>
<thead>
<tr>
<th>Restriction sites not present in pZS197</th>
<th>Unique restriction sites</th>
<th>Restriction sites occurring more than once</th>
</tr>
</thead>
<tbody>
<tr>
<td>AvrII, BstXI, Clal, EcoRI, EcoRV, HpaI, HindIII, MfeI, Mscl, NcoI, NheI, NotI, SalI, SmaI, AflII, AgeI, MluI</td>
<td>NdeI, SacI, SacII, Spel</td>
<td>ApaI, BamHI (3), BglII, KpnI (3), NsiI (2), PstI (2), SphI (2), Xbal, Xhol (2)</td>
</tr>
</tbody>
</table>
shows the results of the restriction analysis for all commonly used restriction enzymes tested including the restriction enzymes that did not have any cleavage sites in pZS197.

Sequencing using the Prrn-Rev, psbA3'UTR-For, psbA3'UTR-Rev, orf512-Rev, M13-For and M13-Rev sequencing oligonucleotides identified the restriction sites present at the sequence junctions in pZS197. The priming region and the direction of sequencing is shown in Figure 3.1 for each sequencing oligonucleotide. Sequencing made it possible to identify restriction sites at junctions that were used in the construction of pZS197 (Figure 3.1). Sequencing with Prrn-Rev and orf512-Rev confirmed that the chimeric \textit{aadA} gene in pZS197 is cloned between the BamHI sites at positions 59,285 and 59,305 in ptDNA and that the \textit{aadA} expression cassette is therefore flanked by single BamHI sites. The 20 bp fragment located between these BamHI sites is not present in pZS197.

The value of using the three techniques together in the elucidation of the pZS197 restriction map was demonstrated by the identification of a unique SpeI cleavage site that proved to be very useful in facilitating subsequent constructions. Database sequence analysis indicated the presence of a SpeI site in the \textit{psbA3'} sequence. SpeI linerised pZS197 (Figure 3.2A, lane 6) indicating that the SpeI site was unique. Double digestion of pZS197 with SpeI and SacII excised a 4.5 and a 2.6 Kb restriction fragment (Figure 3.2B, lane 3) confirming that the SpeI site was located 2.6 Kb from SacII which was the same distance as the \textit{psbA3'} element was from SacII. The presence of the SpeI site in the \textit{psbA3'} sequence in pZS197 was further confirmed by sequencing with the \textit{psbA3'}UTR-Rev sequencing oligonucleotide.

Also the number of NsiI and XhoI sites present in pZS197 was determined in a similar manner. Both NsiI (Figure 3.2B, lane 5) and XhoI (result not shown) appeared to linerise pZS197. However, analysis of the ptDNA in pZS197 using the GeneJockey analysis programme revealed that NsiI sites are found at two closely spaced positions in pZS197 (Figure 3.1). Also, sequencing with the M13-For primer revealed that a second XhoI site is present at the orf512-pBluescript junction approximately 120 bp from the XhoI site at ptDNA position 60,484 in pZS197.
Figure 3.2 A, B and C. Restriction digestion analysis of pZS197. A. Lane 1, Kb ladder; lane 2, undigested pZS197; lanes 3 and 8, SacII/XhoI and BamHI/XhoI double digestions, respectively, lanes 4-7, Dral, NcoI, SpeI and BamHI digestions, respectively. B. Lane 1, Kb ladder; lane 2, undigested pZS197; lanes 3, 4 and 5, SacII/SpeI, XbaI and NsiI digestions, respectively. C. Lane 1, Kb ladder; lanes 2-7, PstI, KpnI, SphI, SacI, NcoI and BstXI digestions, respectively.
3.3.2. Modification of pZS197 to give pCP197

3.3.2.1. Construction of pCP197

pCP197 was constructed in order to introduce multiple unique cloning sites into pZS197 at each end of the aadA gene cassette. The cloning steps in the construction of pCP197 are diagrammed in Figure 3.3 and can be divided into three steps, (1) the construction of pCP197poly, (2) the construction of pCP7-aadA1 and (3) the cloning of the aadA expression cassette from pCP7-aadA1 into pCP197poly to give the final construct pCP197.

(1) pCP197poly was constructed by first linearising pZS197 with Spe I which cuts inside the aadA cassette (Figure 3.3; Figure 3.4). Spe I-linearised plasmid was then partially digested with BamHI and a 5.7 Kb fragment comprising all of pZS197 except the aadA expression cassette was isolated by gel-purification. The desired fragment is the product of cleavage of only two of the 3 BamHI sites in pZS197, those flanking the aadA cassette (i.e. at positions 59285 and 59305 in ptDNA). Isolation of this fragment enabled replacement of the aadA cassette with the synthetic poly linker 197 sequence which was formed by the annealing of the two single stranded oligonucleotides 197a and 197b. This synthetic polylinker possessed BamHI cohesive ends and contained the following restriction sites: AvrII, Sgfl, Sall, BglII, AflII, HindIII, ClaI and Pmel of which BglII was the only one present in pZS197. The polylinker was designed so that the BamHI sites at the ends of the gel-purified 5.7Kb pZS197 partial digestion fragment would be eliminated following ligation with the polylinker.

Figure 3.4A shows the BamHI partial digestion of Spel-linerised pZS197. The partial digest in lane 5 (0.35 units of BamHI per μg of Spel linerised pZS197) was selected because it contained the maximum amount of the partial fragments in the 5-6Kb range. Figure 3.4B shows a scale up of this partial digest that was run slowly on a 0.7% agarose gel. The two partial fragments in the 5-6Kb size range were gel purified and ligated with the annealed 197 polylinker. The ligation reaction was again digested with SpeI to eliminate any products that contained the chimeric aadA gene and was then transformed into E.coli DH5α. DNA was prepared from 12 ampicillin resistant colonies for analysis of plasmid structure.

Figure 3.5 shows the restriction digestion analysis that was carried out to identify and confirm the structure of candidate pCP197poly clones. BamHI digestion of the 12 plasmid DNA clones identified three clones that were linerised
at the remaining BamHI site (in the case of pCP197poly, this site was located in the \textit{rbcL} gene at position 58,048 in ptDNA) and had a molecular weight of approximately 5.7Kb (Figure 3.5A, lanes 1, 4 and 7) as expected for BamHI digested pCP197poly. Figure 3.5B shows confirmation of the structure of pCP197poly in respect of one of these clones. The absence of a SpeI site (Figure 3.5B, lane 2) confirmed that the chimeric \textit{aadA} gene was absent and digestion with NsiI (Figure 3.5B, lane 3) confirmed that the NsiI sites located in the \textit{rbcL} DNA in pZS197 were present. HindIII/XhoI digestion (Figure 3.5B, lane 4) reduced the plasmid size from 5.7Kb to around 4.4 Kb and thus confirmed the presence of the \textit{orf512} fragment. The excised HindIII/XhoI orf512 (\textit{accD}) fragment (of approximately 1.2Kb) migrated out of the gel and therefore is not shown in Figure 3.5B. The presence and orientation of the polylinker sequence in putative pCP197poly clones was checked by sequencing with the orf512 Rev sequencing primer. The pCP197poly clone with the polylinker in the \textit{rbcL-AvrII/PmeI-orf512} orientation was used in the construction of pCP197.

(2) \textbf{pCP-aadA1} was constructed by ligating the synthetic polylinker 197 into BamHI-digested pUC7 to give pCP6 and pCP7 each containing the polylinker in different orientations. pCP7 (with the AvrII to PmeI polylinker sites in the opposite orientation to LacZ expression) was digested with BgIII, treated with calf intestinal phosphatase (CIP) and was then ligated with the gel-purified 1.35Kb BamHI fragment containing the chimeric \textit{aadA} cassette from pZS197. Ligation resulted in the insertion of the \textit{aadA} cassette into pCP7 in either of two orientations to give pCP-aadA1 and pCP-aadA2. Figure 3.3 shows the differences in restriction site positions between pCP7-aadA1 and pCP7-aadA2 and Figure 3.6 shows the restriction digests that were carried out to identify and distinguish between pCP7-aadA1 and pCP7-aadA2. HindIII/KpnI, AflII/PstI and AflII/SpeI double digests result in restriction patterns that distinguish pCP7-aadA1 and 2. For example, HindIII/KpnI double digestion appears to linearise pCP7-aadA2 (Figure 3.6, lane 5) due to the close proximity of the two sites whereas HindIII/KpnI double digestion excises the 1.35Kb \textit{aadA} cassette from pCP7-aadA1 (Figure 3.6, lane 1). Both constructs were also sequenced using the M13-\textit{For} universal primer to confirm the orientation of the \textit{aadA} cassette.

(3) The final construct, \textbf{pCP197}, was constructed by gel purification of the 1.35Kb HindIII/SalI \textit{aadA} cassette fragment from a HindIII/SalI double digest of pCP7-aadA1 and ligating it with a HindIII/SalI double digest of pCP197poly that had been dephosphorylated with CIP. Correct ligation resulted in insertion of the \textit{aadA} cassette into the middle of the 197 polylinker, thus providing a set of unique
Step 1. Construction of pCP197poly

197 polylinker

BamHI

BamHI

rbcL

Spel Linerisation
BamHI partial digestion
gel purification of 5.7Kb partial fragment
CIP treatment
T4 DNA ligation
Spel digestion

BamHI AvrII SgfI SalI BglII AflII HindIII ClaI Pmel BamHI

BamHI

rbcL

ORF512

pZS197

7.10 Kb

pCP197poly

7.10 Kb

Figure 3.3. Construction of pCP197
Step 2. Construction of pCP7-aadA1

Figure 3.3 (contd.). Construction of pCP197
Figure 3.3 (contd.). Construction of pCP197
Step 3. Cloning of aadA from pCP7-aadA1 into pCP197 poly

Figure 3.3 (contd.) Construction of pCP197
Figure 3.4. A. Partial digestion/analysis in identity paterns for pZS197 clones. Mix of 150 clones that were tested by mobility analysis in lanes 1 and 4. Lanes 2, 3, 5 and 6, Spel linerised pZS197 digested with 1.4, 0.7, 0.35, 0.17 and 0.08 units of BamHI per μg of DNA, respectively. B. Scale up of the pZS197 partial digest in lane 3 run for 4.5 hours in a 0.7% agarose gel.
Figure 3.5. A. Restriction digestion analysis to identify putative pCP197poly clones. Six of the twelve clones that were tested by incubation with BamHI are shown in lanes 1-4 and 7 and 8. Lanes 5 and six, Kb ladder and BamHI digested pZS197, respectively. Lanes 1, 4 and 7, 5.7Kb linerised plasmid DNA expected for pCP197poly. B. Restriction digestion analysis to confirm the structure of pCP197poly. Lane 1, undigested pCP197poly DNA; lanes 2-4, Spel, Nsil and HindIII/Xhol digested pCP197poly; lane 5, aliquot of BamHI partial of pZS197 from Figure 3.4A, lane 5; lane 6, Kb ladder.

Figure 3.6. Restriction digestion analysis to confirm the structure of and distinguish between pCP7-aadA1 and pCP7-aadA2. Lanes 1-3, pCP-aadA1 doubly digested with HindIII/Kpnl, AflIII/PstI and AflIII/Spel, respectively; lane 4, Kb ladder; lanes 5-7, pCP-aadA2 doubly digested with HindIII/Kpnl, AflIII/PstI and AflIII/Spel, respectively.
Figure 3.7. Restriction digestion analyses to confirm the structure of pCP197. A. Lane 1, pZS197; lane 2, pCP197; lanes 3, 4, 6 and 7, pCP197 clones digested with BamHI; lane 5, Kb ladder; lane 8, BamHI digested pCP197poly. B. Lane 1, pCP197; lanes 2, 3, 5 and 6, HindIII/Sall doubly digested pCP197 clones; lane 4, Kb ladder. C Lane 1, Kb ladder; lane 2, pZS197; lane 3, CsCl purified pCP197; lanes 4 and 5, pCP197 digested by Pmel and Sgfl, respectively.
restriction sites on either side of the \textit{aadA} cassette (Figure 3.3). The \textit{aadA} cassette from pCP7-aadA1 was used in the construction of pCP197 because ligation into pCP197poly resulted in its orientation being the same as the \textit{aadA} cassette in pZS197 (Svab and Maliga, 1993). Figure 3.7 shows confirmation of the structure of pCP197 by restriction digestion analysis of putative pCP197 clones. BamHI digestion linerised pCP197 to give a larger 7.1 Kb product than the 5.7Kb BamHI linerised pCP197poly (Figure 3.7A). HindIII/SalI double digestion of pCP197 excised a 1.35Kb fragment corresponding in size to the chimeric \textit{aadA} gene (Figure 3.7B). Figure 3.7C shows the Pmel and Sgfl linerisation of pCP197. All the newly introduced restriction sites flanking \textit{aadA} were checked by restriction digestion and sequencing with orf512-Rev and Prrn-Rev primers.

### 3.3.2.2. Modification of pSSH-B1 to give pSSH-B1a

pSSH-B1 (unpublished plasmid) is a modified version of pSSH1 (Kavanagh et al., 1994; Kavanagh et al., 1999) in which the ptDNA insert has been reduced from 7.8 to 6.2 Kb by bi-directional Bal31 deletion. It contains sequences from the plastid genome of the \textit{Solanum nigrum} StSp1 mutant corresponding to nucleotides 136,940 - 143,090 of the plastid genome of \textit{N. tabacum} (Shinozaki, et al., 1986). Deletion of the ptDNA insert in pSSH1 with Bal31 reduced the size of the insert and removed one of two AfIII sites. Thus, ptDNA insert in pSSH-B1 contains a unique AfIII site in the intergenic region between \textit{trnV} and \textit{rps12-3'} at position 141,232 (relative to \textit{N. tabacum}).

pSSH-B1a was constructed by digesting pSSH-B1 at the unique AfIII site, treating it with CIP and ligating it with the synthetic polylinker SSH/B1a. This polylinker was formed by the annealing of two single stranded oligonucleotides, SSH-B1 and SSH-B2. Both ends of the synthetic linker could be ligated with AfIII-digested pSSH-B1 but only one end could be subsequently reclaved with AfIII. Correct ligation of the linker resulted in the following polylinker sites being introduced; AfIII, HindIII, XbaI, Sall and Pmel (Figure 3.8). The insertion of the linker was confirmed by HindIII, Sall, XbaI and Pmel digestion (result not shown). pSSH-B1a plus (+) and minus (-) versions that contain the polylinker in opposite orientations were identified by sequencing with the \textit{A9-For} sequencing oligonucleotide.
Figure 3.8. pSSH-B1a (+/-) contains sequences from the plastid genome of the Solanum nigrum StSp1 mutant corresponding to nucleotides 136,940 - 143,090 of the plastid genome of N. tabacum. (Shinozaki, et al., 1986). The AflIII-Pmel polylinker was inserted at position 141,232 in ptDNA (relative to N. tabacum). The positions in the ptDNA at which the point mutations that confer spectinomycin and streptomycin resistance are shown.
3.3.3. Construction of modified T7 gene 10 expression cassettes for directing T7RNAP-mediated gene expression in chloroplasts (pETCPS and pETCP6)

Modified versions of the T7 gene 10 expression cassette from pET3a (Rosenberg et al., 1987) similar to but not identical to those described by McBride et al., (1994) were constructed in order to facilitate T7 RNAP-mediated expression of transgenes in chloroplasts. The resulting plasmids, pETCPS and pETCP6, whose construction is described below, contain T7 gene 10 promoter, translational and terminator signals but with the chloroplast rps16 3' UTR which functions as an mRNA stability element, inserted between the T7 translation and terminator signals. The modified expression cassettes were also constructed to contain suitable flanking restriction enzyme sites which would facilitate subsequent insertion of the cassette (and an inserted chimeric gene of interest) into the chloroplast transformation vectors pCP197 and pSSH-B1a (described above). Thus, the restriction sites flanking the modified T7 expression cassette in pETCPS made it compatible for cloning into either pCP197 or pSSH-B1a. The sites flanking the T7 expression cassette in pETCP6 made it compatible for cloning into pCP197 but not pSSH-B1a.

(1) Construction of pETCPS

Construction of pETCPS is shown diagrammatically in Figure 3.9. In the first step, pUC9 was digested with HindIII/EcoRI to excise the polylinker region and a new polylinker was introduced by ligation with the CP5 synthetic polylinker obtained by annealing the two single-stranded oligonucleotides, CP5a and CP5b. As a result of ligation, the HindIII and the EcoRI sites used in the cloning were eliminated and a new cloning vector called pUC5 was obtained which contained the following sites in the polylinker: AfIII, HindIII, BgIII, EcoRV, Clal, Pmel. The construction of pUC5 was confirmed by restriction digestion analysis (result not shown) and sequencing with M13 universal primers.

The T7 gene 10 expression cassette was excised from pET3a and gel purified as a 300 bp BgIII/EcoRV fragment. This fragment was ligated with BgIII/EcoRV digested, CIP-treated pUC5 to give pET5. The construction of pET5 was confirmed by excision of a 300 bp EcoRV/BgIII fragment from putative clones (result not shown) and by sequencing with M13 universal primers.

The 3' untranslated regulatory region of the tobacco chloroplast rps16 ribosomal protein gene (rps16-3') located between nucleotides 5092 to 4937
Figure 3.9. Construction of pETCP5
Figure 3.9 (contd.). Construction of pETCP5
Figure 3.9 (contd.). Construction of pETCP5
Figure 3.10. Restriction digestion analysis to confirm the structure of pETCP5. The agarose gels were overloaded with DNA in order to make the small 150bp and 250 bp restriction fragments visible. A. Lane 1, uncut pETCP5; lanes 2, 3, 5 and 6, BamHI/EcoRV doubly digested pETCP5 clones; lane 4, Kb ladder. B. Lanes 1-4 and 5-8, BamHI/EcoRV doubly digested pETCP5 clones; lane 5, Kb ladder.
(Shinozaki et al., 1986) in the ptDNA of N. tabacum was PCR-amplified from total DNA extracted from tobacco leaves using the rps16-For and rps16-Rev oligonucleotides. Amplification with these primers placed a BamHI at the 3' end and EcoRI and BglII sites at the 5' end of the 150 bp rps16-3' product. pET5 was digested with BamHI, treated with CIP and ligated with the BamHI/BglII-digested rps16-3' PCR product to give pETCP5. The BglII site in the rps16-3' element was eliminated upon ligation with the BamHI cohesive end and therefore unique BamHI and EcoRI sites flanked the element in pETCP5.

Figure 3.10A shows the identification and confirmation of clones containing the rps16-3' UTR element by the excision of a 150bp fragment from clones doubly digested with BamHI/EcoRI. Figure 3.10B shows restriction analysis that was performed to identify pETCP5 clones that contained the rps16-3' UTR in the same orientation as it is found in the rps16 gene. Clones that produced a 250bp BamHI/EcoRV restriction fragment (Figure 3.10B, lanes 2, 6, 7 and 9) contain the rps16-3' UTR in the desired orientation (Figure 3.9). Insertion of the regulatory region in the correct orientation in pETCP5 was confirmed using M13 universal primers and the rps16-Rev oligonucleotide in sequencing reactions.


pETCP6 (Figure 3.11) is identical to pETCP5 except that the modified T7 expression cassette is flanked by different restriction sites. pETCP6 was made in the same way as pETCP5 but a different synthetic polylinker (CP6) was cloned into pUC9 at the first step in its construction. The linker was formed by annealing two oligonucleotides, CP6a and CP6b which contains recognition sites for AvrII, BglII, EcoRV and Sall. The structure of pETCP6 was confirmed by performing the same restriction analysis and sequencing as described for pETCP5 (result not shown).

3.3.4. Construction and cloning of chimeric genes into chloroplast transformation vectors

3.3.4.1. Introduction of genes encoding camel single chain antibodies into the chloroplast transformation vectors pCP197 and pSSH-B1a

Two genes encoding camel single-chain antibodies (cAbs) one of which is specific for tetanus toxin (cAb-TT2) and the other specific for chicken egg
lysosome (cAb-Lyk3) were cloned separately into the chloroplast transformation vectors pCP197 and pSSH-B1a (described above) to produce the plasmids pSSH-cAbTT2, pSSH-cAbLyk3, pCP197-cAbTT2 and pCP197-cAbLyk3. The construction pathways are diagrammed in Figure 3.12 and 3.13.

To produce these constructs, the genes encoding cAb-TT2 and cAb-Lyk3 each of which encodes a C-terminal hemagglutinin (HA) tag (Wilson et al., 1984) were first PCR-amplified from the DNA plasmids pHEN4-TT2 and pHEN4-Lyk3, respectively. Figure 3.14 shows the 370bp cAb-TT2 and cAb-Lyk3 PCR products that were then PCR-amplified and digested with AvrII and Ndel and then ligated into pETCP6. The T7 promoter rbs and 12 aa tag rps163' followed by Ndel, Ndel digesting the T7 terminator, was inserted into the pETCP6 vector followed by Ndel, Ndel digesting the T7 terminator, was inserted into the Ndel digesting the T7 terminator vector pETCP6). The partially digested pETCP6 vector is shown below in Figure 3.15.

The resulting plasmid was sequenced and vector integrity was confirmed by digestion with BamHI and Ndel/BluntII (not shown) and then ligated with BamHI/BluntII (not shown) (Figure 3.16A) and pCP197-cAbTT2 and pCP197-cAbLyk3 (Figure 3.16B).

Figure 3.11. pETCP6 is identical to pETCP5 except that the flanking restriction sites are different.

The pETCP6 and pETCP5 plasmids were transformed into M13 universal particles and the resulting PCR of each construct was cloned into a universal plasmid. The controlled expression vector was then tested on the RNA of interest. HA tags had inserted correctly.
lysozyme (cAb-Lys3) were cloned separately into the chloroplast transformation vectors pCP197 and pSSH-B1a (described above) to produce the plasmids: pSSH-cAbTT2, pSSH-cAbLys3, pCP197-cAbTT2 and pCP197-cAbLys3. The construction pathways are diagrammed in Figure 3.12 and 3.13.

To produce these constructs, the genes encoding cAb-TT2 and cAb-Lys3 each of which encodes a C-terminal hemagglutinin (HA) tag (Wilson et al., 1984) were first PCR-amplified from the DNA phagemids, pHEN4-TT2 and pHEN4-Lys3, respectively (Ghahroudi et al., 1997). Figure 3.14 shows the 370bp cAb-TT2 and cAb-Lys3 PCR products that were both PCR amplified using Pfu polymerase and the cAb-For and cAb-Rev primers. The PCR products were digested with BamHI and NdeI for cloning between the T7 promoter/rbs and the rpsl6-3' UTR in pETCP5. Unfortunately the cloning of the antibody gene fragments into the modified T7 expression cassette in pETCPS was complicated by the presence of an extra NdeI site in the vector portion of the plasmid. Therefore pETCPS was prepared for cloning by BamHI linearisation followed by NdeI partial digestion. Figure 3.15 shows the partial digestion of BamHI-linerised pETCP5. The digest shown in lane 9 contained 0.39 units of NdeI per µg of BamHI-linerised pETCP5 which was the concentration at which the NdeI partial product appeared. The partially digested pETCP5 DNA was dephosphorylated with CIP and was ligated with BamHI/NdeI-digested camel antibody PCR products to give pETCP-TT2 and pETCP-Lys3.

Two putative pETCP-TT2 and one pETCP-Lys3 clones were identified by their increased size when linerised with BamHI compared to BamHI-linerised pETCP5 parent plasmid (result not shown). The three clones were analysed by NdeI/BamHI double digestion which excised the 370bp antibody fragments (Figure 3.16A) and the T7 expression cassette excised by HindIII/EcoRV double digestion was found to have increased in size to approximately 800bp as expected upon ligation of the cAb PCR products with the BamHI and the correct NdeI site (Figure 3.16B). In Figure 3.16B, lane 5 the smaller HindIII/EcoRV fragment excised from pETCP5 parent plasmid migrated out of the gel.

The pETCP-TT2 and pETCP-Lys3 clones were sequenced with M13 universal primers and the rps16-Rev primer to confirm that the entire T7-controlled expression cassette was intact and that the antibody fragments and their HA tags had inserted correctly.
Figure 3.12. Construction of pSSH-cAbTT2 and pSSH-cAbLys3
Figure 3.12 (contd.). Construction of pSSH-cAbTT2 and pSSH-cAbLys3
Figure 3.13. Construction of pCP197-TT2 and pCP197-Lys3
Figure 3.14. PCR amplified cAb-TT2 and cAb-Lys3 from pHEN4-TT2 and pHEN4-Lys3, respectively. Lanes 1 and 5, Kb ladder; lanes 2-4, cAb-TT2; lanes 6-8, cAb-Lys3.

Figure 3.15. NdeI partial digestion of BamHI linerised pETCP5. Lane 1, BamHI linerised pETCP5; lanes 2-4 and 6-9, 100ng of 400ng of BamHI linerised pETCP5 digested with 10, 5, 2.5, 1.25, 0.625, 0.325 and 0.156 units of NdeI, respectively; lane 5, KB ladder.
Figure 3.16. Restriction digestion analysis to confirm the structure of pETCP-TT2 and pETCP-Lys3. A. Excision of the 370bp antibody fragment. Lanes 1 and 3, BamHI/Ndel doubly digested pETCP-TT2 clones; lane 2, Kb ladder; lane 4, BamHI/Ndel doubly digested pETCP-Lys3. B. Excision of the 800bp cAb expression cassette. Lanes 1 and 3, HindIII/EcoRV digested pETCP-TT2 clones; lane 2, Kb ladder; lane 4, HindIII/EcoRV doubly digested pETCP-Lys3; lane 5, HindIII/EcoRV doubly digested pETCP5.

Figure 3.17. Restriction digestion identify pSSH-cAbTT2 and pSSH-cAbLys3 clones by the excision of the 800bp expression cassette. Lane 5, Kb ladder; lanes 1-4 and 5-9, AflII/Pmel digested putative pSSH-cAbTT2 and pSSH-cAbLys3 clones, respectively.
The chimeric cAb-TT2 and cAb-Lys3 genes were excised from pETCP-TT2 and pETCP-Lys3 by AfII/Pmel double digestion, gel-purified and ligated with AfII/Pmel digested, dephosphorylated pSSH-B1a (+) to give pSSH-cAbTT2 and pSSH-cAbLys3. Figure 3.17 shows the excision of cAb chimeric genes as 800bp AfII/Pmel fragments from putative pSSH-cAbTT2 and pSSH-cAbLys3 clones. Cloning was confirmed by sequencing across the insert with the D9-For and pSSH-Rev oligonucleotides.

The chimeric camel antibody genes were excised and gel-purified from HindIII/ClaI doubly digested pETCP-TT2 and pETCP-Lys3 and ligated with HindIII/ClaI doubly digested and dephosphorylated pCP197 to give pCP197-cAbTT2 and pCP197-cAbLys3. Figure 3.18A shows the HindIII/ClaI excision of the chimeric genes from putative pCP197-cAbTT2 and pCP197-cAbLys3 clones. Figure 3.18B shows the excision of the aadA and cAb chimeric genes as a 2.2Kb fragment from EcoRV/SalI doubly digested pCP197-cAbTT2 and pCP197-cAbLys3 clones. Sequencing with psbA3′-For and orf512-Rev confirmed that the cAb genes had correctly inserted downstream of the aadA gene.

3.3.4.2. Introduction of genes encoding human α-globin and β-globin into the chloroplast transformation vectors pCP197 and pSSH-B1a

In this section, I describe the construction of a dicistronic hemoglobin gene cassette containing the α-globin and β-globin genes coding for normal adult hemoglobin (HbA) and its introduction into pCP197 and pSSH-B1a to give the plasmids pCP197-Hb and pSSH-Hb, respectively. The construction pathway leading to pCP197-Hb and pSSH-Hb is diagrammed in Figures 3.19 and 3.20.

In the first construction step, the human α-globin gene was PCR-amplified from pHE7 (Shen et al., 1997) using the Hba2-For and Hba2-Rev oligonucleotide pair. The oligonucleotides were designed to introduce an Ndel site at the 5′end and a Hpal and BamHI site at the 3′ end of the amplified α-globin sequences to facilitate cloning into pETCP5. The α-globin PCR product was digested with BamHI/Ndel and ligated with the BamHI-digested, Ndel-partially-digested pETCP5 (Section 3.3.4.1) to give pETCPα. Figure 3.21 shows the restriction digestion analysis that was carried out to confirm the structure of pETCPα. The α-globin gene is 430bp in length and contains a diagnostic HindIII site at position 271. HindIII/EcoRV double digestion excised two fragments of approximately 350 and 450bp that resulted from the excision of
Figure 3.18. Restriction digestion analysis to confirm the structure of pCP197-cAbTT2 and pCP197-cAbLys3. 

A Excision of 800bp cAb expression cassette. Lanes 1-4 and 6-9, HindIII/ClaI doubly digested pCP197-cAbTT2 and pCP197-cAbLys3 clones, respectively; lane 5, Kb ladder.

B Excision of 2.2Kb aadA-cAb fragment. Lanes 1-4, uncut, EcoRV, SalI and EcoRV/SalI digested pCP197-cAbTT2; lanes 5-7, uncut, EcoRV/SalI, and AvrII digested pCO197-cAbLys3; lane 8, Kb ladder.
Figure 3.19. Construction of pCP197-Hb
Figure 3.19 (contd.), Construction of pCP197-Hb
In the second construction step, the β-globin gene was PCR-amplified from pETCP-Hb (Chen et al., 1993) using the oligonucleotide pair HpaI-Rev and HpaI-Rev. The HpaI-Rev oligonucleotide contained the following restriction elements: (5' to 3') a HpaI site at the 5' terminus of the region from the bovine β-globin gene (977–1554) and a HindIII site at the 3' terminus. The PCR product was digested with HindIII and HpaI and ligated with pETCP-Hb double digested, dephosphorylated pETCP-Hb to give pETCP-Hb which contains the downstream β-globin gene pair, pETCP-Hb, and was subsequently identified by the presence of the expected 1.3 Kb HpaI restriction fragments. The sequences of pETCP-Hb were confirmed by sequencing with the following primers: 5' AflIII Xbal Pmel n pSSH-Hb, 5' tmV a-globin P-globin rps12(3').

Figure 3.20. Construction of pSSH-Hb
the T7 expression cassette containing Hbα and its cleavage at the HindIII site in Hbα coding sequence.

In the second construction step, the β-globin gene was PCR-amplified from pHE7 (Shen et al., 1997) using the oligonucleotide pair Hbβ-For and Hbβ-Rev. The Hbβ-For oligonucleotide contained the following sequence elements: (5’-3’) a HpaI site, the ribosome binding site region from the tobacco rbcL gene (nucleotides 57579-57596 in ptDNA (Shinozaki et al., 1986)) and sequences complementary to the 5’ end of the β-globin gene. The Hbβ-Rev oligonucleotide contained a BglII site and sequences complementary to the 3’ end of the β-globin gene. The resulting PCR product was digested with BglII and HpaI and ligated with BamHI/HpaI doubly digested, dephosphorylated pETCPa to give pETCP-Hb which contains the dicistronic hemoglobin gene pair. pETCP-Hb clones were identified by the excision of the dicistronic 1.3 Kb Hb expression cassette fragment by AfII/Pmel double digestion (result not shown). The structure and sequence of pETCP-Hb was confirmed by sequencing with the M13-Rev, rps16-Rev, Hbα-For and Hbα-Rev oligonucleotides.

The Hb expression cassette was introduced into pCP197 by excising it from pETCP-Hb by AfII/Pmel double digestion, gel-purification of the relevant fragment and ligation with AfII/Pmel doubly digested, dephosphorylated pCP197 to give the final construct pCP197-Hb. Figure 3.22 shows restriction digest analysis of a pCP197-Hb clone: digestion with AfII and Pmel releases the 1.3Kb dicistronic Hb expression cassette and digestion with EcoRV/Sall releases the 2.6Kb aadA-Hb fragment. The structure of pCP197-Hb was also confirmed by sequencing with the psbA 3’UTR-For and orf512-Rev oligonucleotides.

The Hb expression cassette was also introduced into pSSH-B1a. This was accomplished by replacing the Xbal-Pmel cAbLys3 fragment in pSSH-cAbLys3 (Figure 3.20) with the Xbal-Pmel Hb cassette fragment from pETCP-Hb to give the final construct pSSH-Hb. The structure of pSSH-Hb was confirmed by sequencing with the Δ9-For and pSSH-Rev oligonucleotides.

3.3.4.3. Introduction of genes encoding a thermostable β-glucosidase and a β-glucanase into the chloroplast transformation vector pSSH-B1a.

Genes encoding the β-glucosidase (celB) and endo-β-1,3-glucanase (lamA) of *Pyrococcus furiosus* (Voorhorst et al., 1995; Gueguen et al., 1997) were cloned.
Figure 3.21. Restriction digestion analysis to confirm the structure of pETCPα.
Lane 1, Kb ladder; lanes 2 and 3, HindIII/EcoRV double digestion of pETCPα clones.

Figure 3.22. Restriction digestion analysis to confirm the structure of pCP197-Hb. Excision of 1.3Kb hemoglobin expression cassette and 2.6Kb aadA-hemoglobin fragment from pCP197-Hb. Lanes 1-4 and 6, uncut, EcoRV, Sall, EcoRV/Sall and AflII/Pmel digested pCP197-Hb, respectively; lane 5, Kb ladder.
into the modified T7 expression cassette in pETCP5 and the resulting expression cassettes were transferred into pSSH-B1a to give pSSH-lamA and pSSH-celB. The construction pathways are diagrammed in Figures 3.23 and 3.24.

1. Construction of pSSH-lamA

The coding sequence of lamA together with the T7 promoter and translational signals (0.9Kb fragment) was excised from the pET9d derivative pLUW532 (unpublished plasmid provided by John van der Oost) by BgIII/BamHI double digestion, gel-purified and ligated with BgIII/BamHI doubly digested, dephosphorylated pETCP5 to give pETCP-lamA. The structure of pETCP-lamA was confirmed by the excision of the 900bp BamHI/BglII T7-lamA cloning fragment and the 800bp NcoI/BamHI lamA coding sequence fragment from pETCP-lamA (result not shown). Also sequencing with the lamA-Rev oligonucleotide was performed to check sequence in the lamA start codon region. The lamA expression cassette minus the T7 promoter (1.2Kb fragment) was excised from pETCP-lamA by XbaI/Pmel double digestion, gel purified and ligated with XbaI/Pmel doubly digested, CIP treated pSSH-cAbLys3 to give the final construct pSSH-lamA. The structure of pSSH-lamA was confirmed by restriction digestion analysis (result not shown) and by sequencing with Δ9-For, rps16-Rev and pSSH-Rev oligonucleotides.

2. Construction of pSSH-celB

The 1.5Kb coding sequence of celB was excised from pLUW511 (unpublished plasmid provided by John van der Oost) by NcoI/BamHI double digestion, gel purified and ligated with NcoI/BamHI doubly digested, dephosphorylated pETCP-lamA to give pETCP-celB. Excision of the 1.5Kb NcoI/BamHI fragment and sequencing with the celB-Rev oligonucleotide was used to identify pETCP-celB clones. The celB expression cassette minus the T7 promoter was excised from pETCP-celB by XbaI/Pmel double digestion, gel purified and ligated with XbaI/Pmel doubly digested, dephosphorylated pSSH-cAbLys3 to give pSSH-celB. The structure of pSSH-celB was confirmed by restriction digestion analysis (result not shown) and by sequencing with Δ9-For, rps16-Rev and pSSH-Rev oligonucleotides. Sequencing with Δ9-For confirmed that the T7 promoter and G10 leader sequence in pSSHcelB was identical to that in pSSH-cAbLys3.

3.3.5. Expression of chimeric genes in E. coli

Because the transcription and translation system in chloroplasts is very similar to that in bacteria, we investigated expression of the various chimeric genes described above in the E. coli prior to their introduction into the chloroplast
Figure 3.23. Construction of pSSH-lamA
Figure 3.24. Construction of pSSH-celB
genome. For these experiments, we used the *E. coli* strain BL21(DE3) which contains a chromosomally integrated T7 RNAP gene under control of the *lacZ* promoter. Production of T7 RNAP itself and consequently T7 RNAP-mediated gene expression is therefore inducible using IPTG.

3.3.5.1. Expression of camel antibody genes in *E. coli*

*E. coli* BL21(DE3) was transformed with pETCP-TT2 (*E. coli* strain DE3-TT2) and pETCP-Lys3 (*E. coli* strain DE3-Lys3) and IPTG-induced cultures were analysed for cAb expression by immunoblotting using polyclonal antibodies against the HA tag which was present as a C-terminal fusion on each cAb. Figure 3.25 shows the immunodetection of a single polypeptide in uninduced cultures and two polypeptides in IPTG-induced cultures of DE3-Lys3. Total *E. coli* protein extracts were prepared from the DE3-Lys3 culture at the mid-exponential growth phase before the addition of IPTG and three hours post-induction. The polypeptide detected in the uninduced DE3-Lys3 culture had a molecular mass of approximately 16Kd. This was within the apparent molecular mass (MW) range reported for cAb proteins resolved by SDS-PAGE (Arbabi Ghahroudi *et al.*, 1997) and the MW calculated for cAb-Lys3 from its amino acid sequence was 16Kd. The two polypeptides detected in the induced DE3-Lys3 culture had apparent MWs of 16 and 17Kd. No corresponding proteins were detected in the non-transformed BL21(DE3) control culture.

SDS-PAGE and western blot analysis of total cellular protein from uninduced and IPTG-induced cells of strain DE3-TT2, using the conditions employed for DE3-Lys3 cultures, did not reveal expression of the cAb-TT2 antibody. For this reason, the following alterations to the expression procedure were employed in an attempt to optimize cAb-TT2 expression: (1) TB broth was used in place of LB broth, (2) after induction the culture was grown at 30°C instead of 37°C, (3) *E. coli* BL21(DE3) pLysS was used as the expression host and (4) the concentration of IPTG was increased from 0.4mM to 1mM. However, no cAb-TT2 expression was detected in the protein extracts prepared from these cultures (result not shown). Also, it was not possible to detect cAb-TT2 or cAb-Lys3 expression in BL21(DE3) cultures transformed with pCP197-TT2 and pCP197-Lys3 (result not shown).

In order to determine the stability of the expression plasmids in *E. coli* BL21 (DE3) cultures, aliquots taken before and after induction were diluted and plated on LB-agar and LB-agar supplemented with ampicillin 100mg/L. The majority of the bacteria in DE3-TT2 and DE3-Lys3 cultures, before the addition
Figure 3.25. Western blot of protein extracts from DE3-Lys3, DE3-TT2 and nontransformed DE3 expression cultures probed with anti-HA antibodies. Lanes 1 and 2, DE3-Lys3 at exponential growth phase (uninduced) and 3 hours post induction (pi), respectively; lanes 3 and 4, uninduced and 3 hours pi DE3 culture; lanes 5 and 6, uninduced and 3 hours pi DE3-TT2 culture.
of IPTG, were ampicillin resistant and consequently the expression construct was stable up to the mid-exponential phase before the addition of IPTG. However, after the addition of IPTG to a final concentration of 0.4mM and incubation until the culture density had doubled, the number of DE3-TT2 colonies that were ampicillin resistant had halved. In contrast greater than 95% of the *E. coli* cells in the DE3-Lys3 culture were ampicillin resistant at this time point (data not shown).

### 3.3.5.2. Expression of hemoglobin in *E.coli*

Prior to carrying out expression studies in *E. coli* BL21(DE3), pETCP-Hb was first introduced into *E. coli* DH5α and overnight cultures were grown for the preparation of plasmid DNA. However, when the cells in overnight cultures were harvested the pellet was reddish-brown in colour instead of the usual beige colour. This colour effect has been previously observed in *E. coli* cultures expressing hemoglobin (Hernan *et al.*, 1992). This effect was also observed for *E. coli* DH5α cultures transformed with pSSH-Hb but not for pCP197-Hb cultures.

When *E. coli* BL21(DE3) was transformed with pETCP-Hb (*E. coli* strain DE3-Hb), it was found that the addition of 2% glucose to the growing medium was sufficient to repress this effect and was if fact required in the overnight culture step in order to achieve a sufficiently high DE3-Hb cell density prior to induction i.e. a cell density similar to nontransformed control cultures.

Figure 3.26A shows SDS-PAGE analysis of total protein extracts of *E. coli* cultures expressing hemoglobin and the corresponding western blot (Figure 3.26B) shows the immunodetection of hemoglobin using anti-human hemoglobin (anti-Hb) antibodies.

A faint signal representing a polypeptide with a MW of 14Kd was detected in the DE3-Hb culture at 2 hours post-induction (Figure 3.26B, lane 8). No corresponding signal was detected in the samples taken immediately prior to induction (lane 6) or 16 hours post-induction (lane 10). Although, both α- and β-globin have a MW of about 14Kd the detected protein was probably β-globin which was found to be more abundantly expressed than α-globin in pHE7 expression cultures (Shen *et al.*, 1997).

SDS-PAGE analysis of *E.coli* JM109 containing pHE7 (designated strain 109-Hb) was also analysed. An extra protein with a MW of 30Kd was clearly visible in the protein gels stained with Coomassie Blue at 16 hours post-induction (Figure 3.26A, lane 3). This protein corresponds to methionine
Figure 3.26. Protein gel showing protein extracts from *E. coli* BL21 DE3 and *E. coli* JM109 expression cultures and corresponding western blot probed with anti-human hemoglobin antibodies. A. 15% SDS-PAGE gel. Lanes 1-3, pHE7 (plasmid for hemoglobin expression) in JM109 immediately before induction (uninduced), 2 hours post-induction (PI) and 16 hours PI, respectively; lane 4, HMW marker; lanes 5, 7 and 9, pETCP5 (modified T7 gene expression cassette with no introduced gene) in BL21 DE3 uninduced, 2 hours PI and 16 hours PI, respectively; lanes 6, 8 and 10, pETCP-Hb (contains the adult human hemoglobin expression cassette) in BL21 DE3 uninduced, 2 hours PI and 16 hours PI, respectively. B. Western blot. Lane identifications are as for A. Arrows indicate methionine aminopeptidase (MAP) and hemoglobin.
aminopeptidase (MAP) which is co-expressed with hemoglobin from pHE7 (Shen et al., 1997). In the corresponding western blot (probed with anti-Hb antibodies), a polypeptide was detected in the 109-Hb culture 2 hours post-induction that corresponded to the polypeptide detected in the DE3-Hb culture in size and intensity (Figure 3.26B, lane 2). However in the 109-Hb culture this polypeptide continued to accumulate and was detected at a higher level after overnight incubation (Figure 3.26B, lane 3).

It was not possible to detect hemoglobin expression in BL21(DE3) cultures transformed with pCP197-Hb either in the presence or absence of induction (result not shown). Expression of hemoglobin in BL21(DE3) transformed with pSSH-Hb was found to be detectable by pellet colour and western blotting in overnight cultures grown even in the absence of IPTG (result not shown).

3.3.5.3. Expression of lamA and celB in E. coli

Because of the inconclusive results obtained above concerning the dependence of chimeric gene expression on T7 RNAP, neither celB nor lamA expression in E. coli was analysed by SDS-PAGE or western blotting. Phenotypic analysis of E. coli DH5α containing pETCP-lamA or pETCP-celB indicated a low level of constitutive expression of both genes in the absence of IPTG: pelleted cells of E. coli DH5α containing pETCP-lamA appeared more milky in colour than the normal beige-colored pellets, a phenotypic difference that has been attributed to the endo-β-1,3-glucanase activity of the lamA gene product while E. coli DH5α colonies containing pETCP-celB were blue on LB-agar plates containing, 2% glucose, ampicillin 100mg/L and X-Gal after overnight incubation at 37°C whereas identical plates containing DH5α-pETCP5 had white colonies. pSSH-celB which is a derivative of pSSH-Bla (+) containing the T7-celB expression cassette was tested in the same manner and was also found (on the basis of X-Gal cleavage) to express celB activity constitutively in DH5α.

We next decided to test whether reversing the orientation of the T7-celB expression cassette in the pSSH-B1a vector would eliminate its constitutive expression in DH5α. This was accomplished by cloning the expression cassette into pSSH-B1a (-) to give pSSH-celB(-) (results not shown). The structure of pSSH-celB (-) was confirmed by sequencing with Δ9-For, celB-Rev and pSSH-Rev. As predicted, DH5α colonies containing pSSH-celB(-) no longer showed constitutive expression of celB and were white when grown on LB agar plates containing 2% glucose and X-Gal.
3.3.5.4. T7 RNAP-dependant expression of GFP in *E. coli*

In order to demonstrate conclusively the functionality of the modified T7 expression cassette prior to its use in chloroplasts, we cloned GFP into pETCP5 and then transferred the T7-GFP expression cassette into pSSH-B1a(-) in which context constitutive expression appears not to occur (as discussed above).

This was accomplished as follows: the 700bp GFP coding sequence (Crameri *et al.*, 1996) was excised from pTpGFP (unpublished plasmid) by NheI digestion and BamHI partial digestion, and cloned into pETCP5 to give pETCP-GFP (results not shown). Because transgenes cloned into pETCP5 show a low level of constitutive expression in *E. coli* DH5α (discussed above), colonies containing pETCP-GFP could be identified by their green fluorescence on exposure to ultraviolet light. The structure of pETCP-GFP was confirmed by sequencing with the M13-Rev and rps16-Rev oligonucleotides. The T7-GFP expression cassette was excised from pETCP-GFP as a 1.1Kb Xbal/Pmel fragment which was used to replace the T7-celB expression cassette in pSSH-celB(-) to give pSSH-GFP(-). Clones containing pSSH-GFP(-) were identified by differences in their NcoI/BamHI and NheI/EcoRI restriction patterns when compared with pSSHcelB(-) (result not shown).

pSSH-GFP(-) was introduced into the T7 RNAP expression strain BL21(DE3) pLysS (Studier, *et al.*, 1990) and a single colony was streaked on two LB-agar plates containing selective antibiotics and supplemented with 2% glucose. One of the plates also included IPTG which induces expression of T7 RNAP. Both plates were incubated at room temperature for several days and photographed while exposed to ultraviolet light (Figure 3.27). The bacterial colonies on both plates expressed green fluorescence but those on the IPTG plate were clearly more intensely fluorescent. The fluorescence in the absence of IPTG can be attributed to background T7 RNAP activity because pSSH-GFP(-) did not express green fluorescence in *E. coli* DH5α cells which lack T7 RNAP.
Figure 3.27. T7 RNAP dependant GFP expression from pSSH-GFP(-) in E.coli BL21 DE3 plysS. The plate on the right contains IPTG whereas the plate on the left does not.
3.4. DISCUSSION

3.4.1. Construction of versatile derivatives of chloroplast transformation vectors

Our aim in this chapter was to construct a suite of vectors with compatible restriction sites that would facilitate the construction of foreign gene expression cassettes based on regulatory sequences of phage T7 and which would enable these cassettes to be introduced into and exchanged between different chloroplast transformation vectors.

The vector pZS197 (Svab and Maliga, 1993) in which selection is based on the bacterial \textit{aadA} gene, has limited utility as a general chloroplast transformation vector because of the rarity of unique restriction sites. For this reason, we decided to construct a versatile, modular derivative, pCP197, containing multiple cloning sites at either side of the \textit{aadA} selectable marker into which we could introduce foreign gene expression cassettes. The choice and sequence of cloning sites was based on the principle that expression cassettes should ideally be capable of being transferred between vectors in a single cloning step. To facilitate the exchange of single or multiple cassettes between pZS197 and pSSH-B1a, sites for the rare-cutting 8 bp recognition restriction enzymes Sgfl and Pmel were included.

Two derivatives of the vector pSSH-B1 (which lacks a multiple cloning site) were also constructed to contain a polylinker cloning site that would facilitate exchange of expression cassettes with pCP197. The derivatives, pSSH-B1a (+) and pSSH-B1a (-) contain the polylinker in alternative orientations. We have found that the introduction of expression cassettes derived from pETCP5 into pSSH-B1a (-) results in the expression cassette being in a transcriptionally silent orientation in \textit{E. coli}. In contrast, when the hemoglobin and celB expression cassettes derived from pETCP5 were introduced into pSSH-B1a (+), the genes were constitutively expressed in \textit{E. coli}. The introduction of pETCP5-derived expression cassettes into pSSH-B1a (+) resulted in the expression cassette being in the same orientation as the LacZ and \(\beta\)-lactamase genes in the pUC plasmid body. Therefore, it is likely that the expression resulted from the LacZ and \(\beta\)-lactamase promoters in the vector portion of the plasmid because the polylinker
in pSSH-B1a is located in a region of chloroplast DNA between divergently transcribed regions.

### 3.4.2. Modification of the T7 G10 expression cassette

The various gene expression cassettes described in this chapter were based on the promoter and terminator regulatory sequences from gene 10 of phage T7. Expression cassettes based on these sequences are most frequently used for directing T7-dependent gene expression in *E. coli* (Studier et al., 1990). We modified the T7 gene 10 expression cassette from pET3a to optimize the expression of chimeric genes in the chloroplasts of transgenic plants expressing a chloroplast-targeted T7 RNAP (Chapter 2) by including the *rpsl6* 3'UTR element which has previously been used in chimeric gene construction for expression in chloroplasts (Zoubenko *et al.*, 1994). The restriction sites flanking the modified T7 regulatory sequences ensured that the expression cassettes (with their chimeric genes) could be easily cloned into both pCP197 and pSSH-B1a. In addition, because of the modular structure of the cassette (each regulatory element is defined by unique restriction sites) it will be possible to further modify it by replacing individual elements with analogous regulatory sequences from other genes e.g. a chloroplast 5' UTR could be used in place of the T7 gene 10 leader sequence.

### 3.4.3. Expression of chimeric genes in *E. coli*

Because chloroplast gene expression systems are prokaryotic-like and T7 RNAP dependent expression systems have been developed in *E. coli* (Studier and Moffatt, 1986) we decided to test the expression of our chimeric genes in *E. coli*. We used the BL21 DE3 expression strain which contains an IPTG inducible T7 RNAP gene. We detected the expression of the anti-lysozyme camel antibody fragment (cAb-Lys3) and hemoglobin by western blotting of protein extracts from *E. coli* cultures that were transformed with pETCP-Lys3 (DE3-Lys3) and pETCP-Hb (DE3-Hb). Although this was a demonstration that the genes had been cloned in the correct frame we could not conclude that the expression was T7 RNAP dependent because of the constitutive nature of expression of genes in pETCP5 derived constructs in DH5α. It is likely that the constitutive expression of chimeric genes in pETCP5 derived constructs was caused by the lacZ promoter in the pUC plasmid body because the addition of glucose 2% in the growing medium was sufficient to abolish the constitutive expression of hemoglobin in DE3-Hb cultures. Glucose is known to inhibit the *lac* operon via the catabolite repression system.
Western blot analysis detected two polypeptide with apparent molecular weights (MW) of approximately 16 and 17 Kd in the IPTG-induced DE3-Lys3 culture. Only the smaller polypeptide was detected in the culture immediately prior to induction. Based on the amino acid sequence of cAb-Lys3 (including the HA tag) its predicted MW is about 16Kd. Given that the larger product detected in the induced culture is IPTG-dependent it probably represents the correct cAb-Lys3 product. The smaller product may have been produced due to the presence of a cryptic promoter close to the orf which results in the expression of a truncated product lacking some of the N-terminal residues of the full-length cAb.

We detected the expression of a single polypeptide in induced DE3-Hb cultures with an apparent MW of 14Kd. It is likely that the polypeptide detected was the β-globin subunit of hemoglobin because this is the most abundantly expressed subunit in E. coli systems expressing hemoglobin cDNAs (Shen et al., 1997; Hernan et al., 1992). It was encouraging that the redish-brown colour observed in the pellets of uninduced overnight cultures indicated that both the α and β- strands of hemoglobin were expressed from the hemoglobin operon in the pETCP-Hb expression cassette.

The instability of expression constructs in IPTG induced E. coli expression strains containing chimeric genes constructed using the pETCPS expression cassette may explain why it was not possible to detect cAb-TT2 expression in DE3-TT2 cultures or hemoglobin expression in DE3-pCP197-Hb cultures. The evidence for plasmid instability comes from the finding that 50% of the cells in the DE3-TT2 culture have lost ampicillin resistance 2 hours after induction. Also hemoglobin expression in the DE3-Hb culture was detectable 2 hours after induction but not after overnight incubation. Whereas, hemoglobin expression from pHE7 in JM109 was detectable 2 hours post induction and increased after overnight incubation. However in the case where hemoglobin is expressed from pSSH-Hb in DE3 without the addition of IPTG, expression was detected after overnight incubation.

A similar finding was reported previously by Guda et al. (1995) using the pET11d expression system for the expression of a synthetic polymer gene in E. coli. They correlated the very low level of polymer production in induced cells with plasmid loss and aberrant cell growth. However they achieved high expression levels in cultures that were not induced and therefore the polymer itself was not responsible for the effect through toxicity. They proposed that expression
in the absence of IPTG was due to the dilution of lac repressor molecules in the rapidly growing expression cultures.

CHAPTER 4

TRANSIENT AND STABLE TRANSFORMATION OF TOBACCO CHLOROPLASTS

4.1 INTRODUCTION

To date several proteins have been expressed at varying levels in transgenic tobacco plants that were either stably or transiently transformed with the encoding DNA. Proteins expressed in stably transformed tobacco plants include: AAO (Patino et al., 1992; GUS (McHarg et al., 1986), bacteriophage promoters (McHarg et al., 1995; Kata et al., 1999), a synthetic protein-based polymer (Su et al., 2001), 5-enzyme polygalacturonase-pectin methylesterase complex (SIPPE; Wang et al., 1999), and a human growth hormone (Ghaderpour et al., 2001). In contrast, transient expression of a gene has been achieved in these stable 

In addition to tobacco, expression of GUS has been reported in transgenic potato plants (Gaff et al., 1999) and the expression of another gene coding for a protein to be expressed in transgenic potato plants has been achieved (Gaff et al., 1999). Other types of chloroplasts and the chloroplasts of tobacco, contain to be expressed from O'Brien et al., 1999).

Recent advances in the manipulation of chloroplasts have made it possible to express various proteins in transgenic plants, with chimeric probes containing the promoter, enhancer, and/or internal transcribed spacer of the chloroplast DNA. Transgenic tobacco plants containing the chloroplast DNA of a petunia (Petunia hybrida) have been produced, in which the expression of the carotenoid biosynthesis enzyme 

expression cassette encoding the GUS gene was expressed during the stage of the tobacco plant. Both the GUS and chloroplast DNA were used as transgenic constructs to deliver the GUS gene to the chloroplast. The expression cassette was placed under the control of a chloroplast gene promoter and rifampicin-selected transgenic tobacco plants were used to test the functional activity of the GUS gene in transgenic plants.

88
CHAPTER 4

TRANSIENT AND STABLE TRANSFORMATION OF TOBACCO CHLOROPLASTS.

4.1. INTRODUCTION

To date several proteins have been expressed at varying levels in chloroplasts that were either stably or transiently transformed with the encoding DNA. Proteins expressed in stably transformed tobacco plants include AADA (Svab and Maliga, 1993), GUS (McBride et al., 1994), insecticidal proteins (McBride et al., 1995; Kota et al., 1999), a synthetic protein-based polymer (PBP) (Guda, et al., 2000), 5-enolpyruvyl shikimate-3-phosphate synthase (EPSPS) (Daniell et al., 1998) and a human growth hormone (somatotropin) (Staub et al., 2000). Furthermore, recombinant protein expression has been achieved in the chloroplasts of other species with reports of GFP expression in stably transformed transplastomic potato plants (Sidorov et al., 1999) and the expression of an AADA-GFP fusion protein in heteroplasmic transplastomic rice plants (Khan and Maliga, 1999). In addition, transient expression of GFP has been reported in the plastids of different tissue types including potato amyloplasts and the chromoplasts of marigold petals, carrot roots and pepper fruits (Hibberd et al., 1998).

Recombinant protein accumulation in chloroplasts has also been achieved by expressing the products of nuclear transgenes as translational fusions with chloroplast targeting sequences. For example, recombinant adult human hemoglobin (rHbA) has been successfully produced in transgenic tobacco containing the genes encoding the α and β-globin strands of HbA in separate expression cassettes in the nucleus (Dieryck and Gruber, 1997). The two expression cassettes were introduced into a single Agrobacterium Ti plasmid and transgenic plants were produced using the method described by Horsch et al. (1985). Both the α and β-globin strands of HbA were translationally fused to the SSU transit peptide from Pisum sativum in order to target rHbαA to the chloroplast. The expression cassettes were driven by the constitutive CaMV 35S promoter and rHbA expression was detected in leaf, root and seeds. rHbA was extracted from seeds by two chromatography steps and was shown to be functional by comparing its biphasic kinetics to native HbA.
Although the above strategy was successful, rHbA was only produced at 0.05% of total extracted seed protein. This is typical of the yields reported for recombinant proteins produced from nuclear transgenes. For example, Staub et al. (2000) only managed to express recombinant human somatotropin from a nuclear transgene to 0.025% of total soluble leaf protein whereas it was produced to levels greater than 7% of total soluble leaf protein from a chloroplast transgene. In addition the chloroplast is potentially a more suitable location for the production of rHbA because many chloroplast genes are naturally grouped in operons. This means that the globin gene pair can be introduced into the chloroplast genome as a single dicistronic expression cassette. Both orfs on a dicistronic mRNA are efficiently translated in chloroplasts (Staub and Maliga, 1995) but not in the cytoplasm (Angenon et al., 1989). For this reason, in the nuclear transformation system it was necessary to introduce the globin gene pair as two separate expression cassettes. However, the introduction of two separate genes, even in a single T-DNA fragment, into nuclear DNA by the Agrobacterium method may result in different relative expression levels from one transformant to another (reviewed in Bogorad, 2000).

In this chapter we describe (i) the transient expression of GUS in tobacco leaves for the optimisation of the biolistic procedure, (ii) the transient expression of celB in Nt-TGS14-7 leaves transiently transformed using the biolistic procedure, (iii) the generation of transplastomic tobacco plants stably transformed with a dicistronic adult human hemoglobin expression cassette and (iv) the crossing of the transplastomic line with a line expressing the ST14 fusion protein in order to produce tobacco plants with the potential to express hemoglobin in the chloroplast in a T7 RNAP-dependent manner.
4.2. MATERIALS AND METHODS

4.2.1. Plasmids
Most of the plasmids used in the experiments described in this Chapter were described in Chapter 3. The following plasmids were also employed:

- **pCTAKI** (unpublished plasmid) contains the GUS gene under transcriptional control of the *CaMV35S* promoter and was used to direct transient expression of GUS in the cytoplasm following biolistic bombardment of detached tobacco leaves.

- **pSSHPrnGUS** (unpublished plasmid) contains the GUS gene under transcriptional control of the chloroplast *rnm* promoter and was used to direct transient expression of GUS in chloroplasts following biolistic bombardment of detached tobacco leaves.

Oligonucleotides for the PCR analysis of Nt-CPHb lines:

4.2.2. Oligonucleotides used in the PCR analysis of Nt-197Hb
In cases where the oligonucleotide was designed from the tobacco ptDNA (Shinozaki *et al.*, 1986) the position of the 5' nucleotide in ptDNA is given in brackets and C denotes that the identifying nucleotide is complementary to the nucleotide at that position in ptDNA.

1. **Hbα-For**: section 3.2.4
2. **Hbα-Rev**: section 3.2.4
3. **Hbβ-For**: section 3.2.4
4. **CP-For** (57521): 5'-GAATTCGTGTCGAGTAGACCTTG-3'
5. **CP-Rev** (60694) (C): 5'-CAATGGGATCTAGAGACCATG-3'
6. **rbcL2-For** (58952): 5'GCTGCTGCTTGTGAAGTATGG-3'
7. **aadA-Rev**: 5'-CAATGGTGACTTCTACAGCGCGG-3'

4.2.3. Preparation of tungsten particles for biolistic transformation
1ml of ice-cold 70% special grade ethanol was added to 30mg of accurately weighed tungsten powder (M-17, average particle size 1.1μm; Alpha) and the suspension was vortexed vigorously for 5 min followed by 15 min soaking on ice. The tungsten was then gently sedimented by spinning at 8000 rpm for 5 seconds in a microfuge and the ethanol was removed using a pipette. The tungsten particles were then washed three times by adding 1ml of sterile doubly distilled H2O, vortexing vigorously for 1 min and microfuging at 8000 rpm for 3
seconds before removing the H$_2$O. During the third wash a further 3 second 8000 rpm spin was performed to ensure as much as possible of the H$_2$O was removed. 500µl of sterile 50% glycerol was added and the suspension was stored at -20°C until use for a maximum period of one week.

4.2.4. Coating washed tungsten particles with DNA

The glycerol stock was vortexed vigorously for five minutes and 50µl of washed tungsten particles were removed to a 1.5ml eppendorf tube before the tungsten had a chance to settle. While vortexing vigorously the following were added to the tungsten aliquot in the order given: (1) 5µl of a 1-2µg/µl plasmid DNA solution; (2) 50µl of 2.5M CaCl$_2$ solution (3) 20µl of 0.1M spermidine (free base; tissue culture grade) solution. The plasmid DNA was purified using Qiagen column purification and the solution was microfuged for 5 min at 15000 rpm immediately before use to sediment residual insoluble matter. The CaCl$_2$ solution was prepared fresh, filter sterilised and stored on ice before use. The spermidine solution was prepared from a 1M stock solution that had been filter sterilised and stored at -20°C for a period no longer than one month. The tungsten particles were then vortexed vigorously for 30 min at 4°C. After vortexing the particles were allowed to settle for 1 min before being sedimented by microfuging for 2 seconds at 8000 rpm. The liquid was removed and discarded and 140µl of 100% special grade ethanol was added. The ethanol was removed without disturbing the sedimented tungsten and 48µl of 100% ethanol was added. The tungsten was gently resuspended by tapping the side of the tube and by low speed vortexing. The DNA coated particles were stored on ice until use, were used within a few hours of preparation and were resuspended by gentle vortexing before loading on the flying disc (macrocarrier).

4.2.5. Biolistic transformation of tobacco leaf tissue

Tobacco plants (Nicotiana tabacum cv. Petit Havana) were grown under sterile conditions as described in section 2.2.8. Healthy leaves 3-5cm in length from plants 4-5 weeks old were removed under sterile conditions and the abaxial side was bombarded. Biolistic transformation was carried out as described in the BioRad manual for the DuPont PDS1000/He Particle Delivery System. The following settings were used for plant leaf transformation unless otherwise stated;

(1) Helium pressure at the tank regulator: 1300-1400 psi
(2) 1100 psi rupture discs were used
(3) Flying disc assembly: level 2 from the top
(4) Flying disc (macrocarrier) was 6mm from the stopping screen
(5) Petri dish holder: level 2 from the bottom
(6) Vacuum at time of shot: 28 in. Hg
(7) The leaf was placed in a petri plate containing MS or RMOP medium to
  cushion the blast.

Following bombardment the petiole was pushed into the medium and
the plate was sealed and stored in the growth room for two days. At this point the
leaves were either incubated with substrate to test for GUS or celB expression or
they were prepared for the regeneration of stable transformants.

4.2.6. GUS and celB histochemical assays
Wild-type tobacco leaves were transiently transformed with pCTAK1 and
pSSHPrnGUS whereas leaves from the Nt-TGS14-7 F17 line were transiently
transformed with pSSHcelB(-). Following the two day incubation in the growth
room the bombarded leaves were vacuum-infiltrated at 20 kPa for 30 min with
staining solution (Silverstone, et al., 1997) containing 0.5mg/ml X-Gluc (5-
bromo-4-chloro-3-indolyl β-D-glucuronide) or 0.25mg/ml X-Gal (5-bromo-4-
chloro-3-indolyl-β-D-galactopyranoside) depending on whether the leaf had been
bombarded with a GUS or celB expressing construct. In the case of celB, the
leaves were heat killed before vacuum infiltration by immersing the leaves in
sterile salt solution (10mM CaCl2, 40 mM KCl), sealing the petri plate and
heating at 70°C in a water bath for 30 min. Leaves that had been bombarded with
pCTAK1 were incubated for 1 day at 37°C in the substrate buffer whereas those
bombarded with pSSHPrnGUS and pSSHcelB(-) were incubated for 4 days at
37°C. Following incubation with substrate buffer the leaves were bleached in
70% ethanol. Leaves were examined under an Olympus stereo microscope and
images were captured using a DP10 Olympus digital camera and manipulated
using Pixera studio and Microsoft PictureIT! Express software.

4.2.7. Regeneration of stable transformants
Following the two day incubation in the growth room on RMOP medium in the
absence of selection the bombarded leaves were cut into 5mm² sections and the
bombarded side was placed on shoot regenerating RMOP (Svab et al., 1990)
medium containing 500mg/L spectinomycin. Resistant tissue that appeared after
6-8 weeks was put through a further round of shoot regeneration before the
shoots were rooted in MS medium containing spectinomycin 500mg/L.
4.3 RESULTS

4.3.1. Transient expression of GUS in tobacco leaves
In order to optimise the biolistic procedure for delivering plasmid DNA into plant cells, GUS expression analyses were performed in transiently transformed cells in detached wild-type tobacco leaves. pCTAK1 (unpublished plasmid) was used as a marker of transient nuclear expression of GUS and pSSHPrnGUS (unpublished plasmid) was used as a marker of transient expression of GUS in chloroplasts. The plasmids were delivered into the tobacco leaf cells by particle bombardment. Of ten leaves bombarded with pCTAK1 four produced greater than 1000 visible blue spots (each denoting a leaf cell that had been transiently expressing GUS) after incubation overnight with GUS substrate buffer and bleaching the leaves with alcohol. Figure 4.1A shows a photograph of one of the tobacco leaves that was bombarded with pCTAK1. The frequency of transient GUS expression in leaves bombarded with pSSHPrnGUS was much lower than for leaves bombarded with pCTAK1 and the blue spots were only visible under magnification and after incubation in GUS substrate buffer for 4 days. The three leaves bombarded had 41, 7 and 4 detectable blue spots. Figure 4.1B, C and D shows photographs taken of magnified leaves showing the detection of GUS activity in tobacco leaves bombarded with pSSHPrnGUS.

4.3.2. T7 RNAP-dependent transient expression of celB in tobacco leaves
*celB* encodes a thermostable β-glucosidase that has a high level of β-galactosidase activity detectable by cleavage of X-Gal. In order to demonstrate T7 RNAP-dependent expression of *celB* in tobacco, detached leaves from tobacco expressing chloroplast targeted T7 RNAP (transgenic line Nt-TGS14-7F17) and non-transgenic tobacco (Nt) were bombarded with pSSHcelB(-). Bombarded leaves were incubated in the growth room for two days, heat-killed in order to destroy endogenous β-galactosidase activity and incubated for 4 days in X-Gal-containing buffer to assay for the expression of thermostable β-galactosidase activity. Alcohol-bleached leaves were examined for the presence of blue spots under high magnification. The photographs in Figure 4.2A and B show blue spots in leaves of Nt-TGS14-7F17 bombarded with pSSHcelB(-) presumably caused by cleavage of X-Gal by the celB-encoded activity. As in the case of leaves bombarded with pSSHPrnGUS the numbers of blue spots denoting regions expressing *celB* were few and the spots were only detectable under very
Figure 4.1. Transient expression of GUS in wild-type tobacco leaf cell cytoplasm and chloroplasts. A. Staining for GUS expression in tobacco leaf cells transiently transformed with pCTAK1. B, C and D. Staining for GUS expression in tobacco leaf chloroplasts transiently transformed with pSSHPrnGUS. The photographs in B, C and D were taken at x22, x28 and x57 magnification, respectively.
Expression of the celB encoded thermostable β-glucosidase activity from *Pyrococcus furiosus* in Nt-TGS14-7F17 leaves transiently transformed with pSSHcelB(-). A. Photograph of bleached leaf (x48 magnification) that was stained with X-Gal buffer. B. Photograph of a different section of another leaf (x57 magnification).
high magnification. No blue spots were detected in non-transgenic leaf tissue bombarded with pSSHcelB(-).

4.3.3. Stable transformation of tobacco chloroplast DNA with a dicistronic human hemoglobin operon.

The hemoglobin expression cassette was introduced into tobacco leaf chloroplasts using the biolistic process. Twenty-eight tobacco leaves (var. Petit Havana) were bombarded with tungsten particles (14 at 1100psi and 14 at 1350psi) coated with pCP197-Hb and were incubated for two days in the growth room before cutting into 5mm² leaf sections and plating on shoot regenerating medium containing spectinomycin. After about 6-8 weeks four shoots and two green calli appeared on two plates containing leaf tissue that had been bombarded at 1100 psi. Three of the four shoots were uniformly green and were growing from the bombarded (underside) surface of the leaf section that was in contact with the regenerating medium. The other shoot was colourless and was growing upwards from the side of a leaf section. Both green calli were growing on the bombarded side of the leaf.

The four shoots and two calli were placed on regenerating medium containing streptomycin for a second round of regeneration. Of the six putative independant transformants only tissue from the four shoots (designated C1-4) remained green and developed shoots on streptomycin. Regenerated shoots were cut for all four and placed in MS medium containing spectinomycin. Of these C1 which originated from the bleached shoot, rooted and grew normally, C2 clones rooted but had abnormally shaped leaves, C3 clones rooted and grew normally and C4 clones did not root and had abnormally shaped leaves.

In order to determine whether the spectinomycin and streptomycin-resistant shoots were genuine plastid transformants (Nt-197Hb) PCR analysis was performed on total cellular DNA. Because chloroplast transformation is mediated by homologous recombination between cpDNA sequences flanking the transgene in the transformation vector (i.e the targeting sequences) and the corresponding homologous sequences present on the plastid genome, it is possible to verify transgene integration by PCR using oligonucleotide primers based on sequences located adjacent to but outside the targeting sequences. With this in mind we designed the oligonucleotides CP-For and CP-Rev from cpDNA sequences that are not included in pCP197-Hb. Figure 4.3A shows a map of the predicted transplastome in Nt-197Hb, the positions of the oligonucleotides that
were used in the PCR analysis and the expected sizes of the PCR products. The homologous regions present in Nt-197Hb and pCP197-Hb are also shown.

Total cellular DNA for clones C1-4 was extracted from the streptomycin resistant callus tissue that resulted from the second round of regeneration and from wild-type tobacco (Nt) leaves. PCR reactions were performed on C1-4 DNA and Nt DNA using the following oligonucleotide primer pairs: (1) CP-For and aadA-Rev (Figure 4.3B), (2) rbcL2-For and aadA-Rev (Figure 4.3C), (3) Hbβ-For and CP-Rev (Figure 4.3D) and (4) Hbα-For and Hbα-Rev (Figure 4.3E). CP-For and aadA-Rev amplified the 2.2Kb predicted transplastome PCR product in C2, C3 and C4 reactions. No products were amplified in C1 or Nt reactions. RbcL2-For and aadA-Rev amplified the predicted 800bp transplastome PCR product for C2, C3 and C4 whereas a smaller incorrect (based on the pCP197-Hb sequence) 400bp product was amplified for C1. Hbβ-For and CP-Rev amplified the predicted 2.1Kb transplastome product for C3 only. Hbα-For and Hbα-Rev amplified a product corresponding in size to the Hbα coding sequence from C1, C2, C3 and C4. The C1, C2 and C4 reactions yielded a very low quantity of Hbα product compared to PCR amplifications using C3 DNA. For this reason and the fact that phenotypically normal, uniformly green, spectinomycin-resistant shoots were obtained that rooted well, the C3 line was investigated in more detail.

4.3.4. Inheritance of spectinomycin resistance in the C3 line
Three uniformly green shoots of the C3 line that rooted normally on MS medium containing 500 mg/L spectinomycin were planted in soil in the greenhouse. When they flowered, reciprocal crosses were carried out with an N. tabacum cv. Samsun NN line in order to investigate maternal inheritance of the spectinomycin resistance trait. When the C3 line was used as the female parent in the cross, progeny seedlings were all resistant (and uniformly green) on medium containing spectinomycin. In contrast, when the C3 line was used as the male parent, all progeny seedlings were spectinomycin-sensitive (bleached white) (data not shown). Thus, the spectinomycin resistance phenotype of the C3 line shows maternal inheritance.

4.3.5. A cross between the C3 transplastomic line and a nuclear transgenic line expressing the ST14 fusion protein produces a novel yellow-bleached phenotype
Seeds were obtained from a cross (C3/ST14) in which the male parent was the ST14-expressing line Nt-TGS14-7 (cv. Samsum NN) and the C3 line (cv. Petit
Figure 4.3 A. Schematic showing the transplastome in Nt-197Hb. The positions of oligonucleotides used in the PCR analysis and the size of the products are shown. The sequence in common between pCP197-Hb and the transplastome is also shown. B, C, D and E. PCR analysis of tobacco clones C1-C4 putatively transformed with pCP197-Hb. B. Oligonucleotides CP-For and aadA-Rev. Lanes 1-4, PCR reactions containing total cellular DNA from C1-C4, respectively; lanes 5 and 6, control PCR reactions containing total cellular DNA from wild-type tobacco (Nt); lane 7, Kb ladder. C. Oligonucleotides rbcL2-For to HbP-Rev. Lanes 1-4, Ka ladder; lane 2, Nt control; lanes 3-6, C1-C4 reactions, respectively. D. Oligonucleotides Hb-For and CP-Rev. Lanes 1-4, C1-C4 reactions; lanes 5 and 6, Nt controls; lane 7, Kb ladder. E Oligonucleotides Hb-For and Hb-Rev. Lane 1, Kb ladder; lane 2, Nt control; lanes 3-6, C1-C4 reactions.
Figure 4.4. Photograph of ST14, C3 and C3/ST14 (C3 as female and ST14 as the male in the cross) seedlings plated on MS medium containing 3% sucrose. A. C3 selfed growing in the presence of spectinomycin 500mg/L. B. ST14 seedlings growing in the presence of spectinomycin 500mg/L. C. C3/ST14 seedlings growing in the presence of spectinomycin 500mg/L and kanamycin 50mg/L. D. C3/ST14 seedlings growing with no antibiotic selection.
Figure 4.5. C3/ST14 phenotype on soil. Row 1, progeny of the C3 line (as female) crossed with wild-type Samsun (as male); row 2, progeny of selfed C3 line; rows 3 and 4, C3/ST14 hybrids.
Havana) was the female parent. It was expected that the resulting progeny would be both kanamycin and spectinomycin resistant having inherited the former from the male parent and the latter from the female parent. The resulting seeds were plated on MS medium containing either kanamycin alone or spectinomycin and kanamycin together or no antibiotics. All seedlings plated under these three conditions grew with a yellow-bleached phenotype i.e. even in the absence of antibiotics. Furthermore, when seeds were germinated on soil, the same yellow-bleached phenotype was observed. This phenotype was judged to be ST14-dependent for the following reasons: (i) in the absence of antibiotics, neither of the parental lines or their selfed progeny showed the yellow-bleached phenotype and (ii) seedlings from a cross between the C3 line (as the female) and a non-transgenic cv. Samsun NN line (as the male) were fully green in the presence or absence of antibiotics. Figure 4.4 shows the C3/ST14 seedlings growing on kanamycin and spectinomycin together and in the absence of antibiotics. In addition, when plated on medium containing spectinomycin, selfed C3 seedlings grew normally and were green (Figure 4.4A) whereas seedlings of the Nt-TGS14-7 are bleached and slow growing (Figure 4.4 B). It was also clear that the ST14-dependent phenotype differed from the spectinomycin-sensitive phenotype because the C3/ST14 seedlings are not as completely bleached (yellow-bleached vs. white bleached) and growth is not as severely retarded as Nt-TGS14-7 seedlings (Figure 4.4 C and D). In some cases, individual C3/ST14 plants developed green sectors and these seedlings grew better than their uniformly bleached siblings (Figure 4.4D). When grown on soil the C3/ST14 plants germinated, were bleached and did not grow past the 2 cotyledon stage (Figure 4.5). This was in contrast to their continued slow growth on antibiotic-free medium containing sucrose.
4.4. DISCUSSION

4.4.1. Detection of T7 RNAP-dependent celB expression in transiently transformed leaf cells

Because particle bombardment of leaves yields very few transformant chloroplasts in a given tissue sample it is necessary to be able to detect transformation by the expression of a reporter gene. Transient transformation of chloroplasts expressing CAT, GUS and GFP has been previously reported (Daniell et al., 1990; Ye et al., 1990; Hibberd et al., 1998). We have demonstrated the expression of a thermostable β-glucosidase encoded by the celB gene of Pyrococcus furiosus in transiently transformed leaves from tobacco expressing a chloroplast targeted T7 RNAP. Because the pattern of celB expression is similar to the expression of GUS from the chloroplast-specific Prrn promoter this result is consistent with the expression of celB in the chloroplast. This demonstrates the potential for celB expression in stably transformed chloroplasts.

4.4.2. Stable transformation of tobacco chloroplast DNA with the genes for adult human hemoglobin

We have obtained using the biolistic process a single transplastomic tobacco line, Nt-197Hb (clone C3), that contains an adult human hemoglobin expression cassette linked to a chimeric spectinomycin resistance (aadA) gene. This line was one of four that was resistant to both spectinomycin and streptomycin. However this was the only line that grew normally in tissue culture and gave the predicted signals for the transplastome in all PCR analyses that were performed. Also the use of ptDNA oligonucleotides that are located adjacent to but outside of the ptDNA sequences in pCP197-Hb gives confidence that the Nt-197Hb line is a genuine chloroplast transformant. The Nt-197Hb line has been taken through three rounds of shoot regeneration on selectable medium which is reported to be more than sufficient to obtain uniformly transformed (homoplasmic) plants using aadA based constructs (Staub et al., 2000).

Recent developments have demonstrated the feasibility of the production of therapeutic proteins in transplastomic plants and high lighted the advantages this system has over production in nuclear transformants (Staub et al., 2000). Staub et al. (2000) generated transgenic tobacco that produced high levels of biological active somatotropin, a human therapeutic protein, from a plastid
transgene. Somatrophin is currently produced in bacterial systems and is used to treat dwarfism in children and potentially has many other applications in human medicine. The chimeric somatotropin gene that was expressed to the highest level in transgenic chloroplasts was under the expressional control of the chloroplast ribosomal RNA operon promoter (Prm), the bacteriophage T7 gene 10 leader sequence (G10L) and the stability element of the chloroplast rps16 gene.

The hemoglobin expression cassette in Nt-197Hb is very similar but contains the T7 promoter instead of the Prm promoter. It is encouraging that the expression cassette containing G10L was associated with the highest level of somatotropin production in transgenic plants. Another expression cassette which contained controlling elements exclusively from chloroplast genes was many times less efficient in expressing somatotropin in transgenic chloroplasts. Therefore it appears that the 63 nucleotide G10L region which is well adapted for initiation of translation in E.coli is also very efficient in promoting translation in chloroplasts.

It would be of interest to compare the performance of our expression system in the production of hemoglobin especially considering that the hemoglobin genes in pCP197-Hb have not been re-synthesised to suit expression in prokaryotic systems as was the case for somatotropin produced in transgenic chloroplasts. It has been found that hemoglobin genes re-synthesised using E. coli preferred codons are much more highly expressed in E.coli than the unmodified genes (Hernan et al., 1992). It may turn out that it is not necessary to re-synthesise eukaryotic genes in order to achieve high levels of expression in transgenic chloroplasts.

It is interesting to speculate on the cause of the ST14-dependent phenotype of C3 plants when crossed with the ST14 expressing line. The yellow-bleached phenotype might be a consequence of T7 RNAP (ST14)-mediated transcription of the Hb expression cassette e.g. if the T7 terminator sequence is inefficient, transcriptional readthrough into adjacent chloroplast genes (perhaps producing antisense transcripts) might be deleterious for chloroplast function. Alternatively, the yellow-bleached phenotype might be caused by a hemoglobin-specific effect in chloroplasts. Both chlorophyll and chloroplast heme are synthesized entirely within the plastid using the same precursor compound protoporphyrin IX (Beale and Weinstein, 1990). The insertion of a magnesium ion or a ferrous ion into protoporphyrin IX by specific chelatases produces either Mg-protoporphyrin IX which is ultimately converted into chlorophyll, or
protoheme which is ultimately converted into heme. It is possible that the C3/ST14 hybrid plants contain a level of α and β-globin sufficient to sequester most of the available heme and that this either deprives chloroplast cytochromes of heme and/or channels protoporphyrin IX into the heme biosynthetic pathway at the expense of the chlorophyll biosynthetic pathway. This might be sufficient to account for the resulting yellow-bleached phenotype.
GENERAL CONCLUSIONS AND FUTURE PROSPECTS

In chapter 2 we described the successful production of transgenic tobacco plants expressing high and low levels of chloroplast-targeted T7 RNAP activity in leaves. We generated three discrete transgenic lines each expressing slightly different ST fusion proteins containing 8, 14 or 25 residues of mSSU translationally fused to the N-terminus of T7 RNAP. The transgenic plants appeared normal, these transgenes were stably inherited in progeny lines, homozygous lines were identified and we determined that the ST25 fusion protein had a half-life of at least 16 days in the dark. We found no apparent differences in the transcription of eight plastid genes in wild-type tobacco leaves and in the leaves of a plant which contained T7 RNAP activity within the chloroplast compartment. For each ST expressing line we determined that a large proportion of total cellular T7 RNAP activity was located within the chloroplast. We have tentatively determined that the ST fusion protein containing 8 mSSU residues (ST8) was the most efficiently targeted to the chloroplast followed by ST14 and ST25. We note however that in vitro import experiments would be required to conclusively determine targeting efficiency.

In chapter 3 we described the modification of the chloroplast transformation vectors pZS197 (Svab and Maliga, 1993) and pSSH-B1 (an unpublished derivative of pSSH1(Kavanagh et al., 1994) which contains a smaller plastid DNA insert) to give pCP197 and pSSH-B1a, respectively. The modification involved the introduction of multiple cloning sites which enabled the insertion of expression cassettes into pCP197 alongside the aadA expression cassette and into an intergenic region of the plastid DNA fragment in pSSH-B1a. In addition we modified the phage T7 gene 10 (T7G10) expression cassette for the expression of chimeric genes in tobacco chloroplasts by introducing the 3' UTR of the plastid rps16 gene upstream of the T7 transcription terminator sequence.

Into the modified T7G10 expression cassette we introduced the genes encoding the following; (i) a thermostable β-glucosidase and a thermostable β-glucanase from the hyperthermophile Pyrococcus furiosus (ii) single chain camel antibodies against chicken egg lysozyme (cAb-Lys3) and tetanus toxin (cAb-TT2) and (iii) adult human hemoglobin (HbA). Most of the resulting
chimeric genes were transferred into both pCP197 and pSSH-B1a. We detected the expression of cAb-Lys3 and HbA in *E. coli* by western blotting but could not conclude that the expression was T7 RNAP-dependent because of the constitutive nature of gene expression from the T7G10 expression cassette in pETCP5 and pSSH-B1a (+) (even in *E. coli* strains not producing T7 RNAP). However we were able to show T7 RNAP-dependent expression in *E. coli* of GFP in the modified T7G10 expression cassette when the chimeric GFP gene was introduced into pSSH-B1a (-). Overall the above findings suggest that the modified T7G10 expression cassette is functional in *E. coli* in a T7 RNAP-dependent manner. In addition we have determined the orientation in which expression cassettes are transcriptionally silent in pSSH-B1a. This knowledge will be useful in the future testing of expression cassettes in *E. coli*. Also in this work we have produced the material that will enable the investigation of the feasibility of expressing several genes of interest in the chloroplasts of transgenic plants.

In chapter 4 we described the successful use of the biolistic procedure in demonstrating the T7 RNAP-dependent expression of the cel B gene of *Pyrococcus furiosus* in transiently transformed chloroplasts. In addition we described the generation of transplastomic tobacco plants (C3 line) stably transformed with a dicistronic adult human hemoglobin expression cassette. We concluded that the C3 line was a genuine chloroplast transformant because of the results of PCR analyses using oligos which anneal to cpDNA sequences not found in the pCP197-Hb transformation vector and because the spectinomycin resistance phenotype of the C3 line showed maternal inheritance. Furthermore the uniformity of the spectinomycin resistance in C3 progeny suggest that the parental line is homoplasmic for the transgene. Interestingly we found that the progeny of a cross between the C3 transplastomic line (as the female) and a nuclear transgenic line expressing the ST14 fusion protein (as the male) produced a novel yellow-bleached phenotype. The C3/ST14 hybrids grew slowly on both kanamycin and spectinomycin when sucrose was supplied in the medium and the phenotype was present even in the absence of selection. We also observed that some of the C3/ST14 hybrids developed green sections which became more prominent over time. This may have been caused by excision of the Hb expression cassette from plastid DNA or it may indicate that some wild-type plastomes persist in the C3 lines. It will be interesting to determine whether the C3 lines used in the cross were homoplasmically transformed as determined by PCR analysis or southern blotting.
The ST14 dependence of the C3/ST14 phenotype (see section 4.3.5) strongly suggests that both the T7 RNAP activity and the dicistronic adult human hemoglobin operon are interacting transcriptionally in the chloroplasts of these plants. It will be interesting to determine whether these plants are producing hemoglobin in the chloroplast and at what level. In addition, it would be particularly interesting to investigate the possibility of rescuing these plants by adding heme and/or iron chelate supplements to the growing medium. Furthermore, it may be possible to produce high levels of HbA in the chloroplasts of transgenic plants and avoid the C3/ST14 phenotype by crossing our C3 line with a line producing a chloroplast-targeted T7 RNAP activity from a tightly regulated promoter which can be induced in mature leaves.
References


Bligny, M., Courtois, F., Thaminy, S., Chang, C-C., Lagrange, T., et al. (2000). Regulation of plastid rDNA transcription by interaction of CDF2 with two different RNA polymerases. EMBO J. 19, 1851-60.


