

## **INFORMATION TO USERS**

**This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.**

**The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.**

**In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.**

**Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps.**

**Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.**

**ProQuest Information and Learning  
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA  
800-521-0600**

**UMI<sup>®</sup>**





**Université d'Ottawa • University of Ottawa**



**EXPRESSION OF BIOLOGICALLY ACTIVE HUMAN  
GRANULOCYTE MACROPHAGE COLONY STIMULATING  
FACTOR IN THE SEEDS OF TRANSGENIC TOBACCO**

**Zaman Alli**

**Thesis submitted to the Department of Biochemistry,  
Microbiology, and Immunology in partial fulfillment of the  
requirements for the degree of Masters in Science**

**University of Ottawa  
Ottawa, Ontario, Canada  
September, 2000**

**© Zaman Alli, Ottawa, Canada. 2000**



**National Library  
of Canada**

**Acquisitions and  
Bibliographic Services**

**395 Wellington Street  
Ottawa ON K1A 0N4  
Canada**

**Bibliothèque nationale  
du Canada**

**Acquisitions et  
services bibliographiques**

**395, rue Wellington  
Ottawa ON K1A 0N4  
Canada**

*Your file Votre référence*

*Our file Notre référence*

0-612-66000-1

**The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.**

**The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.**

**L'auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.**

**L'auteur conserve la propriété du droit d'auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.**

## ABSTRACT

The feasibility of producing recombinant (rt) and biologically active granulocyte-macrophage colony stimulating factor (GM-CSF) in the seeds of transgenic tobacco plants was investigated. The rice seed-specific glutelin promoter (Gt1) was used to direct the expression of the human (h) GM-CSF coding sequence in tobacco seeds. Transgenic tobacco plants producing rthGM-CSF were compared in biological assays with tobacco plants expressing a glutelin/rthGM-CSF fusion protein, driven by the Gt3 promoter. The glutelin/rthGM-CSF fusion construct encoded the first eight N-terminal amino acids from the rice glutelin-3 protein at the N-terminal of the mature hGM-CSF sequence.

The T7 Sequencing kit from Pharmacia was used to sequence and confirm the authenticity of the Gt1 expression construct. The glutelin-1 signal sequence was fused in the correct orientation to the hGM-CSF cDNA. The rthGM-CSF expression cassette (2.5 kb) was subcloned in a plant binary vector pRD400, which contained a kanamycin resistance gene. The pRD400 vector containing the 2.5 kb construct was used to transform *Agrobacterium tumefaciens* cells. Tobacco (*Nicotiana tabacum* cv. Xanthi) leaf sections were transformed by *A. tumefaciens* carrying the complete 2.5 kb construct. Transformed tobacco cells grew on kanamycin medium. Calli developed from these cells and eventually matured into flowering, fertile plants. Flowers were self-pollinated and seeds collected at different days after pollination.

Enzyme-linked immunosorbent assay (ELISA) data from thirty-two Gt1 tobacco plants were obtained. ELISA did not detect any rthGM-CSF in transgenic tobacco leaf extracts or non-transformed seed extracts. Transgenic tobacco seed extracts contained

rthGM-CSF up to a level of 0.03% of total extracted protein. The highest producing rthGM-CSF plants were used for further studies. The monocot glutelin-1 promoter was confirmed as being seed specific in tobacco plants.

Southern analysis of transgenic plants detected a 2.5 kb band corresponding to the original expression construct. A copy reconstruction experiment on two of these plants detected at least three copies of the gene in the highest expressers of rthGM-CSF protein while only one copy was present in a low expresser transgenic tobacco plant. RT-PCR data indicated rthGM-CSF transcripts were present in transgenic plants but absent in non-transgenic plants. Western blots conducted on protein extracts from seeds of transgenic plants showed bands of 19, 21 and 50 kDa.

Biological assays compared the *in vitro* activities of the two tobacco-produced rthGM-CSFs with that of the *Escherichia coli* produced rthGM-CSF protein. Seed extracts containing the rthGM-CSF and the fusion glutelin/rthGM-CSF proteins, as well as the *E. coli* derived rthGM-CSF showed an approximately equal *in vitro* proliferative effect on a TF-1 human GM-CSF dependent cell line. However, addition of non-transformed plant extracts to the assay containing the *E. coli* rthGM-CSF resulted in a five-fold decrease in the proliferation of TF-1 cells. It appeared that the rthGM-CSF produced in the seeds of tobacco plants was glycosylated and this may have stabilized its activity in the *in vitro* biological assay. These findings, which suggest that seed-produced rthGM-CSF is stable and biologically active, are discussed in relation to the global need for low-cost recombinant cytokines.

## TABLE OF CONTENTS

Abstract .....	ii
Table of Contents .....	iv
List of Figures .....	viii
List of Tables.....	ix
Acknowledgments.....	x
Abbreviations .....	xi
Introduction .....	1
Overview.....	1
Classical expression hosts: microbial and mammalian transgenic expression systems.....	3
Recombinant DNA transfer technology: <i>Agrobacterium</i> <i>tumefaciens</i> transformation .....	5
Plants as bioreactors .....	6
Advantages of using plant systems as “pharming platforms” .....	7
Seed storage proteins: glutelin gene regulation and secretion of proteins.....	9
Seed protein homologies .....	9
Glutelin gene regulation .....	9
Glutelin processing and deposition .....	11
Advantages of using rice glutelin promoters for recombinant protein production...	11
Some major groups of proteins successfully produced in plant systems .....	12
Plant derived antigens.....	12
Production of recombinant antibodies in plants .....	13
Granulocyte macrophage colony stimulating factor.....	13
Human GM-CSF receptors .....	15

Transduction of GM-CSF signal into the nucleus .....	16
Clinical applications of rthGM-CSF .....	16
Recombinant hGM-CSF from different expression systems .....	17
Recombinant hGM-CSF from yeast systems .....	17
Recombinant hGM-CSF from baculoviral systems .....	18
Recombinant hGM-CSF from mammalian cell cultures .....	18
Effects of glycosylation on recombinant proteins: GM-CSF .....	19
Application of glutelin promoters in prior seed specific expressions .....	20
Glycoprotein B complex in transgenic tobacco seeds.....	20
Antecedent glutelin/hGM-CSF fusion peptide in tobacco seeds .....	21
Rationale and objectives .....	23
Materials and Methods.....	24
Materials.....	24
Glutelin-1 promoter .....	24
Glutelin-1 signal sequence .....	24
Mature hGM-CSF coding sequence.....	24
Nopaline synthase sequence for polyadenylation.....	25
Tobacco leaves.....	25
Tobacco transformation vector.....	25
Chemicals and enzymes .....	25
Methods.....	26
Sequencing of the Gt1 construct .....	26
Preparation and electrophoretic conditions for sequence analysis on polyacrylamide gels.....	27
Transformation, selection and growth of <i>E. coli</i> .....	28

Preparation of competent <i>A. tumefaciens</i> cells .....	29
Transformation of <i>A. tumefaciens</i> cells.....	29
Inoculant preparation .....	30
Non-transformed tobacco leaf sterilization.....	30
Transformation of sterile leaf sections .....	31
Calli induction and regeneration of transgenic plants .....	31
DNA extraction from tobacco leaves.....	32
Restriction analyses on DNAs.....	33
Agarose gels for DNA electrophoreses.....	33
Southern analysis of total genomic DNA .....	34
Extraction and purification of DNA from bacteria.....	35
Generation of the probe and genecloning.....	35
Radiolabelling of the probe .....	36
Prehybridization and hybridization for Southern blots.....	37
Collection of seeds.....	37
Extraction of total seed RNA .....	38
Preparation of RNA gels.....	39
Reverse transcriptase (RT) and polymerase chain reaction (PCR)	
{RT-PCR} experiments.....	39
Reverse transcriptase experiment.....	39
Polymerase chain reaction experiment .....	40
Extraction of total seed protein.....	41
Ammonium sulphate precipitation.....	41
Sodium dodecyl sulphate-polyacrylamide gel electrophoresis	
(SDS-PAGE).....	42

Protein transfer.....	43
Probing and detection of rthGM-CSF .....	44
Human GM-CSF-specific ELISA.....	45
Biological assays.....	45
Results .....	47
Sequence and restriction analyses on the Gt1 rthGM-CSF expression construct.....	47
Transgenic tobacco plants, integration of hGM-CSF DNA and its transcription .....	52
Selection for acquired antibiotic resistance in calli and plantlets .....	52
Southern analyses .....	52
Correlation of rthGM-CSF expression at the genomic level .....	56
Analyses on transgene transcription .....	56
Human GM-CSF-specific ELISA and Western blot analyses on seed derived rthGM-CSF.....	59
ELISA detection of rthGM-CSF from the seeds of transgenic tobacco plants .....	59
Western blot analysis.....	62
Biological assessment of the seed-derived rthGM-CSF proteins.....	65
Discussion .....	68
Analysis of transgene(s) .....	68
Analysis of recombinant hGM-CSF .....	71
Conclusion.....	79
Future directions .....	80
REFERENCES .....	81

## LIST OF FIGURES

<b>Figure 1.</b> Recombinant hGM-CSF expression constructs .....	49
<b>Figure 2.</b> Sequencing confirmation of construct integrity. ....	50
<b>Figure 3.</b> The coding sequences of the Gt1 expression construct. ....	51
<b>Figure 4.</b> Continuous selection for transformed tobacco leaf tissues. ....	54
<b>Figure 5.</b> Genomic organization of three independent transgenic events. ....	55
<b>Figure 6.</b> RT-PCR detection of the rthGM-CSF mRNA transcripts in seeds of transgenic tobacco plants.....	58
<b>Figure 7.</b> A standard curve for the estimation of rthGM-CSF in the seed extracts of transgenic tobacco plants. ....	60
<b>Figure 8.</b> Immunoblot detection of the partially purified, ammonium sulphate precipitated, rthGM-CSF from tobacco seeds. ....	63
<b>Figure 9.</b> Loading of protein samples on SDS gel. ....	64
<b>Figure 10.</b> Biological activity of seed derived rthGM-CSF from Gt1 and Gt3 plants after 48 hours.....	66
<b>Figure 11.</b> The rthGM-CSF expression construct is stably inherited in the second generation.....	67

LIST OF TABLES

**Table 1.** Human GM-CSF immunoassay on rthGM-CSF extracted from tobacco seeds.....61

## ACKNOWLEDGMENTS

I thank all those involved in my academic career for their patience, understanding and support. I wish to gratefully acknowledge my supervisors' expert guidance and assistance throughout my research years, in particular, Dr. Illimar Altosaar and Dr. Peter R. Ganz. Dr. Ravinder Sardana was always willing to give technical advice and directions. Further gratitude is extended to my former and current thesis advisory committee members Dr. Denis G. Williamson, Dr. Anil Dudani and Dr. Eileen Tackaberry. My appreciation extends to the technical staff that assisted with the experiments, Connie Sauder, Gary Rozon, Sue Porter and Danielle Bowness. This work was supported by the Canadian Red Cross (now Canadian Blood Services), the Natural Sciences and Engineering Research Council of Canada and Monsanto Canada.

## ABBREVIATIONS

<b>AP</b>	Alkaline phosphatase
<b>BA</b>	6-benzyladenine
<b>BCIP</b>	5-bromo-4-chloro-3-indolylphosphate p-toluidine salt
<b>Bis-Acrylamide</b>	N,N'-methylene bisacrylamide
<b>bp</b>	Base pair
<b>CaMV</b>	Cauliflower mosaic virus
<b>cDNA</b>	Complementary DNA
<b>DAP</b>	Days after pollination
<b>DEPC</b>	Diethylpyrocarbonate
<b>DTT</b>	Dithiothreitol
<b>EDTA</b>	Ethylene diamine tetraacetic acid
<b>ER</b>	Endoplasmic reticulum
<b>GUS</b>	$\beta$ -glucuronidase
<b>kDa</b>	Kilo Dalton
<b>LB</b>	Luria Bertani medium
<b>MW</b>	Molecular weight
<b>mRNA</b>	Messenger RNA
<b>MS</b>	Murashige and Skoog medium
<b>NAA</b>	$\alpha$ -naphthalene acetic acid
<b>NBT</b>	Nitroblue tetrazolium chloride
<b>ng</b>	Nanogram
<b>NOS-TER</b>	Nopaline synthase terminator
<b>nt</b>	Nucleotide
<b>PCR</b>	Polymerase chain reaction
<b>PVPP</b>	Polyvinyl-polyrrolidone
<b>rpm</b>	Revolutions per minute
<b>RT-PCR</b>	Reverse transcriptase-polymerase chain reaction
<b>SDS</b>	Sodium dodecyl sulfate
<b>SSC</b>	Sodium chloride and sodium citrate solution
<b>TAE</b>	Tris-acetic acid-EDTA buffer
<b>TBS</b>	Tris-buffered saline
<b>T-DNA</b>	Transfer DNA
<b>TEMED</b>	N,N,N',N'-tetramethylethylenediamine
<b>TRIS</b>	Tris (hydroxymethyl)aminomethane
<b>YEP</b>	Yeast extract, peptone and NaCl medium

## Chapter 1

### INTRODUCTION

#### Overview

Historically, when one needed a protein drug of animal origin, like insulin or interferon, the only choice was to extract it from the original same tissue like dog pancreas or human blood. With the advent of recombinant DNA technology, it is possible to produce the therapeutic protein in several expression systems including bacteria, yeast, animals, insect cells and plants (Ganz *et al.*, 1996 and references therein). For recombinant protein production, promoters usually drive transgene-coding sequences from the new host organisms. For instance, if a human coding sequence is to be expressed in a plant system, a promoter from plants (or plant virus) may be used in the recombinant expression construct. Stable expression is possible if an organism incorporates the recombinant DNA into its genome. An expression plasmid, for example in bacteria, may need an antibiotic resistant gene for maintaining selection for expression. Once a protein of interest is produced, it can be purified from the recombinant source or production organism and administered during therapy. Insulin was the first recombinant therapeutic protein that was produced in bacteria (Bourgaize *et al.*, 2000).

Bacteria belong to the prokaryotic kingdom, implying that they are mainly simple, single celled organisms. Eukaryotic organisms are extremely more complex than bacteria. For example, heteronuclear messenger RNA requires a series of modifications before it could leave the nucleus of a eukaryotic cell, for subsequent translation into a protein. When structural and functional relations are compared, bacteria do not need the complex mRNA and peptide processing enzymes as eukaryotes do. Therefore, if a therapeutic

protein, such as human erythropoietin, requires glycosylation for its *in vivo* activities. bacteria will not serve as a suitable recombinant production system (Tsuda *et al.*, 1990).

Recombinant proteins were first produced in bacteria and later, eukaryotic systems were investigated. Yeast and mammalian cells allowed for glycosylation of recombinant proteins that is not possible with bacterial systems, however, the yields were low. Further, like bacterial systems, mammalian cell culture systems required constant temperature, pH, and oxygen levels. This meant that scale-up of bioreactors was difficult. Plant systems were investigated in an attempt to eliminate some of the disadvantages associated with the earliest recombinant systems (Ganz *et al.*, 1996).

The first attempt to produce a recombinant antibody in a plant system was conducted in 1989 (Hiatt *et al.*, 1989). Transformation of tobacco leaf tissues leads to stable expression that resulted in a 1.3% yield of antibodies in the total soluble leaf proteins. Since then, plant systems have been continuously investigated in feasibility studies to produce a wide range of medical and industrial proteins. Some medically important proteins that were synthesized in plants include human insulin (Arakawa *et al.*, 1998b) and erythropoietin (Matsumoto *et al.*, 1995). Upon testing, these plant-produced proteins were found to have biological activities equivalent to the native proteins. This is why many life science companies (e.g. Monsanto, Pharmacia, Mogen) have commercial interests in transgenic research.

Various cDNAs for human cytokines have been harvested from the natural source or cloned and expressed in several recombinant expression systems. The purified cytokines were tested to better understand the molecular basis of diseases and this led to the use of cytokines for therapeutic applications. However, the costs associated with obtaining

cytokines have significantly impaired discoveries of additional potential applications of these cytokines. Therefore, the supply of recombinant cytokines is considered insufficient and limiting requiring considerable and immediate efforts to innovate new production strategies (Ganz *et al.*, 1996).

The production of human granulocyte macrophage colony-stimulating factor (GM-CSF) in plants was postulated as a prospective abundant source. Prior work by this laboratory used the glutelin-3 promoter to drive the synthesis of a syllogistic peptide in tobacco plants. This new word was coined to reflect the N-terminal fusion peptide that allowed for mimicking of a plant-seed storage protein thereby conferring preferential processing and stability of the fusion peptide in the plant environment. This syllogistic strategy may also be referred to as an antecedent strategy since the mature protein consists of an N-terminal syllogen. The resulting protein synthesized in plant seeds was a truncated glutelin protein fused to the human (h) GM-CSF mature protein. The fusion peptide was found to be biologically active (Ganz *et al.*, 1996). This is where my project started and it involved the production and assessment of the stability of the mature hGM-CSF in tobacco seeds without the syllogistic peptide. For this purpose, a new recombinant expression construct for hGM-CSF was initiated.

### **Classical expression hosts: microbial and mammalian transgenic expression systems**

Protein supplement therapy requires production and purification of proteins from heterologous recombinant systems, be they bacterial, viral, yeast, animal or plant systems. One must consider the system carefully, however, to deliver cheap, safe, biologically active therapeutic agents. Many eukaryotic proteins produced in *E. coli* were found to mimic the effects of the native proteins both *in vitro* and *in vivo*. However, several reports have also

described the aggregation of recombinant proteins in bacteria as inclusion bodies, and this required additional costs in refolding (Burgess *et al.*, 1987). In addition, bacteria cannot glycosylate proteins. Both bacterial systems and animal cell culture systems share disadvantages in the requirement for growth media and controlled environments. Human systems, for example, blood fractionation systems, have fluctuations in supplies as well as the possibility of contaminations by human specific pathogens (Ganz *et al.*, 1996). Yeast systems may have additional glycosylation recognition sites leading to incorrectly glycosylated proteins (Ernst *et al.*, 1987).

Mammalian expression systems can accomplish complex post-translational steps including glycosylation and correct folding of proteins to produce biologically active peptides. This is a major advantage over bacterial systems, mammalian expression systems however have several inherent disadvantages at the core of the system. Slow cell multiplication results in low yield while maintenance costs are high; human pathogenic organisms could thrive in the cell culture medium and product contaminants albeit trace may infect patients; lack of cell walls make mammalian cells more susceptible to stirring damage (shear forces) resulting in cell death; additional inputs including fetal bovine serum and the need for purification of recombinant products result in higher prices. Complications arise during scale-up of bioreactors because it is difficult to maintain constant pH, temperature, the level of metabolic wastes and oxygen levels (Ganz *et al.*, 1995, 1996). Therefore, these classical expression systems may not be competitive in terms of costs and yields for bioactive peptides if a suitable plant host is found.

The complete human genomic sequence will help to rapidly identify diseases with correlation of the defective genes. Hence, an upsurge in demand for more human therapeutic proteins is impending. Recombinant systems capable of rapid scale-up, high

yields of safe biologically active human proteins will be best suited for this task. To date, there exists no transgenic expression system that results in inexpensive, safe and biologically active human therapeutics. Plant-based genetic expression systems may serve as a vehicle for the biological assembly and production of human therapeutic proteins for global humanitarian purposes.

### **Recombinant DNA transfer technology: *Agrobacterium tumefaciens* transformation**

A recombinant expression vector usually consists of several regulatory sequences: promoter, translational enhancer, targeting signals, coding region, terminator and polyadenylation signal. Once assembled, and the DNA sequence of the construct verified by sequencing, it may be introduced into a host organism via several available transformation techniques such as microinjection, electroporation, viral delivery, particle gun bombardment or via *Agrobacterium tumefaciens*. Gene transfer methods in plants were reviewed by Walden *et al.*, 1995.

Improvements in transformation vectors lead to the likelihood of transforming both monocots and dicots (Walden *et al.*, 1995; reviewed by Lindsey, 1996). While other transformation methodologies are possible, the use of *A. tumefaciens* is now routine for the invention of transgenic crops. In this system, the bacterial tumour-inducing (ti) plasmid contains a segment of DNA (transfer DNA or T-DNA) that normally integrates randomly into the genomes of host-specific plants. The gene products encoded by the regions outside the T-DNA borders of ti-plasmid allow for the genomic integration in plants. Wild-type *A. tumefaciens* infection of roots of plants leads to tumours. The host plant then synthesizes opines (nopaline and octopine). Modification of the T-DNA plasmid of *A. tumefaciens* allowed for the insertion of recombinant expression constructs between the T-DNA borders of binary vectors (Datla *et al.*, 1992). Additional antibiotic resistant gene

insertions allow for the selection of cells that have integrated the genes of interest into their genomes.

The initial transformation step involves transforming *A. tumefaciens* with a binary vector containing the gene of interest that is then co-cultivated with damaged plant tissues. Injured leaves seem to attract the bacterium to the wound site. The region contained between the T-DNA gets randomly integrated into the host genome (Horsch *et al.*, 1985). Selection for transformed cells is possible when an antibiotic resistance gene such as kanamycin is included between the T-DNA regions. Transformed cells grow into a clump of undifferentiated cells known as a callus. Callus tissues may eventually differentiate and grow into shoots. From these shoots, transgenic plants will later develop. Unfortunately, successive regeneration of plantlets via tissue culture has not been predictable in all species, and slow progress has been accomplished on a gene-by-gene, plant-by-plant basis.

### **Plants as bioreactors**

Plants are the main source of the world's staple diet. Fruits, leaves, roots and seeds are the storage sites for the deposition of a variety of nutrients like carbohydrates, proteins, and vitamins. Food crops such as maize, rice and wheat are easy to propagate and the biomass obtained from these fields is low-cost (Cramer *et al.*, 1996). If a recombinant protein is expressed in the leaves of grain crops, it is possible to enhance the returns to the farmer if both the leaves and grains are utilized. Applied biotechnology has been used to produce many plant-based proteins: viral antigens, entomocidal proteins, human and industrial enzymes and cytokines (Herbers *et al.*, 1995; Cramer *et al.*, 1996; Ganz *et al.*, 1995, 1996). Specifically, some examples of human blood proteins that have been successfully expressed in plants are albumin, erythropoietin (Epo), granulocyte macrophage colony stimulating factor (GM-CSF), growth hormone, human protein C,

insulin, interferon, somatostatin, and tissue plasminogen activator (Cramer *et al.*, 1996; Ganz *et al.*, 1996). It was shown that these proteins could perform the same functions as their native proteins.

As many plant tissues are generally recognized as safe, the probability of health risks due to contamination with potential human pathogens and toxins is greatly reduced. These considerations have led to an increased interest in expressing important gene sequences from diverse origins (e.g. human, bacteria, virus) in plants (Tacket *et al.*, 2000: [www.prodigene.com](http://www.prodigene.com)). Since the seeds of many plant species are edible, therapeutic proteins produced in seeds, in some instances, may not require further processing and purification, offering the option of oral delivery. In fact, tubers from transgenic potato that produced 38-nm Norwalk virus-like particles were implicated in acquired oral immunity in humans (Tacket *et al.*, 2000). Although seeds can remain dormant for years, it was only recently shown that recombinant seed proteins were easily stored at room temperature without degradation (Fiedler and Conrad, 1995; Stoger *et al.*, 2000). Therefore, plants may be effective systems for the production of valuable proteins (Sijmons *et al.*, 1990; Ganz *et al.*, 1996).

#### *Advantages of using plant systems as "pharming platforms"*

Traditionally, transformation by *Agrobacterium tumefaciens* and the biolistic particle gun were used to generate transient or stable expression in both monocots and dicots. However, it is simple to introduce recombinant expression constructs into the genomes of plant systems using *A. tumefaciens*. Gene inserts into host plant genomes result in stable transgenic tissues. The short regeneration time for harvesting therapeutic proteins makes plants more attractive, contrasted with, a transgenic cow for example, that takes at least one year to produce transgenic milk.

Plants, unlike bacteria, have the ability to process recombinant proteins in a similar manner as is observed in humans. For a review of protein processing in plants. (e.g. pre and pro peptide cleavages, N- linked glycosylations, disulphide bond formations, beta-hydroxylations, gamma-carboxylations) please see Cramer *et al.*, 1996.

The high yielding biomass of plants offers an additional incentive to investigate plants as potential protein production factories. Increased recombinant protein production in plants may be possible by using plant signal sequences, increased pyrimidine (G + C) content of the gene/cDNA (Horvath *et al.*, 2000) and endoplasmic reticulum retention sequences (Schouten *et al.*, 1996).

Transgenic plant start-up costs are relatively low compared to industrial fermentation, cell culture, or blood fractionation facilities. Rapid scale-up is possible with plant systems, but it takes time to buy industrial space and equipment for cell culture systems. Additional costs are incurred by mammalian and bacterial systems for inputs such as cell culture media and bioreactors. Plant systems' inputs such as sunlight, carbon dioxide, oxygen, water and soil are mostly free or obtained at minimal costs. Storage of recombinant proteins after industrial production is required in many expression systems, while recombinant proteins in leaves and seeds of plants were shown to be stable for long periods (Fiedler and Conrad, 1995; Fiedler *et al.*, 1997; Stoger *et al.*, 2000). Targeting of recombinant protein to edible plant seeds, roots and leaves may result in reduced purification costs. There is minimal risk associated with plant systems acting as vectors for delivery of infective human pathogens during therapeutic administration. After considering these advantages of plant systems for recombinant protein expression, the choice was made to assess tobacco seed capacity for rthGM-CSF accumulation.

## **Seed storage proteins: glutelin gene regulation and secretion of proteins**

### *Seed protein homologies*

Homologies between seed storage proteins may help elucidate shared regulatory and processing pathways. To determine the degree of homology between seed storage proteins from various plant species, legume-like polypeptide homologies were found between pea (dicot) and oat 12S globulin (monocot); among wheat, rye, and corn (Robert *et al.*, 1985a); between rice glutelin (monocot) and soybean glycinin (dicot) (Robert *et al.*, 1985b; Higuchi and Fukazawa *et al.*, 1987) and between glutelin subfamilies (Gt1, Gt2, and Gt3) and the 11S legume storage protein (Okita *et al.*, 1989).

Glutelins comprise the major portion of rice reserve proteins and are encoded by three subfamilies of genes: Gt1, Gt2 and Gt3. There are approximately five to eight gene copies within each subfamily. The degree of homology varies in each subfamily: Gt1 and Gt2 proteins are 87% homologous but both showed significant divergence when compared with the Gt3 protein sequence (Okita *et al.*, 1989). Glutelin genes appear to share a common regulatory pathway. Hence, use of the Gt1 or Gt3 promoter along with a glutelin signal sequence in recombinant constructs may result in identical seed regulation.

### *Glutelin gene regulation*

Glutelin genes are subjected to transcriptional regulation as other genes are in living organisms. The location of *cis*-acting sequences responsible for spatial and temporal endosperm expression of glutelin was investigated in transformed tobacco endosperm cells (Leisy *et al.*, 1990; Zhao *et al.*, 1994; Takaiwa *et al.*, 1996). The beta-glucuronidase (GUS) and chloramphenicol acetyl transferase (CAT) fusion-gene activities allowed determination of the location and strength of endosperm expressions reported for the glutelin promoters. Glutelin promoters fused to these reporter genes directed the synthesis

of proteins only in tobacco seed endosperm cells, further confirming the specificity of the glutelin promoters in both monocots and dicots (Leisy *et al.*, 1990; Zhao *et al.*, 1994; Takaiwa *et al.*, 1996). The promoter region contained between -441 and -237 bp conferred endosperm specific expression (Takaiwa *et al.*, 1991). Transgenic tobacco plants expressing a rice Gt3 promoter fused to a CAT reporter coding sequence showed that glutelin accumulation begins at eight days after flowering (DAF) and peaks at sixteen DAF (Leisy *et al.*, 1990). The rice seed storage prolamins, also follow a similar spatial and temporal expression pattern as rice glutelins (Shyur *et al.*, 1992).

In another experiment, different mutated regions of the Gt3 promoter fused to reporter genes were analysed in tobacco (Zhao *et al.*, 1994). Two glutelin promoter domains were needed for optimized endosperm expression: domain I (-346 to -263 bp) and domain II (-945 to -726 bp). Deletion of domain II led to a delay of approximately eight days when compared to the peak at which glutelin accumulates. Recently, additional data were reported of the existence of duplicate conserved AACA (from -73 to -61 bp and -212 to -200) and GCN4 (-165 to -158 bp and -96 to -92 bp) motifs responsible for seed-specific endosperm expressions (Takaiwa *et al.*, 1996). Plants have similar gene regulatory mechanisms as animals (reviewed by Kuhlemeier, 1992) and thus recombinant constructs utilizing glutelin promoters may be recognized and processed correctly in plants. Therefore, expression constructs utilizing wildtype glutelin promoters are expected to maintain the same glutelin spatial and temporal expression in tobacco seeds. Since the glutelin promoter and signal sequences were used in the recombinant expression construct, its gene product processing will next be detailed as a possible pathway of secretion.

### *Glutelin processing and deposition*

Glutelins are initially synthesised as large preproteins at the rough endoplasmic reticulum (Yamagata and Tanaka *et al.*, 1986). Co-translational-dependent transport into the ER results in cleavage of the N-terminal 24 amino acid residue signal peptide. Additional processing in the Golgi apparatus results in matured glutelins being deposited into type-II protein bodies (PB-II). Type-I protein bodies were found to contain the next major rice storage proteins - prolamines. It is expected that the rthGM-CSF produced in transgenic tobacco plants will undergo similar processing as the glutelins. For a complete review on compartmentalization of proteins in rice and wheat, please see Okita, 1996.

### *Advantages of using rice glutelin promoters for recombinant protein production*

Protein bodies provide a low-hydration environment necessary for the long-term stability of these proteins that are synthesized during embryogenesis. Only at specific times during the development of seeds are the seed storage promoters active and this ensures a smooth and orderly series of events to proceed (Leisy *et al.*, 1990; Shyur *et al.*, 1992; Zhao *et al.*, 1994). A developmental time frame of 40-60 days results in a final protein content in mature rice grain of 6% (Ganz *et al.*, 1996). In rice endosperm cells, glutelins comprise 80% of the seed reserve proteins (Okita *et al.*, 1989). During storage of seeds, enzymatic activity in the dormant embryo is minimal, but upon germination, seed storage proteins are hydrolysed to supply amino acids for the growing plant. If mature grain is stored in a dry cool place, years may elapse between the time of seed maturation and germination. The same seed protein stability may be afforded to recombinant proteins of pharmaceutical interest.

It may be possible to express recombinant human proteins in seeds and study their expressions. The protein reserves remain stable as was recently shown for antibody

production in transgenic tobacco leaves and seeds (Fiedler and Conrad, 1995; Fiedler *et al.*, 1997). Long-term stability of recombinant protein was also demonstrated in rice and wheat (Stoger *et al.*, 2000). Production of therapeutic peptides in the “dry” seed environment is expected to reduce refrigeration costs and increase the possibility of edible forms of therapy. Therefore, seeds may be effective systems for the production of high-value pharmaceutical proteins (Tackaberry *et al.*, 1999). Several experiments to produce recombinant peptides in tobacco seeds were conducted using glutelin promoters and these will be discussed below.

### **Some major groups of proteins successfully produced in plant systems**

#### *Plant derived antigens*

Several reports have detailed antigen production in plants and these include: *E. coli* heat-labile enterotoxin  $\beta$ -subunit (Haq *et al.*, 1995), hepatitis B surface antigen (Thanavala *et al.*, 1995), rabies glycoprotein (McGarvey *et al.*, 1995), insulin fused to the cholera toxin  $\beta$ -subunit (Arakawa *et al.*, 1997, 1998a, b), Norwalk virus capsid protein (Mason *et al.*, 1998) and human cytomegalovirus immunodominant glycoprotein B complex (Tackaberry *et al.*, 1999). In several of these reports, the plant produced viral antigens were implicated in the development of oral immunity (Arakana *et al.*, 1998a). The vaccines that were extracted from various plant tissues showed approximate equivalent ability to stimulate the immune system as the commercially available vaccines (Thanavala *et al.*, 1995). For example, oral immunity resulted in production of IgA antibodies against the rabies virus in mice (Modelska *et al.*, 1998). These studies demonstrated the feasibility of oral immunity using viral and bacterial antigens (Modelska *et al.*, 1998). This technology may result in

future development of low-cost edible vaccines in plant systems (Mason *et al.*, 1992, 1996, 1998; Arakana *et al.*, 1998a, b).

#### *Production of recombinant antibodies in plants*

In addition to antigens, antibodies have been produced in many different cellular compartments of plants and animals. In 1989, Hiatt *et al.*, became the first group to produce antibodies in plants. Since then, other recombinant antibodies have been produced. Mucosal antibody against genital herpes was produced in soybean and mammalian cell cultures (Zeitlin *et al.*, 1998) and anti-tumour specific antibodies were recently produced in rice cell cultures (Torres *et al.*, 1999). Rice cell cultures produced therapeutic antibodies against tumours, resulting in a maximum yield of 3.8 µg/g of callus tissue. Glycosylated antibodies from plant source were reported to confer high specificity for their native antigens and hence plant glycosylation patterns did not affect the antibody functions (Tavladoraki *et al.*, 1993; Zeitlin *et al.*, 1998). Further information on antibody production in plants may be obtained from the recent review of Conrad and Fiedler, 1998.

#### **Granulocyte macrophage colony stimulating factor**

Growth factors were originally identified by *in vitro* assays and each growth factor was named due to its effects on different cell types. The growth factor that stimulated the growth of certain types of white blood cells (eosinophils, neutrophils and monocytes) was designated Granulocyte Macrophage Colony Stimulating Factor (GM-CSF). In 1977, GM-CSF was discovered, but it was not until eight years later that the cDNA of human (h) GM-CSF was cloned (Wong *et al.*, 1985a, b). The murine cDNA was isolated one year earlier (Gough *et al.*, 1984). The mature secreted hGM-CSF is a single-chain polypeptide of 127 amino acid residues (Wong *et al.*, 1985a, b; Cantrell *et al.*, 1985; Lee *et al.*, 1985).

Even though hGM-CSF showed approximately 70% amino acid sequence identity with the murine GM-CSF, they can only bind and activate their own receptors (Maliszewski *et al.*, 1988; Leong *et al.*, 1989).

Since the pleiotropic cytokine GM-CSF was discovered, it was found to stimulate the proliferation, maturation and function of a variety of hematopoietic cells (Sieff, 1987; Gasson, 1991; Klingemann *et al.*, 1991). Additionally, GM-CSF has a proliferative effect on non-hematopoietic cell lineages (Gasson, 1991; Nicola, 1989). However, more insights into its actions have accumulated over the years to include the modification of functional activity of eosinophils, neutrophils, monocytes and macrophages (Kleinermann *et al.*, 1988; Sullivan *et al.*, 1989) stimulation of T cell proliferation (Santoli *et al.*, 1988) and induction of expression of interleukin-1 in neutrophils (Lindermann *et al.*, 1988).

When hGM-CSF was extracted from different human tissues, it was found to be glycosylated resulting in an apparent MW of 23 kDa (reviewed by Quesniaux and Jones, 1998). The mature unglycosylated hGM-CSF has a MW of approximately 14 kDa as reported by: Wong *et al.*, 1985a, b; Cantrell *et al.*, 1985; Lee *et al.*, 1985 and reviewed by Quesniaux and Jones, 1998. The bovine GM-CSF gene from a BT2 T-cell line was isolated and found to code for a 143 amino acid polypeptide. The mature protein results from cleavage of a 17 amino acid signal peptide leaving 126 amino acids (Maliszewski *et al.*, 1988). The MW of mature unglycosylated bovine GM-CSF was predicted to be 14.3 kDa while expression of the bovine GM-CSF cDNA in the periplasm of *E. coli* resulted in a protein of 14.5 kDa on SDS-PAGE (Leong *et al.*, 1989). The bovine GM-CSF protein alignment predicted greater than 50% homology with human and murine GM-CSF. Conservation of both the "location of cysteine residues and glycosylation sites." was reported (Leong *et al.*, 1989).

In humans, GM-CSF production is tightly regulated. It is produced in extremely low concentrations by many cells including T cells, B cells, macrophages, endothelial cells and fibroblast cells (Gasson, 1991; Metcalf, 1991a). Although, these cells can produce GM-CSF, increased GM-CSF production in both T and B lymphocytes may result from antigens or interleukin stimuli (Brizzi *et al.*, 1995 and reviewed by Quesniaux and Jones, 1998). For example, increased transcription of hGM-CSF mRNA was found in a human myeloid leukaemic cell line (M-07e) when it was stimulated with interleukin-3 or mast cell growth factor (Brizzi *et al.*, 1995). Thus GM-CSF's production is necessary for the proper functioning of the immune system.

#### *Human GM-CSF receptors*

The receptors for hGM-CSF are composed of a ligand specific 80 kDa  $\alpha$ -subunit and a 120 kDa  $\beta$ -subunit (Gearing *et al.*, 1989; Hayashida *et al.*, 1990; Kitamura *et al.*, 1991). The receptor's  $\beta$ -subunit can also bind IL-3 and IL-5. Both hematopoietic cells and non-hematopoietic cells express receptors for GM-CSF. These three cytokines (GM-CSF, IL-3 and IL-5) have similarities in their structures and have been classified as belonging to the type I cytokine receptor superfamily. The  $\alpha$ -subunit binds GM-CSF with low affinity ( $K_d \sim 10$  nM) while the  $\beta$ -subunit requires the assistance of the  $\alpha$ -subunit to bind GM-CSF (Rajotte *et al.*, 1997). Aspartate 112 of hGM-CSF was identified as a possible amino acid responsible for interacting with arginine 280 of the  $\alpha$ -subunit of the receptor (Hercus *et al.*, 1994). Mutant mice lacking the receptor's  $\beta$ -subunit showed no GM-CSF receptor binding (Robb *et al.*, 1995). In the absence of  $\alpha$ -subunit, GM-CSF cannot bind and activate the  $\beta$ -subunit. As a result, it was suggested that the  $\alpha$ -subunit must recruit the  $\beta$ -subunit and convert the complex into a high affinity one ( $K_d \sim 50$  pM) (Hayashida *et al.*, 1990). Once receptors are activated, the signals must be transmitted to the nucleus.

### *Transduction of GM-CSF signal into the nucleus*

The  $\alpha$  and  $\beta$ -subunits of the GM-CSF receptor require signal transduction proteins to regulate the cell cycle (Polotskaya *et al.*, 1993; Ronco *et al.*, 1995). Mutational analyses showed that the GM-CSF  $\alpha$ -subunit receptor is responsible for cell growth (Polotskaya *et al.*, 1993, 1994) while the intracellular  $\beta$ -subunit activates the Janus kinase-2 signalling pathway (Brizzi *et al.*, 1996; reviewed by D'Andrea and Gonda, 2000). Janus kinase-2 activation during GM-CSF stimulation resulted in transcription of the cell cycle genes c-fos and c-myc (Watanabe *et al.*, 1996). The  $\beta$ -subunit was reported to activate the mitogen-activated protein (MAP) kinase pathway (reviewed by Quesniaux and Jones, 1998).

### **Clinical applications of rthGM-CSF**

A significant amount of data on GM-CSF has accumulated on its molecular mechanisms leading to the observed effects on multiple cells. This growth factor plays a vital role in regulating the production and function of white blood cells (granulocytes and monocytes) and thus is of great importance in fighting infections (Metcalf 1991a, b). A very low count of granulocytes is associated with susceptibility to serious infections. In recent years, it has been found that GM-CSF has many beneficial clinical uses in the areas of chemotherapy, bone marrow transplantation, cancer and AIDS (Metcalf 1991a, b). Thus, it offers potential opportunities for the treatment of certain medical conditions (e.g. congenital neutropenia, infections in individuals diagnosed with low numbers of granulocytes and defects in blood cell formations). As a result of successful clinical trials, GM-CSF has been approved in different countries for clinical use (reviewed by Quesniaux and Jones, 1998).

## **Recombinant hGM-CSF from different expression systems**

Practical methods to produce hGM-CSF in sufficient quantities for clinical use have relied on recombinant DNA methods. For example, rthGM-CSF was produced in COS-1 (Wong *et al.*, 1985b), yeast (Ernst *et al.*, 1987) and Namalwa cells (Okamoto *et al.*, 1990). Transfection of hGM-CSF cDNA into COS-1 monkey kidney cells resulted in production of rthGM-CSF with identical properties as the rthGM-CSF derived from a human T cell line (Wong *et al.*, 1985b). When *E. coli* was used to produce hGM-CSF, only one band of approximately 15 kDa was seen on a Western blot (Burgess *et al.*, 1987). Production of rthGM-CSF in yeast resulted in both O-linked and N-linked glycosylations (Ernst *et al.*, 1992).

### *Recombinant hGM-CSF from yeast systems*

To secrete rthGM-CSF in yeast, a fusion construct was made with the yeast  $\alpha$ -mating factor secretion peptide and the mature hGM-CSF cDNA (Ernst *et al.*, 1987). Both O-linked and N-linked glycosylations were reported for the secreted rthGM-CSF. Westerns blots showed 14.5 kDa, 15.5 kDa, and 18 kDa bands. A smear was observed between 30-100 kDa. Endoglycosidase H treatment (cleaves core N-linked mannose chains) or site directed mutagenesis of both N-linked glycosylation sites resulted in only two observable bands at 14.5 and 15.5 kDa. From these experiments, the 15.5 kDa band was identified as an O-linked glycosylated species. It was concluded that some form of O-linked and N-linked glycosylations were present in the higher MW species. Higher secretion rates (90%) were observed for the glycosylated 50 kDa protein when compared to the unglycosylated 14.5 kDa protein (20%).

### *Recombinant hGM-CSF from baculoviral systems*

Several biologically active proteins were produced in baculoviral-infected cells including hGM-CSF (Schuh and Morrissey, 1999), porcine GM-CSF (Inumaru *et al.*, 1998) and porcine IL-18 (Muneta *et al.*, 2000). Expression of the hGM-CSF gene in *Spodoptera frugiperda* {(Sf9), common name: Fall Armyworm} cell culture using viral infection resulted in secretion of biologically active hGM-CSF. Bands of apparently 14.5 kDa, 15.5 kDa and 16.5 kDa were detected on western blots (Chiou and Wu, 1990). The observation of a single band when cells were treated with tunicamycin, or when rthGM-CSF from Sf9 cell culture was N-link deglycosylated led the authors to conclude that the two higher MW species were N-linked glycosylated proteins. Baculoviral infection of *Bombyx mori* (silkworm) nuclear cells with the hGM-CSF gene also yielded three different molecular masses of rthGM-CSF {15 kDa, 18 kDa and 20 kDa} (Shi *et al.*, 1996). Other reports showed Sf9 cell culture secreted functional rthGM-CSF (Au *et al.*, 1996).

### *Recombinant hGM-CSF from mammalian cell cultures*

COS-1 cells from monkey kidney or Chinese hamster ovary (CHO) cells have been used to produce recombinant bovine (Maliszewski *et al.*, 1988) and human (Luo *et al.*, 1995) GM-CSF polypeptides as well as hGM-CSF-EPO fusion proteins (Coscarella *et al.*, 1997). Transfection of hGM-CSF cDNA into COS-1 monkey kidney cells resulted in production of rthGM-CSF with identical properties as the rthGM-CSF derived from a human T cell line (Wong *et al.*, 1985b). Further, the rthGM-CSF protein produced by monkey kidney COS-1 cells was reported to be biologically active, as assessed by proliferation studies on a hGM-CSF dependent leukemic cell line (Luo *et al.*, 1995). It was suggested that the adeno-associated virus 2 used by this group of researchers to transfect COS-1 cells might prove valuable in therapeutic conditions requiring administration of GM-CSF. In fact, several adenoviral vectors for expression of Lac Z, interleukin 6, or

canine GM-CSF were injected into the bone marrow of dogs to assess the therapeutic effectiveness for administration via viral infection (Foley *et al.*, 1997). Analysis of infected bone marrow cell supernatants revealed that the canine GM-CSF gene was actively transcribed and biologically active resulting in a "localized myeloid expansion." In another related experiment to deliver recombinant protein to the lungs of cystic fibrosis patients, viral infection lead only to short-term therapy (Crystal *et al.*, 1994).

It is obvious from the various expression studies in different systems that the quest for delivery of safe biologically active therapeutics is ongoing. Human therapy, requiring erythroid differentiation, may utilize a fusion protein of human GM-CSF with human Epo since the chimera was shown to be better at stimulating erythroid differentiation than the same combination of both of the individual growth factors (Coscarella *et al.*, 1997). In the near future, the demand for rthGM-CSF is likely to increase. The high cost of this recombinant product, however, is a limiting factor. Therefore, alternative methods of production should be investigated to ensure its safe and economical supply (Ganz *et al.*, 1996).

#### **Effects of glycosylation on recombinant proteins: GM-CSF**

In plants and insects, N-linked glycosylated proteins usually have two sugar residues { $\beta$ -(1,2)-xylose and  $\alpha$ -(1,3)-fucose} that are not found in mammals. In plant systems, sialic acid is not present, however it was found in mammalian glycosylated proteins (reviewed by Chrispeels and Faye, 1996). During therapeutic application of recombinant proteins from plants or insect sources, the above mentioned sugars may stimulate the immune system and eventually, the recombinant proteins may become inactivated. This inactivation may be due to the development of antibodies directed against the plant or insect produced proteins. The ability to harvest biologically active human

proteins from plants in large quantities may require modification of a suitable plant host for achievement of this goal (Cabanes-Macheteau *et al.*, 1999). In an attempt to modify tobacco suspension cells to glycosylate proteins with mammalian sugar residues, an N-linked glycosylation enzyme ( $\beta$  1,4-galactosyltransferase) was introduced (Palacpac *et al.*, 1999). This enzyme modified the host N-linked glycosylation pattern and its substrate was found on plant cell glycosylated proteins.

Human erythropoietin (Epo) showed similar proliferative activity as GM-CSF and IL-3 on a human TF-1 cell line (Kitamura *et al.*, 1989). Deglycosylation of hEpo was reported to increase its *in vitro* activity. Its *in vivo* activity, however, was significantly impaired (Tsuda *et al.*, 1990; Higuchi *et al.*, 1992). GM-CSF differs from Epo in that its deglycosylated form is still active *in vivo*. Therefore, one may conclude that *in vitro*, the carbohydrate moieties are not necessary for the biological activity of some proteins, but are necessary for others. The glycosylation restraints imposed by the *E. coli* expression system was not worth further investigation for proteins requiring glycosylation for their *in vivo* activities. Therefore, an ideal plant system for human cytokine production is needed that will save time and lives.

### **Application of glutelin promoters in prior seed specific expressions**

#### *Glycoprotein B complex in transgenic tobacco seeds*

The first attempt to spatially and temporally produce an oral subunit vaccine as a "farmaceutical" occurred in our Ottawa laboratory. This was achieved in seeds utilizing the well-characterized glutelin-3 promoter and its associated signal sequence (Tackaberry *et al.*, 1999). The human cytomegalovirus immunodominant glycoprotein B complex (HCMV IG-BC) was targeted to the seeds of transgenic tobacco plants but not to leaves or

stems. The HCMV IG-BC was detected by electron microscopy in protein bodies of tobacco endosperm cells (E. Tackaberry; personal communication). Again, the Gt3 promoter demonstrated its seed specific expression status. Further, the glutelin signal sequence was needed for correct targeting to the protein bodies of endosperm cells.

ELISAs specific for the HCMV IG-BC detected a synthesis range of 70-146 ng/mg of total soluble extracted seed protein. These values for the recombinant protein represent a maximum of 0.01% of total soluble proteins that accumulated in transgenic tobacco seeds. Although the recombinant protein accumulation levels were relatively low when compared to the 4% accumulation level observed for the phaseolin protein expressed in transgenic rice seeds under the control of a different glutelin promoter (Zheng *et al.*, 1995), tobacco was used only as a test organism for processing and stability of the HCMV IG-BC. Antigenic specificity of the plant synthesized HCMV IG-BC was demonstrated using a human fibroblast cell line. This suggested that tobacco could process the viral subunit vaccine with high fidelity for possible vaccine applications. It was suggested that other antigens might be targeted to the seeds of plants for synthesis of human and animal vaccines (Tackaberry *et al.*, 1999).

#### *Antecedent glutelin/hGM-CSF fusion peptide in tobacco seeds*

Our laboratory became the first to demonstrate the feasibility of producing a recombinant human fusion cytokine in transgenic tobacco seeds (Ganz *et al.*, 1996). This group recognized the need for an ample supply of rthGM-CSF and so turned to plant systems to alleviate the impending shortage of cheap, safe and biologically active cytokines. Since GM-CSF was the most widely studied cytokine and its cDNA was commercially available, it was a good candidate protein to study its expression in plant systems. The recombinant expression construct that was made would putatively result in

the synthesis of an N-terminal signal sequence composed of the first eight residues of the glutelin-3 protein fused in frame with the mature hGM-CSF sequence (Figure 1). For targeting the synthesis of the protein into seeds, the glutelin-3 promoter was utilized. This approach was highly successful, resulting in seed specific accumulation of the fusion protein.

Correspondingly, for amassing of the recombinant fusion cytokine explicitly in the protein bodies of seeds, the glutelin-3 signal sequence was used. Whether the glutelin-3 signal sequence was cleaved or not and if the fusion cytokine accumulated in the endosperm compartments was not demonstrated. However, the antecedent cytokine did accumulate in transgenic tobacco seeds and was found to be biologically active. On Western blots, two distinct bands were detected that were 19 kDa and 21 kDa. This experiment demonstrated that tobacco seeds have the ability to process a human fusion cytokine such that it retained its biological functions. The stability of the recombinant cytokine may be due to its true human/foreign nature being masked by the antecedent glutelin-3 fusion peptide. Alternately, this antecedent peptide might allow for preferential processing of the fusion cytokine in the plant seed environment.

## **Rationale and objectives**

The production of human cytokines in plants may offer a cheap, safe and biologically active source of human therapeutic proteins. Before this project, no matured human cytokines were expressed in plant seeds. However, in a previous study from our laboratory, rthGM-CSF was produced in tobacco seeds as a fusion peptide. The fusion peptide consisted of the first eight amino acids of the glutelin-3 mature protein as an antecedent peptide. This antecedent peptide was postulated to offer added stability to the hGM-CSF and to impart preferential processing in the tobacco seed environment. Alternately, syllogistic cytokine stability in tobacco seeds may be due to its true human/foreign nature being masked by the antecedent glutelin-3 fusion peptide. Thus, it became crucial for the present Gt1 experiment to proceed to resolve the effects of the antecedent peptide on hGM-CSF processing in transgenic tobacco seeds. The fusion cytokine, although it accumulated in tobacco seeds and was found to be biologically active, may not be approved for human cytokine therapy. The final goal of expressing human cytokines in recombinant expression systems is for human therapeutic application. The objectives of this project were to express the mature hGM-CSF in tobacco seeds and determine its stability in the plant environment. Rice glutelin promoters are among the strongest endosperm promoters known. The Gt3 promoter is considered weaker than the Gt1 promoter and hence the Gt1 rice glutelin promoter was used for targeting the production into tobacco seeds. Further to the objectives of this project was to resolve whether the seed environment was capable of processing the non-syllogistic rthGM-CSF peptide such that it retains its full biological activity.

## Chapter 2

### MATERIALS AND METHODS

#### **Materials**

##### **Glutelin-1 promoter**

The glutelin-1 promoter was obtained from Dr. Tom Okita (Washington State University, Pullman, WA). I had subcloned this promoter into the rthGM-CSF expression construct at the start of this thesis (Figure 1). However, the sequences that encoded the amino acids (mature hGM-CSF coding sequence and the glutelin-1 signal sequence) had not been sequenced to confirm the integrity of the Gt1 expression construct. Therefore, its sequencing became part of this thesis project.

##### **Glutelin-1 signal sequence**

The coding region of the glutelin-1 signal sequence (SS) was amplified from *Japonica* rice genomic DNA using the polymerase chain reaction. This PCR experiment was conducted by Dr. Ravinder Sardana and Ms Connie Sauder (University of Ottawa, Ottawa, Ontario, Canada). The PCR amplified product was given to Dr. Eileen Tackaberry (Bureau of Biologicals and Radiopharmaceuticals, Health Canada, Ottawa, Ontario, Canada) who subcloned the SS into the multiple cloning site of a second pGEM4Z plasmid.

##### **Mature hGM-CSF coding sequence**

The mature hGM-CSF coding sequence was obtained from R&D Systems (Minneapolis, MN, USA) in the BBG 12 plasmid. An ATG sequence had been fused to the mature hGM-CSF coding sequence such that the native human signal sequence was not included in the BBG 12 plasmid.

### **Nopaline synthase sequence for polyadenylation**

The nopaline synthase sequence for termination of transcription (NOS-TER) was obtained from a 35S/hGM-CSF/NOS-TER vector. This plasmid was previously used to transform tobacco in Dr. Illimar Altosaar's laboratory.

### **Tobacco leaves**

Young tobacco leaves from non-transformed tobacco plants (*Nicotiana tabacum* cv. Xanthi) were available in the laboratory. Sterilized leaf sections were transformed with the rthGM-CSF expression construct driven by the glutelin-I promoter. This construct is referred to as the GtI construct (Figure 1, Results section).

### **Tobacco transformation vector**

The National Research Council of Canada, Saskatoon, Saskatchewan, provided the plant transformation vector pRD400 which contained the neomycin phosphotransferase II gene for resistance to the antibiotic kanamycin (Datla *et al.*, 1992).

### **Chemicals and enzymes**

Solutions were prepared with chemicals from the following suppliers: Sigma, Fisher Scientific, Promega and Gibco BRL. Restriction and modifying enzymes were obtained from Boehringer Mannheim, Gibco BRL, Promega, and New England Biolabs. Radionucleotides (deoxycytidine triphosphate: <sup>32</sup>P-dCTP and deoxynucleoside triphosphates: <sup>35</sup>S-dNTP) were from Amersham. Antibiotics (ampicillin, carbenicillin, and kanamycin) were from Gibco BRL. Customized oligonucleotide primers for PCR amplification and sequencing were synthesized by Dr. G. Alvarado (Synthaid Biotechnologies Inc., Nepean, Ontario).

## **Methods**

Generally, experimental techniques were slightly modified from the previously published standard procedures (Sambrook *et al.*, 1989).

### **Sequencing of the Gt1 construct**

The instructions contained in the T7 sequencing kit (Procedure C: Standard Annealing of Primer to Double-Stranded Template, Pharmacia Biotech) were used to sequence approximately 2 µg of the Gt1 plasmid DNA as described below. Four different primers (10 pmol each) were used in both forward and reverse reactions. The four primers are as follows (1) 5'- GCA CGA TGA TTT CTC ATT G - 3', (2) 5' - ATG GCA CCC GCC CGG TCA - 3', (3) 5' - TGG GAG CAT GTG AAT GCC AT - 3' (4) 5' - AGC AGG AAG TCC TTC AGG - 3'. Only primer (1) was from outside the coding region of the Gt1 construct and, it bound the construct at approximately -60 bp from the start of the glutelin signal sequence. The Gt1 construct contained in the pGEM4Z plasmid was denatured and then ethanol precipitated as recommended by Pharmacia Biotech. Primers were added to the resuspended DNA along with the supplied primer-annealing buffer. After heating the reaction at 65°C for 5 min, the microfuge tubes were immediately transferred to a 37°C water bath for 10 min. The reaction tube was removed from the water bath and placed at room temperature for a further 5 min. Additional components were added to the labelling reaction: dCTP, dGTP, dTTP, NaCl, T7 DNA polymerase (3.2 units) and <sup>35</sup>SdATP. After a brief centrifugation step, the reaction was incubated at room temperature for five minutes. Aliquots of this reaction were transferred to four sequencing mix tubes. Each of the four tubes contained four different dNTPs along with one of the four ddNTP but in a lower concentration. After incubating for five minutes at 37°C, the reaction was terminated by the addition of the supplied stop solution {0.3% bromophenol blue, 0.3% xylene cyanol FF, 10 mM EDTA (pH 7.5), 97.5% deionised formamide}.

## **Preparation and electrophoretic conditions for sequence analysis on polyacrylamide gels**

Slight modifications were made to the previously published, standard procedure (Sambrook *et al.*, 1989). The glass plates and spacers for the polyacrylamide gel assembly were soaked with warm water and thoroughly scrubbed with a sponge. Soap was applied to the sponge and the above mentioned gel assembly parts were scrubbed a second time. If residual silicone was present on the glass plates, a KOH/methanol solution (5 g KOH/100 mL methanol) was used to remove it. After washing with warm tap water, distilled water was used in a final wash. The gel assembly parts were wiped with ethanol and allowed to dry. The short glass plate for the inside gel assembly was placed in a chemical hood and siliconized with Sigmacote. The gel apparatus was assembled with two flexible 0.4 mm plastic spacers positioned on the outer side edges but between the two glass plates. The top and bottom of the gel assembly had no spacers. To prevent leakage, the bottom edge of the gel assembly was sealed with 50 mL of a 6% acrylamide solution {5.7% acrylamide, 0.3% bis-acrylamide, 46% urea, 45 mM Tris base, 45 mM boric acid, 2 mM EDTA (pH 8.0)} containing 250  $\mu$ L of 25% ammonium persulphate and 250  $\mu$ L TEMED. The sequencing gel was also composed of a 6% acrylamide gel solution prepared by the addition of 150  $\mu$ L of 25% ammonium persulphate and 150  $\mu$ L TEMED to 150 mL of a 6% acrylamide solution as above. The gel solution was swirled to mix and immediately poured between the two glass plates. The shark teeth comb was positioned upside-down at the top of the assembly but between the two glass plates. Distilled water was carefully added to the top of the comb and on top of the gel solution. The sequencing gel assembly was wrapped with plastic wrap to prevent further oxidation and the gel was allowed to polymerize overnight. The next day, the plastic wrap was removed along with the comb and the bottom gel seal. The top of the gel was briefly washed with distilled water, and the comb was inserted such that its teeth slightly penetrated the gel front. The

sequencing gel electrophoresis buffer {1X TBE: 45 mM Tris base, 45 mM boric acid, 2 mM EDTA (pH 8.0)} was added to the Bio-Rad electrophoresis apparatus assembly and the gel was pre-run for an hour at 40W. The T7 sequenced reaction samples were denatured at 80°C for two minutes before loading into the wells. The samples were electrophoresed into the gel for 3 h (short run) and 6 h (long run) at 40 W. It was possible to load the samples for the long run, three hours in advance of the short run and this allowed for analysis of up to 300 bp of sequenced data. The determination of the sequences of both long and short runs on the same gel saved time. Once the electrophoresis was completed, the gel apparatus was disassembled and the gel was allowed to bind to a 3MM Whatman filter paper. The gel, supported by the filter paper, was wrapped with plastic wrap and dried for 1 h in a Gel Slab Dryer at 80°C (BIORAD). The paper was then unwrapped and placed into an X-ray cassette along with an X-ray film. The X-ray film was developed using a standard procedure within 2 days (Sambrook *et al.*, 1989). Sometimes a longer period of exposure was required to increase the intensity of the signals. The derived sequencing data was compared to the predicted sequence of the construct.

### **Transformation, selection and growth of *E. coli***

The glutelin-I-expression construct contained in the pRD400 vector (Gt1/pRD400) was used to transform 50 µL of competent cells. Competent *E. coli* DH5  $\alpha$ -cells were obtained from Gibco BRL and transformed using the protocol provided by the supplier. The Gt1 construct was previously assembled and was made available at the start of this project. Briefly, competent cells were thawed on ice for 30 min and 1 µg of the Gt1/pRD400 plasmid DNA was added. The content was gently pipetted five times and left on ice for 30 min. After a heat shock treatment at 37°C for 30 s, one mL of Luria-Bertani (LB) medium (10 g/L bactotryptone, 5 g/L yeast extract, 10 g/L NaCl, pH 7.0) was added.

The mixture was inverted several times and incubated at 37°C for 1 h. Following a 1 min centrifugation step (Micro-Centrifuge Model 235B, Fisher Scientific), the pelleted cells were resuspended in approximately 60 µL of LB medium. Thirty microlitres was spread on LB/1.3% agar plates containing 50 µg/mL kanamycin. After incubating overnight at 37°C, the antibiotic resistant colonies were picked and grown in LB medium containing 50 µg/mL kanamycin. These bacterial cultures were used for DNA isolation as described below.

### **Preparation of competent *A. tumefaciens* cells**

About 10 µL of *A. tumefaciens* cells (LBA 4404) were plated on YEP/agar plates (10 g/L yeast extract, 10 g/L peptone, 5 g/L NaCl, 1.5% agar, pH 7.0) containing no antibiotics. These plates were incubated at 28°C for 2-3 days, within which time colonies appeared. Several single colonies were each grown in 5 mL of YEP medium for 15 hours. Each overnight culture was transferred to 45 mL of fresh YEP medium (contained no agar) contained in 500 mL sterile flask. Cultures grew until an O.D. of approximately 0.5 was achieved. The bacterial cultures were poured into sterile 50-mL centrifuge tubes and centrifuged (3000 rpm, 4°C, 5 min). After discarding the supernatants, the pellets were resuspended in 1 mL of 20 mM CaCl<sub>2</sub>. Aliquots of 100 µL competent cells were dispensed in pre-chilled 1.5 mL sterile microfuge tubes. Competent *A. tumefaciens* cells were stored at -70°C until needed.

### **Transformation of *A. tumefaciens* cells**

The freeze-thaw method from Clontech was used to transform an aliquot of competent *A. tumefaciens* cells (100 µL) with the Gt1/pRD400 construct (1 µg). After adding the plasmid DNA, the cells were gently mixed by pipetting and then incubated on ice for 30 min. After the incubation period, liquid nitrogen was used to freeze the cells.

The cells were thawed at 37°C for 10 min, afterward, 1 mL of YEP medium was added and the tube was inverted several times to mix. The mixture was incubated at 28°C for 4 hours. Following a 30 s centrifugation step, the pelleted cells were resuspended in approximately 60 µL of the remaining YEP medium and 30 µL was spread on a YEP plate (1.5% agar, 50 µg/mL kanamycin). The plates were incubated at 28°C for 2-3 days to allow for antibiotic resistant colonies to appear. These colonies were used to prepare the inoculant.

### **Inoculant preparation**

A transformed *A. tumefaciens* colony carrying the Gt1 construct (pRD400 backbone) was used to inoculate 5 mL YEP medium containing 50 µg/mL kanamycin in a sterile 30 mL test-tube. After shaking (260 rpm) at 28°C for 24 h, a second inoculation using 45 mL YEP/kanamycin medium (50 µg/mL kanamycin) was initiated using the first 5 mL inoculated culture. The second inoculated culture was incubated as above but only for 4 h. A sterile 50-mL centrifuge tube was used to collect the culture (O.D. ~ 0.5) and for centrifugation (3,000 rpm, 4°C, 7 min). The supernatant was discarded and the tubes were inverted to remove most of the antibiotic medium. After resuspending the pellet in 5 mL of YEP medium without antibiotics, the culture was poured into a sterile Petri dish for transformation of the sterilized tobacco leaf sections.

### **Non-transformed tobacco leaf sterilization**

Young tobacco leaves (*Nicotiana tabacum* cv. Xanthi) were freshly harvested and placed into sterile petri dishes contained in a laminar flow hood. Leaves were sterilised for 30 s in 70% ethanol followed by a 10 min soak in 1% Javex (6% sodium hypochlorite). After three washes with sterile dH<sub>2</sub>O, the leaves were placed into a second sterile petri dish

containing sterile dH<sub>2</sub>O. The leaves were cut with a sterile surgical blade into small sections, approximately one cm square. Leaf sections were transformed by *A. tumefaciens* harbouring the Gt1 rthGM-CSF expression construct as described below.

### **Transformation of sterile leaf sections**

A slight modification was made to the previously published procedure for transformation of tobacco (Horsch *et al.*, 1985). Distilled water from sterilized leaf sections was removed by briefly blotting the leaf sections onto sterile filter paper. The leaf sections were transferred into the *A. tumefaciens* suspension. After 5 min, the leaf sections were blotted onto sterile filter paper before being placed onto tobacco co-cultivation medium {1X MS (Murashige and Skoog medium, Gibco BRL), 3% sucrose, 1.0 mg 6-benzyladenine [BA], 0.1 mg/L alpha-naphthalene acetic acid [NAA], pH 5.8, 0.8% agar}. Approximately four leaf sections were placed into each petri dish before sealing with Parafilm. Subsequently, the plates were incubated in the dark at 20°C for two days.

### **Calli induction and regeneration of transgenic plants**

Leaf sections from the above co-cultivation plates were transferred to a regeneration medium (1X MS, 3% sucrose, 1 mg/L BA, 0.1 mg/L NAA, pH 5.8, 0.8% agar, 500 mg/L carbenicillin and 300 mg/L kanamycin). Carbenicillin was used to kill *A. tumefaciens* cells while kanamycin selected for transformed leaf cells. Calli were induced from transformed leaf cells at 20°C and a photoperiod of 16 h light and 8 h dark. Within one month, calli developed and eventually some shoots appeared. When shoots were at least 4 mm long, they were dissected and placed into magenta jars containing a rooting medium (1X MS, 2% sucrose, pH 5.8, 0.8% agar, 500 mg/L carbenicillin, 300 mg/L kanamycin). After one month, plantlets developed roots and were transferred

into pots containing soil. To prevent desiccation, plants were placed in the greenhouse and covered with transparent plastic bags. The greenhouse photoperiod was the same as above but the temperatures were different (light: 26°C. dark: 20°C). After one week, plastic bags were removed and plants were allowed to grow to maturity. Plants were watered when needed and fertilized with a water-soluble fertilizer every two weeks. Transgenic tobacco plants normally matured and flowered within six to nine months.

### **DNA extraction from tobacco leaves**

Young developing leaves were collected from transformed and non-transformed tobacco plants. The leaf surfaces were washed with distilled water and approximately 3 grams of leaves were placed into a sterile mortar that was pre-chilled with liquid nitrogen (LN). Leaves were grounded to a fine powdery white flour. Liquid nitrogen was added several times during the course of grinding to keep the tissues hard and to prevent degradation of the DNA. Each sample was added to a 50-mL centrifuge tube containing 5 mL of buffer S {100 mM Tris.HCl (pH 8.5), 100 mM NaCl, 50 mM EDTA (pH 8.0), 2% SDS, 0.1 mg/mL proteinase K, 10 mM DTT}. The centrifuge tubes were inverted several times to mix their contents and afterwards, they were incubated at 37°C for 1 h. After the incubation period, 5 mL of buffered {100 mM Tris.HCl (pH 8.5), 100 mM NaCl, 50 mM EDTA} saturated phenol was added to each sample and the tubes were placed at 37°C for 30 min. The mixtures were centrifuged (3,000 rpm, 4°C, 10 min) and the top aqueous phases were transferred into new 50-mL centrifuge tubes. The DNA/RNA was precipitated with 0.1 volume of 3 M sodium acetate (pH 5.2) and 2.5 volumes of 99% ethanol. The precipitation step proceeded at 20°C on a rotary shaker with gentle rocking. Eventually, nucleic acids settled at the bottom of the tubes and the mixtures were centrifuged (3,000 rpm, 4°C, 5 min). The supernatants were removed using clean pipettes that were flame drawn. The pellets were allowed to dry and then were resuspended in 0.5 mL of TE buffer

(10 mM Tris.HCl, 1 mM EDTA, pH 8.5). To remove contaminating RNAs, 100 µg of RNase A was added to each sample. After incubating at 37°C for 2 hours, one volume of 1:1 buffered saturated phenol to chloroform was added. The samples were inverted several times to mix and centrifuged (3,000 rpm, 4°C, 10 min). The top (aqueous) phases were carefully removed and the DNA was precipitated as above but the samples were placed at 0°C overnight. The DNA pellets were collected after a centrifugation step (3,000 rpm, 4°C, 10 min). The DNAs were resuspended each in 200 µL of dH<sub>2</sub>O and the concentrations were determined by measuring the optical densities at 260 nm using a spectrophotometer. The integrity of the genomic DNAs was assessed via electrophoresis on a 0.8% agarose gel.

### **Restriction analyses on DNAs**

Plasmid or genomic DNA was digested with the appropriate restriction endonucleases as recommended by the enzyme supplier. For Southern analyses, 40 µg of total genomic DNA was digested overnight at 37°C while plasmid DNAs were digested for 3 hours at 37°C.

### **Agarose gels for DNA electrophoreses**

Digested genomic DNAs were electrophoresis in 0.8% agarose gels (0.8 g agarose/100 mL 1X TAE) that contained ethidium bromide. For RT-PCR experiments, 1% agarose gels were used (1 g agarose/100 mL 1X TAE buffer). Upon solidifying, the gel was submerged in 1X TAE buffer {50X TAE buffer: 242 g Tris base, 57.1 mL glacial acetic acid, 100 mL 0.5 M EDTA (pH 8.0)}. After the DNAs were digested, 6X gel loading buffer (30% glycerol, 0.25% xylene cyanol FF, and 0.25% bromophenol blue) were added to obtain a final 1X gel loading buffer concentration. A brief centrifugation step followed and the samples were loaded into the gel lanes. Phage lambda DNA/*Hind* III

marker and  $\Phi$ X 174 RF DNA/*Hae* III marker were loaded as standards. The DNA was electrophoresed at a constant 5 volts per square cm of gel. After the electrophoresis was completed, the gel was UV illuminated (Fotodyne) and photographed using a Polaroid camera (Polaroid MP.4 LandCamera).

### **Southern analysis of total genomic DNA**

After running the restriction-digested samples on an agarose gel as above, the DNA was transferred onto a nitrocellulose membrane (High Bond N, Amersham) and probed with the radiolabelled *Eco*R I/*Hind* III Gt1 rthGM-CSF expression construct. The gel was photographed as above and placed into a DNA denaturation solution (1.5 M NaCl, 0.5 M NaOH). After 1 h, the gel was rinsed with dH<sub>2</sub>O and placed into a neutralization buffer {1 M Tris.HCl (pH 7.5), 1.5 M NaCl}. During the hour-long neutralization step, the transfer apparatus was set-up as previously described (Sambrook *et al.*, 1989) using 20X SSC (3.0 M NaCl, 300 mM sodium citrate, pH 7.0) as the transfer buffer. Two pieces of 3MM filter paper were wrapped halfway around a glass plate with the ends directly in the transfer buffer. The gel was placed on top of the 3MM filter papers, and a nitrocellulose membrane, presoaked in dH<sub>2</sub>O, was placed on top of the gel. Two additional pieces of 3MM filter paper were presoaked in 2X SSC and placed on top of the membrane. A stack of paper towels was placed onto the transfer set-up. A 500-gram weight was placed on top of the paper towels. The transfer proceeded at 20°C for 18 h. As the paper towels became wet, they were replaced with dry ones. After the transfer was completed, the membrane was soaked in 2X SSC to remove any agarose that was sticking to it. The membrane was allowed to air-dry before the DNA was fixed by baking under vacuum (80°C, 2 h).

### **Extraction and purification of DNA from bacteria**

Miniprep kits were obtained from QIAGEN (Mississauga, Ontario). The QIAprep Spin Miniprep kit protocol was used to extract and purify plasmid DNA from bacteria. In this procedure, transformed *E. coli* cells carrying the plasmid of interest were grown overnight at 37°C in 3.0 mL of LB or YEP medium supplemented with the appropriate antibiotic (ampicillin 100 µg/mL or kanamycin 50 µg/mL). Cells were poured into 1.5 mL microfuge tubes and centrifuged (Model 235B, Fisher Scientific) for 1 min to obtain the bacterial pellet. The supernatants were poured off and the pellets were completely resuspended in 250 µL of Buffer P1 (QIAGEN). Then 250 µL of lysis buffer (P2, QIAGEN) was added to each sample. The alkaline lysis was allowed to proceed for 4 min before being neutralized with 350 µL of buffer N3 (QIAGEN). The lysates became cloudy and were centrifuged (10 min at 4°C). The supernatants were applied to QIAprep spin columns and centrifuged for 30 s. After washing the spin columns, the DNA was eluted with 50 µL of TE (10 mM Tris.HCl, 1 mM EDTA, pH 8.5). Generally, 5 µL of DNA was electrophoresed on 0.8% agarose gels to assess the quality and yield. More accurate quantifications were performed by spectrophotometry using a wavelength of 260 nm. The DNA was used for cloning, making probes, or restriction endonuclease analyses.

### **Generation of the probe and genecloning**

The original Gt1 transformation construct contained in the pGEM4Z plasmid was digested with *EcoR* I and *Hind* III to release the 2.522 kb fragment. The restriction endonuclease *Sca* I was also added to the reaction. The backbone of the pGEM4Z plasmid (2.7 kb) was cut with *Sca* I to enhance separation on a 0.8% agarose gel into three fragments: 2.522 kb transformation construct and the *Sca* I-restricted plasmid backbone consisting of a 1.7 kb and a 1.0 kb fragment. The 2.522 kb fragment was excised under UV and recovered from the agarose gel using a genecloning kit (QIAEX II) from

QIAGEN. Briefly, three volumes (3 mL QX I/g of agarose) of buffer QX I was added to a 1.5 mL microfuge containing the agarose fragment and 30  $\mu$ L of QIAEX II silica particle mixture. The QIAGEN supplied high chaotropic salt contained in the QX I buffer used large anions to disrupt sugar monomers and polymers of agarose. The mixture was incubated for 10 min at 50°C and vortexed every 2 min. The sample was centrifuged for 30 s and the supernatant discarded. The pellet was washed with 500  $\mu$ L of QX I buffer and centrifuged for 30 s. Two additional washes followed using 500  $\mu$ L of the supplied PE buffer. The pellet was air-dried for 15 min and eluted with 50  $\mu$ L of 10 mM Tris.HCl (pH 8.5). The DNA was recovered after a 30 s centrifugation step to pellet the silica particles. The DNA concentration was estimated on an agarose gel (0.8%) and used for the probe labelling reaction.

#### **Radiolabelling of the probe**

The probe for Southern hybridization was labelled using the Ready.To.Go DNA Labelling Beads (dCTP) Kit from Pharmacia. The Manufacturer's Standard Protocol was followed for the labelling reaction. Briefly, boiling for 3 min at 100°C denatured approximately 50 ng of the gene-cleaned *EcoR I/Hind III* digested expression construct. The reaction was placed immediately on ice for 2 min. After a brief pulse centrifugation step, the DNA and 50  $\mu$ Ci of [<sup>32</sup>P]dCTP were added to the supplied reaction tube. The solution was mixed by vortexing and briefly pulse centrifuged before incubating at 37°C for 15 min. The supplied reaction mixture consisted of the following components: reaction buffer, dATP, dGTP, dTTP, Klenow fragment (7-12 units) and random oligodeoxyribonucleotides (mainly 9-mers). Unincorporated nucleotides were removed from the labelled DNA fragments by centrifuging through a MicroSpin S-300 HR column (Pharmacia). The purified DNA was collected as the flow-through from the column.

### **Prehybridization and hybridization for Southern blots**

The baked nitrocellulose membrane was placed in 5X SSC solution for 5 min and transferred to a tray containing 50 mL of prehybridization buffer. The prehybridization and hybridization buffers were composed of the following components: 50% deionised formamide, 5X Denhardt's solution {0.1% ficoll, 0.1% polyvinyl-polypyrrolidone (PVPP) and 0.1% BSA}, 5X SSC, 0.1% SDS, and 100 pg/mL denatured salmon sperm DNA. The membrane was incubated at 42°C in the prehybridization buffer with constant, gentle shaking (50 rpm). After two hours, the prehybridization buffer was replaced with 30 mL of fresh hybridization buffer. The purified radioactive probe was boiled for 3 min and rapidly cooled on ice. After 5 min on ice, the probe was added to the hybridization solution. The hybridization was allowed to proceed at 42°C for 16 hour with gently constant shaking at 50 rpm. The nitrocellulose membrane was carefully removed and placed into a tray containing 400 mL of 2X SSC and 0.1% SDS solution. The first wash was conducted at room temperature (20°C). The second and third washes were at 65°C (15 min each) in 400 mL of a 1X SSC solution containing 0.1% SDS. The final wash was performed at 65°C (15 min) but the composition of the 400 mL solution was 0.4X SSC and 0.1% SDS. The membrane was removed and placed onto a piece of filter paper to air dry for half an hour before being wrapped in plastic wrap. The wrapped membrane was placed in a cassette and an X-ray film was placed on top of the membrane. Two intensifying screens sandwiched the radioactive membrane and the X-ray film. The film was developed using a standard procedure (Sambrook *et al.*, 1989).

### **Collection of seeds**

At maturity, flowers were self-pollinated and tagged (dated). Seeds were harvested at specific days after pollination (DAP). Seeds were collected from the early to the mid-maturation stage (6-15 DAP). After harvesting the seeds, they were kept on ice and

quickly scraped away from the pods. Seeds from different plants were transferred into separate cryovials and kept frozen at  $-70^{\circ}\text{C}$  until needed for mRNA or protein extractions.

### **Extraction of total seed RNA**

A previously published procedure for the collection of RNA was slightly modified for the collection of RNA from seeds (Verwoerd *et al.*, 1989). Seeds were collected between 6-11 DAP to get a wide range of seed mRNAs. Extreme care was taken to avoid contamination: gloves and sterile surgical blades were used for different samples. Seeds, scraped away from pods, fell directly into liquid nitrogen. Seeds were grounded in liquid nitrogen to a fine powdery flour using a sterile mortar and pestle. A pre-weighed 1.5 mL microfuge tube was used to collect approximately 100 mg of ground seeds and, immediately, 500  $\mu\text{L}$  of hot ( $80^{\circ}\text{C}$ ) extraction buffer was added  $\{(1 : 1)$  phenol : 0.1 M LiCl, 100 mM Tris.HCl [pH 8.0], 10 mM EDTA, 1% SDS $\}$ . To homogenize the mixtures, the tubes were vortexed for 30 seconds and 250  $\mu\text{L}$  of 24 : 1 (chloroform : isoamylalcohol) was used to extract the ribonucleic acids. A five-minute centrifugation step followed and the top aqueous phases were transferred to new tubes. One volume of 4 M LiCl was added to each sample. The RNAs were allowed to precipitate overnight at  $4^{\circ}\text{C}$ . After centrifuging for five minutes, the pellets were collected and each re-dissolved in 250  $\mu\text{L}$  of diethyl pyrocarbonate (DEPC)-treated water. After addition of 0.1 volume of 3 M NaOAc (pH 5.2), the RNAs were precipitated using two volumes of 99% ethanol, and placed at  $-70^{\circ}\text{C}$  for two hours. Following a 10 min centrifugation step, the supernatants were removed and the RNA pellets were washed with 70% ice-cold ethanol (containing 30% DEPC water). The pellets were collected after centrifuging for 10 min and were dried at  $37^{\circ}\text{C}$ . The RNA pellets were re-dissolved in 50  $\mu\text{L}$  of DEPC water and stored in 10  $\mu\text{L}$  aliquots at  $-70^{\circ}\text{C}$ . No RNA was extracted from leaves since the hGM-CSF specific ELISA kit detected no rthGM-CSF in transgenic and non-transgenic tobacco leaves.

## **Preparation of RNA gels**

Gels, buffers and samples were prepared in DEPC water. The integrity of the above extracted RNAs was assessed by electrophoresis on a 1.0% agarose gel containing 1.9% formaldehyde. The gel tray, gel comb and electrophoresis chamber were washed with distilled water and presoaked overnight in a DEPC solution containing 1% SDS. The 1X RNA electrophoresis buffer {MOPS [3-(N-morpholino) propanesulfonic acid]/EDTA} was prepared from a 10X stock solution (10X MOPS/EDTA buffer: 0.2 M MOPS, 50 mM sodium acetate, 10 mM EDTA, pH 7.0). To each 5  $\mu$ L of RNA sample were added 25  $\mu$ L of RNA loading buffer (0.75 mL deionized formamide, 0.15 mL of 10X MOPS electrophoresis buffer, 0.24 mL of 37% formaldehyde, 0.1 mL DEPC-treated dH<sub>2</sub>O, 0.1 mL glycerol, 0.08 mL of 10% (w/v) bromophenol blue). The RNA samples containing the RNA loading buffer were placed at 65°C for 15 min. After cooling to 50°C, 1  $\mu$ g of ethidium bromide was added to each sample. Prior to loading the RNA samples, a pipette was used to flush the sample wells with fresh 1X RNA electrophoresis buffer. After a brief centrifugation step, the RNA samples were loaded into the flushed wells and the gel ran at a constant 30 V. After electrophoresis, the gel was visualized and photographed under UV illumination (Fotodyne) as described for DNA agarose gels. Quantification of the RNAs by spectrophotometry (260 nm) was also conducted. Total RNAs were used in RT-PCR experiments.

## **Reverse transcriptase (RT) and polymerase chain reaction (PCR) {RT-PCR} experiments**

### *Reverse transcriptase experiment*

Total RNA was isolated from the developing seeds of transformed and non-transformed tobacco plants as described above. RT-PCR experiments were performed to detect the presence of hGM-CF mRNA transcripts using the Advantage<sup>TM</sup> RT-for-PCR

Kit from CLONTECH. The supplied total RNA from human placenta along with PCR specific glyceraldehyde-3-phosphate-dehydrogenase (G3PDH) primers was used as a positive control for the RT-PCR reactions. The cDNA synthesis reaction contained 0.5 µg of total seed RNA or 1 µg human placenta total RNA or water (negative control), reaction buffer {50 mM Tris.HCl (pH 8.3), 75 mM KCl, 3 mM MgCl<sub>2</sub>}, dNTP mix (0.5 mM each), 20 pM of oligo (dT)<sub>18</sub> primer, 1 unit of RNase inhibitor and 200 units of Moloney-Murine Leukemia Virus reverse transcriptase. After incubating for 1 h at 42°C, the cDNA reaction was stopped by heating at 94°C for 5 min. The reaction was diluted five fold and the cDNAs were amplified as described below.

#### *Polymerase chain reaction experiment*

The PCR reaction was performed to amplify a 384 bp fragment from the hGM-CSF cDNA that was achieved in the RT experiment. The PCR reaction was performed using a PCR buffer {50 mM KCl, 20 mM Tris.HCl (pH 8.4), 2.5 mM MgCl<sub>2</sub>} and the following components were added to the reaction: 0.1 mM of each dNTP, 20 ng of both the forward and reverse primers, 10 µL of the diluted cDNAs, and one unit of Taq DNA polymerase. The forward and reverse primers were: forward primer, 5' - ATG GCA CCC GCC CGG. TCA - 3'; and reverse primer, 5' - CTC CTG GAC TGG CTC CCA - 3'. A 100 µL aliquot of light mineral oil was placed into each reaction tube and the tubes were briefly centrifuged (3,000 rpm, 4 s). Initially, the samples were denatured at 95°C for 3 min. Thirty-two cycles (45 s at 95°C, 45 s at 55°C, 45 s at 72°C) were performed using a Microcycler apparatus (Eppendorf). Finally, an extension step at 72°C for 3 min was performed. The PCR reactions were analysed on a 1% agarose gel.

### **Extraction of total seed protein**

Approximately 500 mg of early to mid-maturation tobacco seeds were placed into a pre-chilled mortar. A pestle was used to grind the seeds to a fine powder. The pulverized seeds from each sample were scooped up using sterilized spatulas and placed into pre-weighed 1.5 mL microfuge tubes. The tubes containing the samples were weighed again and the net weights were determined. Protein extraction buffer {50 mM Tris.HCl (pH 8.0), 50 mM NaCl, 1 mM EDTA, 0.1% (v/v) triton-X 100, 1% (w/v) insoluble polyvinyl-polyrrolidone (PVPP), 1% (w/v) ascorbic acid, 25  $\mu$ L of a 200 mM solution of phenylmethyl-sulfonyl fluoride} was immediately added in a one to one (v/w) ratio. For example, 500  $\mu$ L of protein extraction buffer was added to 500 mg of seeds. The samples were vortexed to ensure complete mixing and centrifuged (13,000 rpm, 4°C, 10 min). The supernatants were transferred to new tubes and the pellets discarded or re-extracted as desired using the same extraction buffer ratio as above. The total protein content of the seed extracts was determined with the Bio-Rad Protein Assay Kit that uses the method of Bradford (Bradford, 1976). Bovine serum albumin was used as a protein standard. The extracted seed proteins were used in rthGM-CSF specific ELISA, biological activity assays and ammonium sulphate precipitations for Western blots.

### **Ammonium sulphate precipitation**

Ammonium sulphate precipitations were conducted on total protein extracts from both transformed and non-transformed tobacco seeds as described previously (Wong *et al.*, 1985a; Harris, 1989). Briefly, a two-step process was conducted to partially purify the plant-produced rthGM-CSF from other seed proteins. The first 0-65% precipitation step involved the addition of 404 mg of ammonium sulphate per mL of total protein extract. The second precipitation step, 65-95%, used 208 mg ammonium sulphate per mL of initial starting total protein extract. After addition of the 404 mg of ammonium sulphate to the

1 mL of total protein extracts, the tubes were tapped gently to mix and placed on ice (4°C). During the 2 h incubation period, the tubes were occasionally inverted. The 0-65% precipitated proteins were pelleted via centrifugation (13,000 rpm, 4°C, 10 min). The supernatants were carefully removed and added to new 1.5 mL microfuge tubes containing the second amount of ammonium sulphate. The tubes were inverted several times to mix and allowed to precipitate overnight at 4°C. After 16 hours, the tubes were centrifuged as above and the supernatants removed. The pellets were resuspended in 70 µL of sterile Tris (pH 8.0) containing 0.1% BSA. Bradford and ELISA analyses were conducted on the reconstituted protein samples. The following quantities of reconstituted protein samples were used in the SDS-PAGE gel: NT 191.7 µg, NT/rthGM-CSF from *E. coli* 274.5 µg, EB/rthGM-CSF from *E. coli* 23.25 µg, plant #14 163.32 µg and plant F1A(#14) 62.2 µg. Please note that the NT and NT/rthGM-CSF reconstituted protein extracts contained more total proteins than the other samples.

#### **Sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE)**

The SDS-PAGE gel was prepared as described previously (Sambrook *et al.*, 1989). The separating gel was a 15% Tris-glycine SDS-PAGE gel while the stacking gel was a 5% Tris-glycine SDS-PAGE gel. The 15% SDS-PAGE gel contained 4.6 mL H<sub>2</sub>O, 10 mL of 30% acrylamide mixture (29.0% acrylamide and 1.0% bis-acrylamide, Bio-Rad), 5.0 mL of 1.5 M Tris (pH 8.8), 0.2 mL of 10% SDS, 0.2 mL of 10% ammonium persulfate and 0.008 mL of TEMED. The 5% stacking gel contained 3.4 mL dH<sub>2</sub>O, 0.83 mL of 30% acrylamide mixture (29.0% acrylamide and 1.0% bis-acrylamide, Bio-Rad), 0.63 mL of 1.0 M Tris (pH 6.8), 0.05 mL of 10% SDS, 0.05 mL of 10% ammonium persulfate and 0.005 mL of TEMED. The electrophoresis buffer contained 25 mM Tris.HCl (pH 8.3), 250 mM glycine and 0.1% SDS. Ammonium sulphate precipitated samples were diluted with distilled water to reduce the ammonium sulphate concentration to less than 1%. Gel

loading buffer were added to the diluted precipitated samples to obtain a final concentration of 1X SDS buffer {20 mM Tris (pH 8.0), 1 mM EDTA, 2.5% SDS, 0.1% (w/v) bromophenol blue and 5% 2- $\beta$ -mercaptoethanol}. After heating at 100°C for 3 min, the samples were centrifuged (13,000 rpm, 4°C, 3 min) and loaded into the wells. The prestained SDS-PAGE MW standards were from Bio-Rad (catalogue number 161-0318). The prestained MW standards as shown in Lane 1, Figure 9 were A:  $\beta$ -galactosidase, 127.0 kDa; B: bovine serum albumin, 84.0 kDa; C: ovalbumin, 49.5 kDa; D: carbonic anhydrase, 35.3 kDa; E: soybean trypsin inhibitor, 28.1 kDa; F: lysozyme, 20.5 kDa; G: a rthGM-CSF band that appeared to migrate at 18 kDa; H: aprotinin, 7.0 kDa. The BIO-RAD Mini-PROTEAN® II Cell electrophoresis apparatus was used for the SDS-PAGE gel electrophoresis. The gel ran at 9.0 V until the dye touched the bottom of the resolving gel.

### **Protein transfer**

The proteins in the SDS-PAGE gel were transferred onto a nitrocellulose membrane using the BIO-RAD TRANS-BLOT® SD, SEMI-DRY TRANSFER CELL apparatus as per the manufacturer's specifications. Briefly, the SDS-PAGE gel was placed onto two pieces of 3MM filter paper that were presoaked in transfer buffer (39 mM glycine, 48 mM Tris.HCl, 0.037% SDS, 20% methanol, pH 8.3). The gel and the chromatographic papers were positioned onto the transfer apparatus such that the gel was on top of the assembly. The nitrocellulose membrane was placed into dH<sub>2</sub>O and then soaked in the transfer buffer for three minutes. The nitrocellulose membrane was placed in direct contact with the gel followed by two pieces of chromatography paper. The chromatographic papers were pre-soaked in transfer buffer before positioning onto the nitrocellulose membrane. Five mL of transfer buffer was pipetted on top of the assembly and the air bubbles were removed by gently rolling a glass pipette over the transfer

assembly. Excess transfer buffer was blotted away from the base of the apparatus. The transfer was allowed to proceed at 10 mA for two hours. At the end of the transfer, the paper sheets were carefully removed and the nitrocellulose membrane was immediately placed into 40 mL of cold (4°C) blocking buffer {20 mM Tris (pH 7.4), 0.9% NaCl, 0.25% gelatine, 0.1% triton X-100 (v/v), 0.02% SDS}. After the transfer, the SDS-PAGE gel was stained overnight with a 0.25% Coomassie brilliant blue R250 solution (45 mL ethanol, 45 mL dH<sub>2</sub>O, and 10 mL glacial acetic acid). The gel was destained using a solution composed of 45% ethanol and 10% glacial acetic acid.

### **Probing and detection of rthGM-CSF**

The primary antibody solution was prepared by diluting the polyclonal hGM-CSF specific antibody (R&D Systems) to approximately 1/2200 times in the blocking buffer. All antibody solutions were stored at 4°C. The nitrocellulose membrane was removed from the overnight blocking buffer and placed at 4°C for 16 h in the primary antibody solution. The membrane was then washed three times with wash buffer {20 mM Tris (pH 7.4), 0.9% NaCl, 50 µL Tween 20 (Sigma)}. Each wash was performed on a rotary shaker with gentle rocking for 20 min at room temperature (20°C). The rabbit anti-goat secondary antibody (alkaline phosphatase conjugated) was diluted 1/3333 times in the blocking buffer and the membrane was incubated for two hours at room temperature. Two brief washes with wash buffer were conducted for 15 min each. Detection of antibody binding was in alkaline phosphatase (AP) buffer {100 mM Tris.HCl (pH 9.5), 100 mM NaCl, 50 mM MgCl<sub>2</sub>}. Each membrane was incubated with 10 mL of AP buffer, 44 µL of nitroblue tetrazolium chloride (NBT) and 33 µL of 5-bromo-4-chloro-3-indolylphosphate p-toluidine salt (BCIP). Once bands developed, the membrane was washed with dH<sub>2</sub>O and allowed to air dry.

## **Human GM-CSF-specific ELISA**

Tobacco seeds from early to mid-maturation stage (9-15 DAP) were used for extraction of total soluble proteins as described above. The extracts were centrifuged (13,000 rpm, 4°C, 10 min) and the supernatants were transferred to new tubes. The rthGM-CSF concentrations were determined using a Quantikine™ kit with instructions from R&D systems. The Quantikine™ is a hGM-CSF enzyme linked immunosorbent assay kit (ELISA). Total seed extracted proteins were determined by the Bradford method (Bradford, 1976). Briefly, the ELISA was performed using a microtitre plate coated with a monoclonal antibody for hGM-CSF. Plant protein extracts as well as *E. coli* produced rthGM-CSF were allowed to bind to the antibodies. The wells were washed three times with the supplied wash buffer to remove unbound proteins. The second antibody which is linked to horseradish peroxidase was allowed to bind to rthGM-CSF at a different site resulting in a sandwich arrangement. The wells were washed a second time and the colour reaction was initiated. After development of the colour reaction, the reaction was stopped and the samples were transferred to a new microtitre plate. The microtitre plate was read in a spectrophotometer at 450 nm. The data from the *E. coli* derived rthGM-CSF was used to generate a standard curve as recommended by R&D Systems. The concentration of rthGM-CSF was determined for each plant extract. The seeds of all transgenic plants produced rthGM-CSF as assessed by ELISA. Further, analyses were also conducted on total leaf protein extracts from some transgenic tobacco plants and no rthGM-CSF was detected.

## **Biological assays**

The human bone marrow, erythroleukemia, TF-1 cell line (Kitamura *et al.*, 1989) was obtained from American Type Culture Collection (ATCC). The cells were grown as suspension cultures in RPMI 1640/glucose (pH 7.0) medium, supplemented with 1 ng/mL

*E. coli* derived rthGM-CSF (R&D Systems) and fetal bovine serum (10%). The cells were washed twice with 1X PBS and resuspended at a concentration of  $2.0 \times 10^5$  cells/mL in RPMI 1640/glucose (pH 7.0). A 0.5 mL volume of cells ( $1 \times 10^5$ ) was transferred into each well of a 24-well tissue culture plate. An aliquot of 0.5 mL RPMI 1640/glucose/20% fetal bovine serum (pH 7.0), containing one of the following samples were added to each well: 1 ng/mL rthGM-CSF (*E. coli* derived), transgenic tobacco seed extract containing 1 ng/mL of plant derived rthGM-CSF, seed extract from a NT plant, seed protein extraction buffer. Final concentrations are RPMI + 10% fetal bovine serum +/- 1ng/mL of rthGM-CSF or buffer or NT seed protein extract. Cells were grown at 37°C and monitored every 48 hours. All experiments were done in duplicate under sterile conditions. For the experiment where the cell growth was measured at 96 hours, additional aliquots (seed derived protein, as well as *E. coli* derived rthGM-CSF) were added to the growing cells at 48 hours.

Live cells were counted using haemocytometry/trypan blue exclusion. A 50  $\mu$ L aliquot of cells was removed from the assay and added to 50  $\mu$ L of a 0.05% trypan blue solution. After mixing, the solution was then pipetted into the wells of a hemocytometer chamber having a depth of 0.1 mm. The hemocytometer was placed under a microscope with a 40X magnification. Only clear, live cells were counted from the four large corner squares. The concentration of cells was then determined taking into account the volume occupied by the cells and the dilution factor. The averages of two independent experiments were used in Microsoft Excel to generate the biological assay graphs. The standard deviations were also calculated using Microsoft Excel.

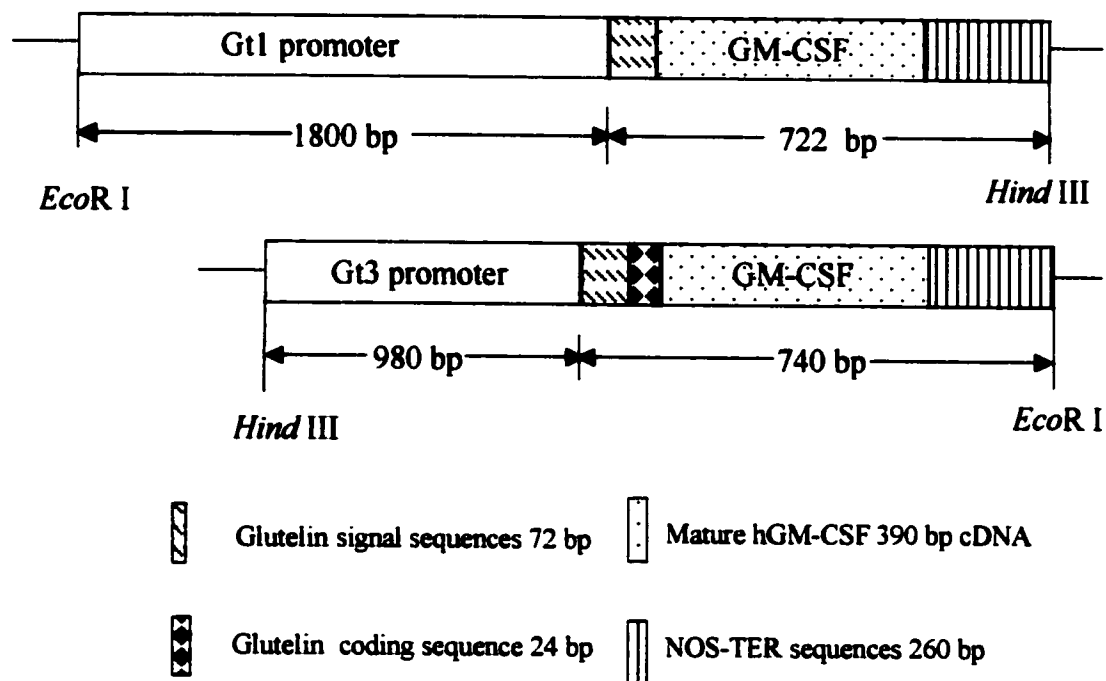
## RESULTS

### **Sequence and restriction analyses on the Gt1 rthGM-CSF expression construct**

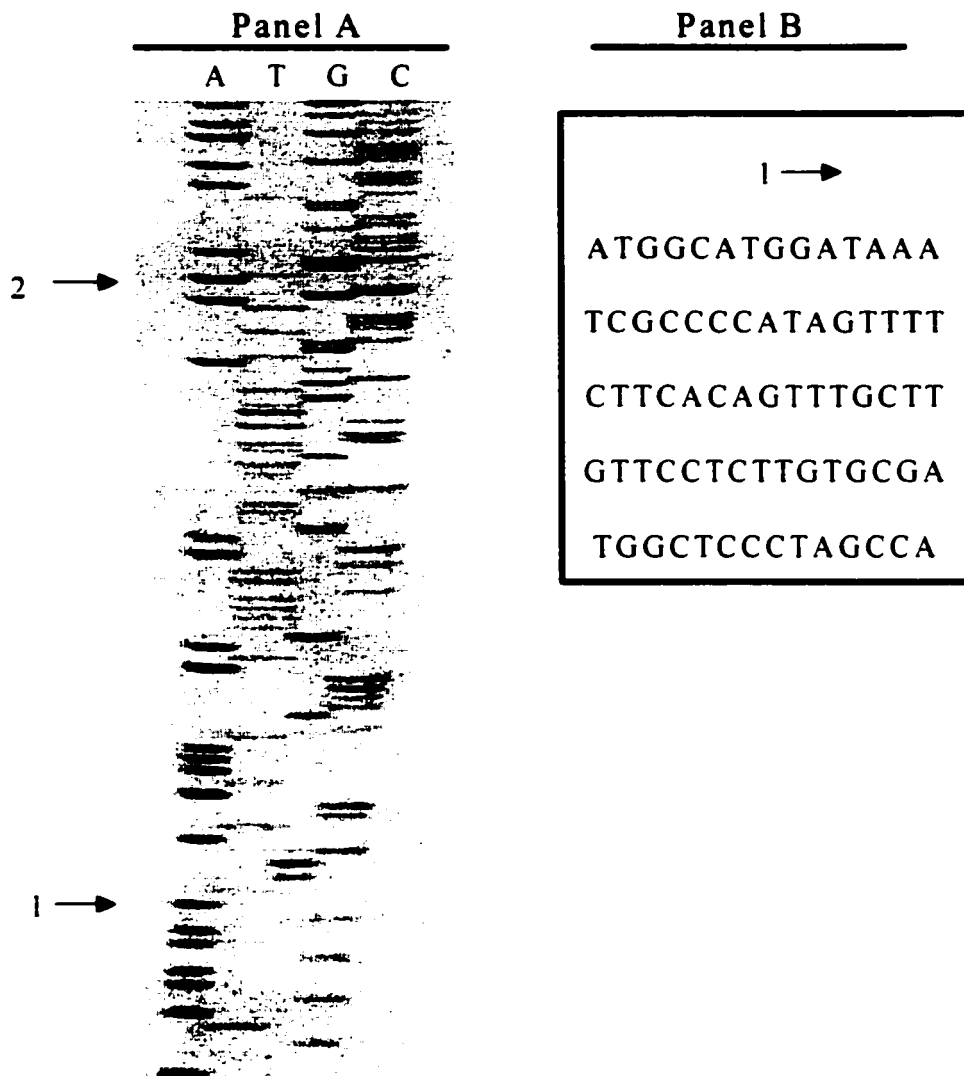
The DNA fragment encoding the mature hGM-CSF was placed under the control of the rice glutelin-1 promoter (Gt1). The rthGM-CSF obtained from two Gt1 transgenic plants was used for comparison in biological assays with the glutelin/hGM-CSF fusion protein produced by a Gt3 plant (Figure 1). As stated before, transformed Gt3 plants were available in the laboratory of Dr. Illimar Altsaar at the time this thesis commenced. Data obtained for the Gt1 construct by the dideoxy sequencing method using the T7 Sequencing Kit confirmed that the complete glutelin-1 signal sequence was fused in-frame with the coding sequence of the mature hGM-CSF sequence (Figures 2-3).

Restriction analyses were used to deduce the orientation of the glutelin-1 promoter and the NOS-TER sequences in the rthGM-CSF expression construct as present in the pGEM4Z plasmid (Figure 1). When the Gt1 expression construct contained in the pGEM4Z plasmid was digested with *EcoR* I, *Hind* III and *Sca* I, a 2.5 kb was observed. This 2.5 kb fragment was expected for the Gt1 construct. When the above mentioned plasmid was digested with *EcoR* I and *Bgl* II (*Bgl* II cuts at the 3' end of the promoter and the 5' end of the glutelin signal sequence), an expected band of 1.8 kb was observed that represented the Gt1 promoter. Further analysis using *EcoR* I, *Hind* III and *Bgl* II resulted in expected fragments of 1.8 kb and 0.722 kb fragments. Restriction analyses using *EcoR* I and *Hind* III conducted on the Gt1/pRD400 plasmid also showed a 2.5 kb band representative of the Gt1 rthGM-CSF expression construct (data not shown). Further.

restriction analyses using *EcoRI* and *Hind* III conducted on DNA extracted from *A. tumefaciens* cells that grew on YEP/kanamycin medium showed that cells harboured the 2.5 kb Gt1 rthGM-CSF expression construct (data not shown).



**Figure 1.** Recombinant hGM-CSF expression constructs. The mature hGM-CSF coding sequence (390 bp) was placed under the control of two rice glutelin promoters Gt1 and Gt3 as shown. Note that the Gt3 construct, in addition to the 72 bp signal sequence, contained the N-terminal 24 nucleotide glutelin coding sequence (encoding an eight amino acid residue sylogistic peptide) fused in-frame with the mature hGM-CSF coding sequence. This additional glutelin sequence is absent from the Gt1 construct. These constructs were each subcloned into separate pRD400 plant transformation vectors. The NOS-TER sequences from *A. tumefaciens* (260 bp) were included for termination (of transcription and polyadenylation). The diagrams are not to scale and are further discussed in the text.



**Figure 2.** Sequencing confirmation of construct integrity. Autoradiograph detection of the junction of the glutelin-1 signal sequence and the mature hGM-CSF coding sequence as discussed in the “Methods” sections on sequencing. Panel A: arrow 1 points to the translation start site of the glutelin-1 signal sequence while arrow 2 points to the start of the mature hGM-CSF coding sequence. The letters at the top of the autoradiogram represent the dideoxynucleotides where the sequences were terminated. Panel B: from top to bottom, the determined nucleotide sequences, reading from arrow 1 to 2 in the autoradiogram.

**ATGGCATGGATAAATCGCCCCATAGTTTTCTTCACAGTTTGCTTGTTT**  
**CTCTTGTCGATGGCTCCCTAGCCATGGCACCCGCCGGTCACCCAGC**  
**CCAGCACGCAGCCCTGGGAGCATGTGAATGCCATCCAGGAGGCCCGG**  
**CGTCTCCTGAACCTGAGTAGAGACACTGCTGCTGAGATGAATGAAACA**  
**GTAGAAGTGATATCAGAAATGTTTGACCTCCAGGAGCCGACTTGCCTA**  
**CAGACCCGCCTGGAGCTGTACAAGCAGGGCCTGCGGGGCAGCCTCACC**  
**AAGCTCAAGGGCCCCTTGACCATGATGGCCAGCCACTACAAGCAGCAC**  
**TGCCCTCCAACCCCGGAACTTCCTGTGCAACCCAGATTATCACCTTT**  
**GAAAGTTTCAAAGAGAACCTGAAGGACTTCCTGCTTGTTCATCCCCTTT**  
**GACTGCTGGGAGCCAGTCCAGGAGTGATAA**

**Figure 3.** The coding sequences of the Gt1 expression construct. The start of the sequence shows the first 72 nt for the glutelin-1 signal sequence (brown and green indicate codons) and is immediately followed by the mature hGM-CSF coding sequence as was present in the BBG 12 plasmid from R&D Systems (red). The hGM-CSF sequence is preceded by an ATG codon and ends with duplicate stop codons (black). For RT-PCR experiments, only the red region was amplified (384 bp). The length of the entire construct is 462 bp and was determined using the T7 sequencing kit as described in the “Methods” sections on sequencing.

## **Transgenic tobacco plants, integration of hGM-CSF DNA and its transcription**

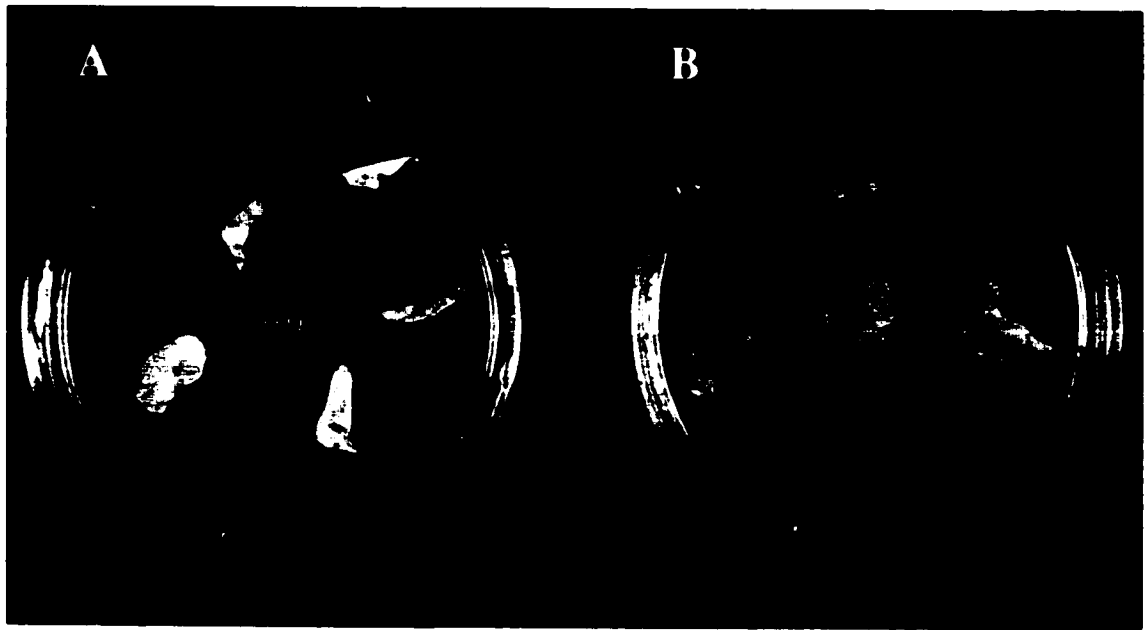
### *Selection for acquired antibiotic resistance in calli and plantlets*

Putatively transformed *A. tumefaciens* cells containing the chimeric pRD400/Gt1 expression construct were used to transform tobacco leaf sections. Selection for antibiotic resistance leaf cells was conducted on a regeneration medium containing two antibiotics: carbenicillin and kanamycin. Carbenicillin is known to kill *A. tumefaciens* cells. Kanamycin was expected to kill non-transformed tobacco cells. Putatively transformed tobacco cells developed into clumps of dedifferentiated tissues known as calli. In Figure 4. Petri dish A shows non-transformed (NT) leaf sections that necrotized while Plate B shows vigorously growing calli tissues and plantlets. One hundred percent of leaf sections on the antibiotic selection medium developed one to several pieces of green calli tissue. Plantlets were further selected for antibiotic resistance when they were placed into a rooting medium containing the same two antibiotics. Roots developed from all plantlets and thereafter, whole plants were transferred to soil. In total, 32 independent Gt1 transgenic plants were regenerated and chosen for further studies.

### *Southern analyses*

The molecular basis for the acquired antibiotic resistance in transgenic Gt1 plants was investigated by performing Southern blots. Genomic DNA from several Gt1 plants was isolated from leaves, purified, digested with *EcoR* I and *Hind* III, fractionated by agarose gel electrophoresis, transferred onto nylon membrane and probed with labelled *EcoR* I-*Hind* III fragments (Figure 1) containing the hGM-CSF mature coding sequence. An expected fragment of 2.522 kb was seen for three Gt1 plants (Figure 5: lanes 4-6). No

bands were observed for the NT plant (Figure 5: lane 3). In addition to the three plants shown in Figure 5, eleven other transformed and regenerated tobacco plants were analysed via Southern blotting and all plants tested positive for the 2.522 kb hGM-CSF expression construct (data not shown).



**Figure 4.** Continuous selection for transformed tobacco leaf tissues. The photograph shows leaf sections on a regeneration medium after one month of selection on 300 mg/L kanamycin and 500 mg/L carbenicillin as described in the “Methods” section on calli induction. Leaf sections in Petri dish A were not transformed and therefore necrotized. Plate B shows leaf sections that were transformed with *A. tumefaciens* harbouring the rthGM-CSF expression construct: proliferation of resistant cells into calli and morphogenesis into plantlets was observed.



**Figure 5.** Genomic organization of three independent transgenic events. Genomic DNA (40 ug) from each tobacco plant was digested with restriction enzymes and loaded onto a 0.8% agarose gel. After electrophoresis, the DNA was transferred to a nitrocellulose membrane and probed with a gel purified and radiolabelled *EcoR* I-*Hind* III Gt1 construct as shown in Figure 1 and described in the "Methods" sections on Southern. Lanes 1 and 2: the *EcoR* I-*Hind* III insert released from the plasmid containing the Gt1 construct; lanes 3-6: DNA cleaved with *EcoR* I and *Hind* III from a non-transformed plant and three Gt1 transgenic plants {14, F1A(14) and 38} respectively. Detection of independent genomic insertion events, lanes 7-10 and lanes 11-14: DNA cleaved with *EcoR* I alone and *Hind* III alone respectively. For lanes 7-10 and lanes 11-14, the same plant DNA order is maintained as for lanes 3-6: DNA from a non-transformed plant followed by three transgenic plants {14, F1A(14) and 38} respectively. Arrows indicate the sizes of the detected bands in the autoradiogram.

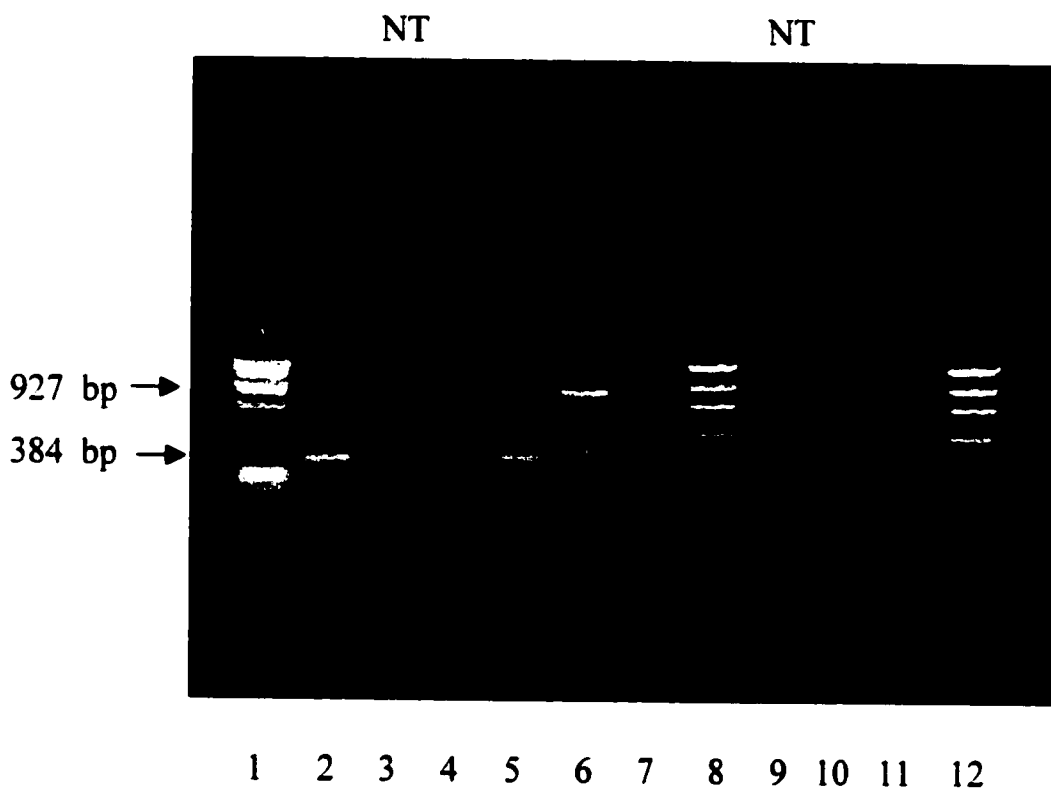
### *Correlation of rthGM-CSF expression at the genomic level*

Further analyses on the Gt1 plants that expressed high {plant #14 and its progeny F1A(14)} and low levels (#38) of rthGM-CSF were conducted to determine the number of stably integrated expression constructs within their genomes. Plants 14 and F1A(14) were easy to study due to the high level of rthGM-CSF that could be derived from their seeds. For comparison with the high expressers, a low Gt1 rthGM-CSF expresser plant 38 was used. Purified genomic DNA was digested with either *EcoR* I or *Hind* III, fractionated by agarose gel electrophoresis, transferred onto a nitrocellulose membrane and probed with labelled *EcoR* I-*Hind* III fragments (Figure 1) containing the mature hGM-CSF coding sequence. At least three copies of the Gt1 expression construct were present in the plants producing high levels of rthGM-CSF {Figure 5: plant 14 and its progeny F1A(14)}. However, at least one copy of the Gt1 expression construct was present in the low expresser Gt1 plant 38 (Figure 5). This data served to clarify the inheritance pattern for the progeny plant F1A(14) of three copies of the Gt1 expression construct.

### *Analyses on transgene transcription*

To check if the transgene was actively transcribed, total RNA was extracted from the developing seeds {9-11 days after pollination (DAP)} of Gt1 plants 14 and its progeny F1A(14). The presence of the hGM-CSF mRNA was detected by a combination of reverse transcriptase (RT) and polymerase chain reaction (PCR) experiments using hGM-CSF sequence-specific primers. The PCR amplified fragment of 384 bp was observed for plants that produced the highest levels of rthGM-CSF. The data was obtained for plants 14 and F1A(14) and is presented in Figure 6: lanes 4-5. Without the RT experiment, no fragment was PCR amplified from total RNA and hence the RNA was not contaminated with

genomic DNA (Figure 6: lanes 10-11). The negative controls (non-transformed tobacco seed RNA and water) did not show any amplified product further confirming that the amplified fragments were not artifacts. The glyceraldehydes-3-phosphate-dehydrogenase (G3PDH) mRNA was present in the total human placenta mRNA that was supplied as a positive control with the RT-PCR kit. When the G3PDH specific primers were used in RT-PCR experiments, a 927 bp band was observed (Figure 6: lane 6).

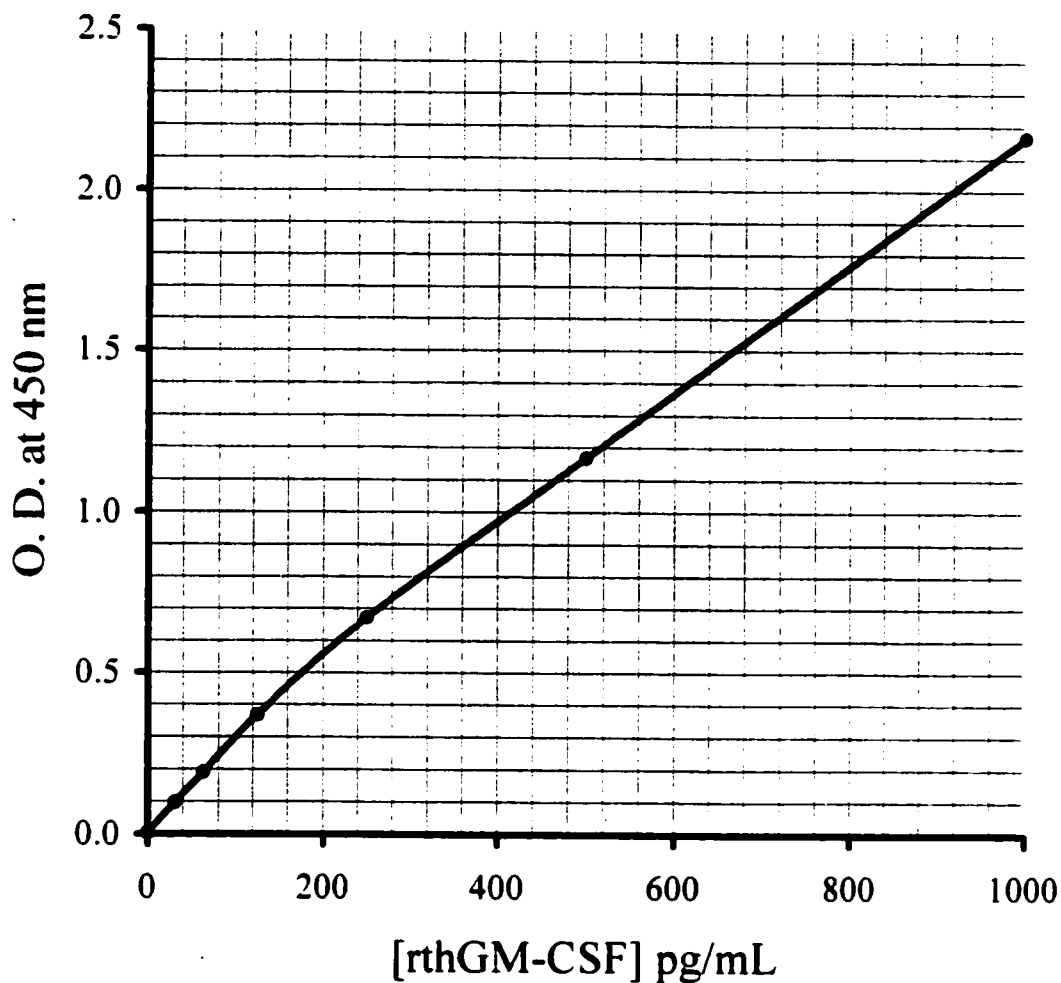


**Figure 6.** RT-PCR detection of the *rthGM-CSF* mRNA transcripts in seeds of transgenic tobacco plants. The cDNAs generated using the RT-PCR kit were PCR amplified and subjected to gel electrophoresis as described in the “Methods” sections on RT-PCR. Lanes 1, 8 and 12: *Hae* III markers; lane 2: mature hGM-CSF coding sequence (384 bp) amplified from a pGEM4Z plasmid containing the Gt1 expression construct; lane 3: total RNA from a non-transformed plant; lanes 4-5: total RNA from transformed plants 14 and F1A(14); lane 6: G3PDH, a positive control (927 bp) from the RT-PCR kit; lane 7: no template; lanes 9-11: total RNA from a non-transformed plant and transformed plants 14 and F1A(14) respectively without the reverse transcription step.

## **Human GM-CSF-specific ELISA and Western blot analyses on seed derived rthGM-CSF**

### *ELISA detection of rthGM-CSF from the seeds of transgenic tobacco plants*

To test if the seeds from the transgenic plants produced rthGM-CSF, protein extracts from transgenic and NT plants were assayed using the Quantikine™. hGM-CSF immunoassay kit. The assay tested developing seeds that were collected at different days after pollination (DAP). A standard curve was generated using *E. coli* derived rthGM-CSF (e.g. Figure 7). The optical density values obtained at 450 nm, for unknown samples, were used to estimate the rthGM-CSF concentrations from the standard curve. The results of the ELISA data on representative plants are shown in Table 1. The transgenic seed extracts contained rthGM-CSF to a level of 736 ng/mL of total protein extract (Table 1: plant 14). This value corresponds to 251 ng/g (rthGM-CSF/total extracted seed proteins) or 0.025% of total soluble proteins. However, plant F1A(14) produced rthGM-CSF to a level of 612 ng/mL of extract or 287 ng/g of extracted seed proteins. This value is 0.029% of total soluble seed proteins. Contrast these results for these two transgenic tobacco plants with the NT tobacco plant that give a reading of 0.0 ng/mL of rthGM-CSF. Thirty-two Gt1 plants were analysed via ELISA for the presence of hGM-CSF in the seed extracts. All transgenic plants produced rthGM-CSF in seeds that ranged from 8 ng/mL to 736 ng/mL of seed extract (data not shown). The seed extracts from an NT plant or the leaf extracts from transformed plants showed no immunoreactive material.



**Figure 7.** A standard curve for the estimation of rthGM-CSF in the seed extracts of transgenic tobacco plants. The rthGM-CSF specific ELISA kit from R&D Systems was used to generate the standard curve. The data was generated using *E. coli* derived rthGM-CSF from R&D Systems. Only the linear portion of the graph was used to derive the unknown rthGM-CSF concentrations in transgenic tobacco seed extracts. Additional information can be found in the "Methods" section on ELISA.

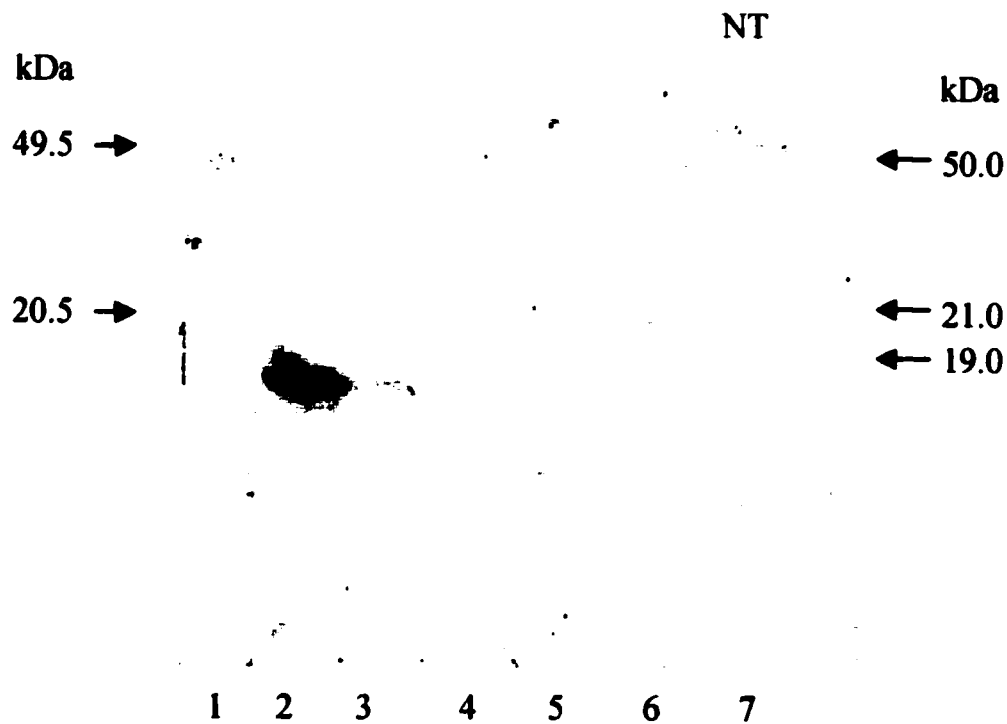
**Table 1.** Human GM-CSF immunoassay on rthGM-CSF extracted from tobacco seeds.

Transgenic event	[GM-CSF] ng/mL of protein extract <sup>a</sup>
NT	0
14	736
38	13.5
F1A(14)	612

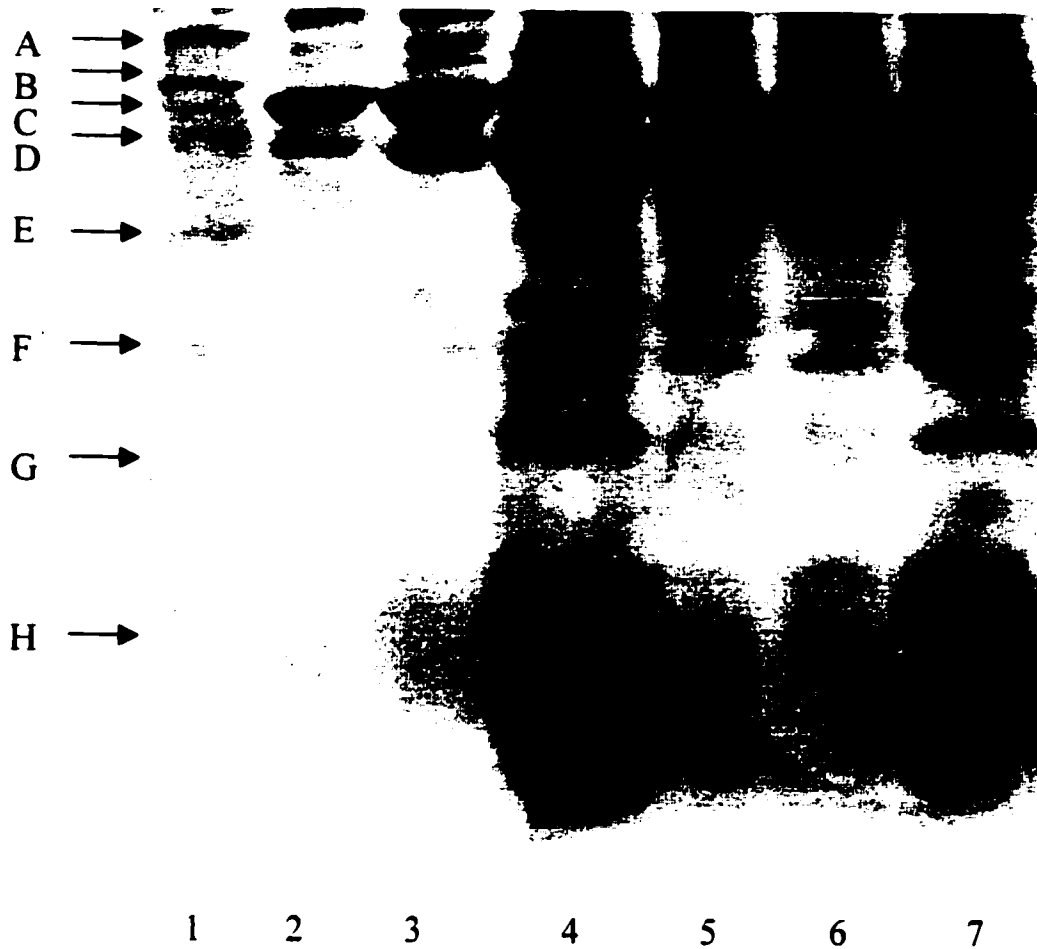
<sup>a</sup>The quantity of rthGM-CSF contained in the primary seed extracts was estimated using the Quantikine™ kit from R&D Systems. Diluted seed extracts were added to pre-coated microtitre wells containing hGM-CSF specific monoclonal antibodies. The unbound substances were washed away with the wash buffer. The amount of rthGM-CSF was quantified using a horseradish peroxidase conjugated secondary antibody and a chromogen solution. A standard curve was generated using *E. coli* produced rthGM-CSF and the unknown concentrations from plant extracts were determined as described in the “Methods” section on ELISA. NT, non-transformed tobacco plant as control.

### *Western blot analysis*

Further characterization of the expressed recombinant protein utilized Western blotting experiments on soluble protein extracts from seeds of transgenic plants 14 and F1A(14). Distinct bands ranging in size from 19-21 kDa from the seed extracts of two transgenic plants were observed (Figure 8: lanes 5-6). No bands of this size range were detected from the seed extracts of a non-transformed plant. In addition to the distinct 19-21 kDa bands, higher molecular weight protein(s) of about 50 kDa from seed extracts of transgenic tobacco plants also reacted with the hGM-CSF antibodies. The result is shown for the two plants in Figure 8: lanes 5-6. A higher MW band (approximately 55 kDa) appeared in all four lanes that contained plant proteins (Figure 9, Lanes 4-7). The *E. coli* derived rthGM-CSF appeared as a distinct band at 18 kDa. After the proteins were transferred to the nitrocellulose membrane, the SDS-PAGE gel was stained with Coomassie blue to determine the efficiency of protein transfer (Figure 9). The prestained standards are shown in lane 1 (Figures 8 and 9).



**Figure 8.** Immunoblot detection of the partially purified, ammonium sulphate precipitated rthGM-CSF from tobacco seeds. After electrophoresis in an SDS-PAGE gel, the proteins were transferred to a nitrocellulose membrane and probed with hrtGM-CSF specific antibodies. Detection utilized alkaline phosphatase linked secondary antibodies as described in the “Methods” section on probing and detection of rthGM-CSF. Prestained SDS-PAGE standards (additional bands are shown in Figure 9) lane 1: ovalbumin, 49.5 kDa and lysozyme, 20.5 kDa. Lanes 2-3: 80 ng and 30 ng of *E. coli* derived rthGM-CSF (14 kDa); lane 4: *E. coli* derived rthGM-CSF added to seed extract from a non-transformed plant; lanes 5-6: protein extracts from seeds of transgenic plants 14 and F1A(14); lane 7: seed extract from a non-transformed plant. Arrows on the right point to bands in the transformed plant lanes 14 and F1A(14).



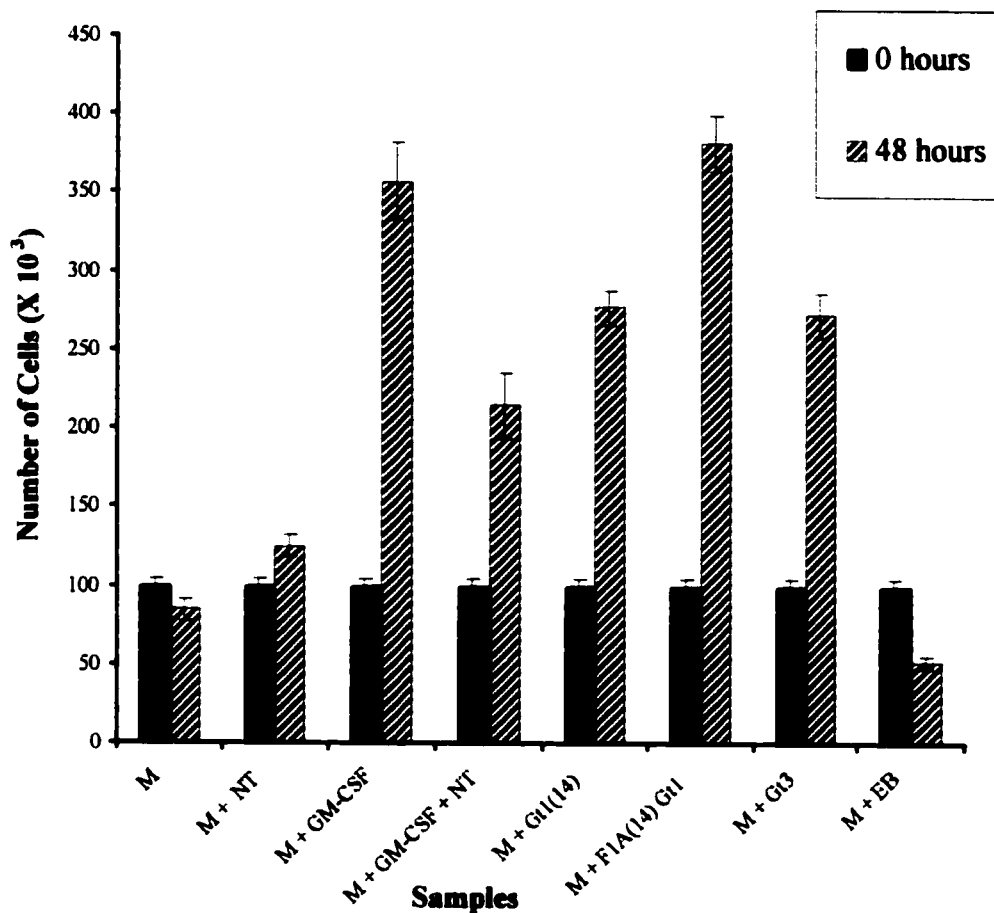
**Figure 9.** Loading of protein samples on SDS gel. After transferring to the nitrocellulose membrane in Figure 8, the SDS-PAGE gel was stained with Coomassie blue as described in the "Methods" section on protein transfer. After destaining, the gel was photographed. Lane 1: prestained MW standards from Bio-Rad. A:  $\beta$ -galactosidase, 127.0 kDa; B: bovine serum albumin, 84.0 kDa; C: ovalbumin, 49.5 kDa; D: carbonic anhydrase, 35.3 kDa; E: soybean trypsin inhibitor, 28.1 kDa; F: lysozyme, 20.5 kDa; G: a rthGM-CSF band that appeared to migrate at 18 kDa; H: aprotinin, 7.0 kDa. Lanes 2-7 represent the same samples as shown in Figure 8. Note that the NT samples (Lanes 4 and 7) had more protein in order to detect any non-specific binding of antibodies in Figure 8.

### **Biological assessment of the seed-derived rthGM-CSF proteins**

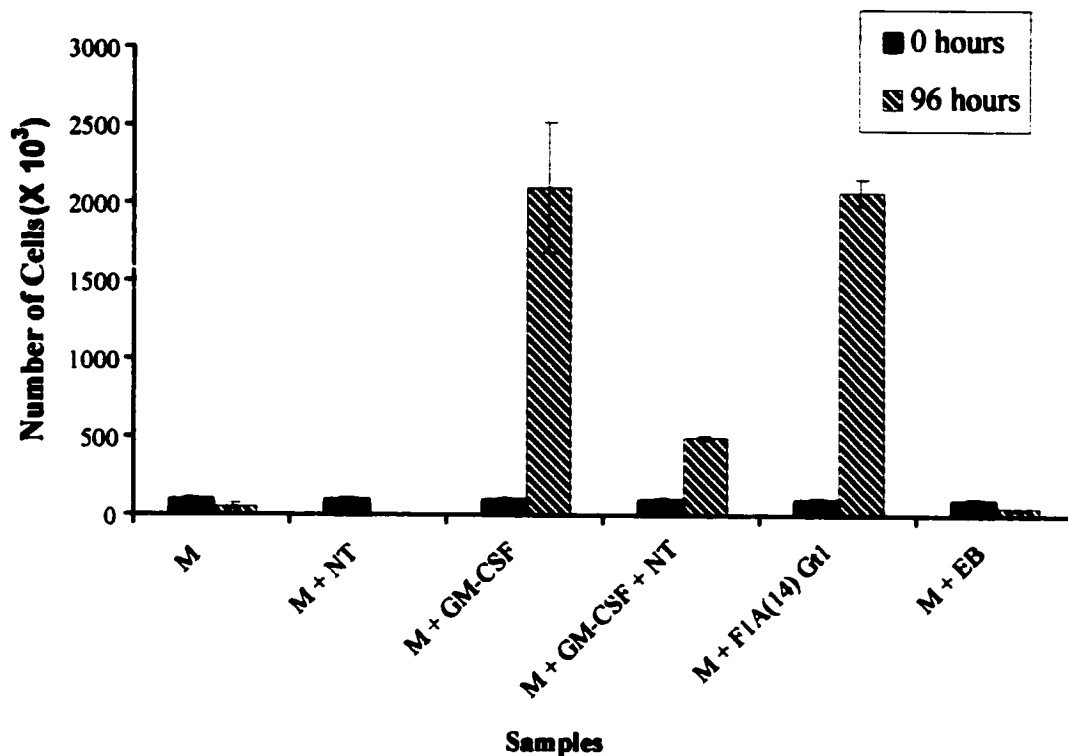
Further analyses were pursued to determine if the rthGM-CSF produced in transgenic seeds were biologically active. Both Gt1 and Gt3 seed-derived extracts were analysed. The hGM-CSF dependent TF-1 cell line was used (Kitamura *et al.*, 1989) to determine cell viability after incubation with or without commercially available rthGM-CSF or aliquots of tobacco seed extracts. The results of these experiments are presented in Figure 10. The bioassay medium alone (without rthGM-CSF), NT seed extracts and the protein extraction buffer (EB) did not support TF-1 cell proliferation. The seed-derived extracts from the Gt1 and Gt3 plants as well as the *E. coli* derived rthGM-CSF when added individually to the medium supported TF-1 cell proliferation.

The extracts from NT tobacco seeds decreased the proliferation of TF-1 cells (Figures: 10-11). For example, with the addition of both *E. coli* derived rthGM-CSF and NT seed extract to the same medium, a 40% decrease in the number of TF-1 cells was observed (Figure 10). This decrease in rate of proliferation was obtained by comparing the number of viable cells found at 48 hours when *E. coli* derived rthGM-CSF alone was added to the medium.

To gain information on the activity of rthGM-CSF expressed in the first generation of a progeny plant F1A(14), biological assays were continued until 96 hours. The results of these experiments are presented in Figure 11. At 96 hours, the seed extracts from the second regenerant F1A(14) plant stimulated the growth of the initial number of starting TF-1 cells, by 20-fold. This increase is equivalent to proliferation by *E. coli* derived rthGM-CSF. When both *E. coli* derived rthGM-CSF and NT extracts were added together to the medium with TF-1 cells, the increase in the number of viable cells achieved was only 5-fold.



**Figure 10.** Biological activity of seed derived rthGM-CSF from Gt1 and Gt3 plants after 48 hours. Human TF-1 cells were grown as suspension cultures. At the start of the experiment,  $1 \times 10^5$  cells were pipetted in duplicate into the wells of a sterile culture plate. The cells were incubated at 37 °C in the presence or absence of aliquots of seed extracts from non-transformed or transformed plants or *E. coli* derived rthGM-CSF. All wells containing rthGM-CSF had a final concentration of 1 ng/mL of rthGM-CSF. Cell viability measurements were determined using haemocytometry/trypan blue exclusion. M: medium; NT: non-transformed seed extract; Gt1(14), F1A(14) Gt1 and Gt3: seed extracts; EB: protein extraction buffer. The experiment is further discussed in the “Methods” section on biological assays.



**Figure 11.** The rthGM-CSF expression construct is stably inherited in the second generation. Seed extract from a progeny plant was analysed for its proliferative effect on TF-1 cells. The data represent the proliferation of cells as measured at 96 hours. TF-1 cells were grown as suspension cultures and  $1 \times 10^5$  cells were pipetted in duplicate into the wells of a culture plate. The cells were incubated at  $37^\circ\text{C}$  in the presence or absence of *E. coli* derived rthGM-CSF or aliquots of seed extracts from either a non-transformed plant or a transformed plant F1A(14). All wells containing rthGM-CSF had a final concentration of 1 ng/mL of rthGM-CSF. Cell number and viability were determined using haemocytometry/trypan blue exclusion. M: medium; NT: non-transformed seed extract; F1A(14) Gt1: seed extract; EB: protein extraction buffer. The experiment is further discussed in the “Methods” section on biological assays.

## Chapter 4

### DISCUSSION

#### **Analysis of transgene(s)**

Here is the first reported expression of a matured human cytokine that specifically accumulated in the seeds of transgenic tobacco plants. The previously reported cytokine that was produced in transgenic tobacco seeds by Ganz *et al.*, 1996 consisted of an N-terminal sylogen composed of the first eight residues of the glutelin-3 protein fused in frame with the mature hGM-CSF sequence. This is also the first report of the use of the glutelin-1 promoter and glutelin-1 signal sequence fused together in-frame with a cytokine for expression in a plant system. Alternately, the Gt3 sylogistic cytokine was driven by the glutelin-3 promoter and utilized the glutelin-3 signal peptide for targeting to the endosperm cells. Although, the Gt1 and Gt3 signal sequences encoded identical signal peptides, the cytokine expression constructs utilized two different promoters from the same family and resulted in the putative syntheses of two different proteins. Therefore, these two experiments, though similar in nature, cannot be equated.

The hGM-CSF coding sequence lacked its own signal sequence but included an N-terminal methionine codon derived from the commercially available BBG12 plasmid. Previous reports showed that the N-terminal methionine was not cleaved from the mature rthGM-CSF protein when it was expressed in *E. coli*. However, this did not affect the protein's biological activity (Burgess *et al.*, 1987). Further, the Gt3 rthGM-CSF from tobacco was shown to be biologically active and this protein should also contain an N-terminal methionine (Ganz *et al.*, 1996). Therefore, the ATG codon at the 5' end of the

mature hGM-CSF coding sequence was retained in the new Gt1 construct with no apparent destabilizing effect on the protein structure.

The utilization of *A. tumefaciens* for generating transgenic tobacco plants was simple and required no specialized apparatus. After the co-cultivation period of the leaf sections with *A. tumefaciens* harbouring the Gt1 expression construct, the selection process was initiated. From Figure 4, it appeared that the concentration of the two selection antibiotics, carbenicillin and kanamycin, was adequate for this experiment. One may deduce that the selection procedure was effective in eliminating non-transformed tobacco cells thereby preventing development of non-transformed plantlets. Hence, all of the tobacco cells surviving selection developed into transgenic calli and eventually differentiated into transgenic plantlets. Since one hundred percent of leaf sections that were co-cultivated with *A. tumefaciens* cells developed one to several pieces of calli tissue, *A. tumefaciens* is ideal for transforming tobacco. Other transformation methods including the particle bombardment method result in significantly lower transformation rates.

No phenotypic abnormalities were observed for any of the regenerated plants. This may indicate that the insertion site(s) of the Gt1 expression construct(s) did not disrupt any genes that are responsible for the overall visual characteristics. However, if mutation of an absolutely essential gene occurred in some transformed tobacco cells, most likely, these cells will not survive and hence could not grow into matured plants. Further, self-pollination and germination of seeds indicated that the parental lines were fertile and viable. For example, plant F1A(14) was a progeny of plant 14 and both demonstrated almost identical genotypic and phenotypic characteristics.

The molecular basis for the acquired antibiotic resistance in transgenic Gt1 plants was investigated by performing a Southern blotting experiment. Southern analysis served several purposes. In particular, Southern blots confirm that the putatively transformed tobacco plants acquired the Gt1 expression construct and that the antibiotic selection procedure was effective. Further, the integrity of the Gt1 expression construct was also deduced from the Southern blot and this indicated the expected size of 2.5 kb (Figure 5). No rearrangement of the Gt1 expression construct was observed which suggests that the rthGM-CSF constructs were functional. The stringency of the hybridization conditions was adequate and hence, no bands appeared for the NT plant lane.

Further, the kanamycin gene for antibiotic resistance appeared to also have had no rearrangements that may result in a loss of its function. If the kanamycin gene was active but no Gt1 expression construct was detected, this would indicate a rearrangement or breakage of the T-DNA during *A. tumefaciens* mediated transformation. Growth on kanamycin and detection of the Gt1 expression construct on a Southern blot seem to indicate that the entire DNA sequence contained between the T-DNA borders were transferred into the genomes of transgenic plants. Therefore, *A. tumefaciens* mediated transformation of tobacco leaf sections coupled with selection on the antibiotic media was an effective method of generating transgenic tobacco plants.

The possible molecular mechanism that resulted in higher and lower production of rthGM-CSF in three Gt1 plants was investigated. The number of stably integrated Gt1 expression constructs within the genomes of high expressor plants 14 and F1A(14) was determined to be at least three copies while the low expressor plant 38 had at least one copy (Figure 5: lanes 7-14). This pattern seem to correlate well with the expression levels observed for plant 14 and its progeny F1A(14). The inheritance pattern of transgenes for

the progeny plant F1A(14) suggests that there were no genomic rearrangements. Although only three Gt1 plants were analysed, it may be possible to increase recombinant protein accumulation in the seeds of transgenic tobacco plants by increasing the transgenes from one to three copies.

The RT-PCR experiment allowed for the detection of the mRNA transcripts arising from the rthGM-CSF gene. The stability of the seed mRNA obtained from the Gt1 construct was not known. Therefore, seeds were collected from different days after pollination (9-11 DAP). The presence of the mRNA transcripts in plants 14 and F1A(14) and their absence in the NT plant seem to strengthen our hypothesis that the rthGM-CSF protein was being produced via translation of the transgenes mRNAs (Figure 6: lanes 3-5). To further confirm that the RT-PCR steps were true and the mRNA was not contaminated with DNA, no product was PCR amplified when the mRNA was used as a template for the PCR reaction (Figure 6: lanes 9-11). Total mRNA from human placenta was used in the RT-PCR experiment as a positive control. The observation of a 927 bp band indicated that the RT-PCR experiment was successful at every stage (Figure 6: lane 6). The RT-PCR results indicated that the mRNA supplied was intact and that the RT-PCR steps were authentic. Taken together, the antibiotic selection, Southern and RT-PCR analyses seem to indicate that the putatively transformed plants are real and are not artifacts. However, this experiment will only be successful if the transgene(s) are correctly processed resulting in function recombinant hGM-CSF accumulation.

#### **Analysis of recombinant hGM-CSF**

To determine whether putatively transformed plants produced rthGM-CSF in the leaves or seeds, the Quantikine™, hGM-CSF ELISA kit was used. The sensitivity of the

assay was reported at 2.8 pg of rthGM-CSF by R&D Systems. The analyses required for this project were somewhat limited by the very nature of the experiments being conducted. Collection of seeds required plants to flower and seeds to be collected at specific DAP. In some cases, regenerated plants became infected and died before analyses were conducted. Hence more than 32 transgenic plants were generated, for example, the previously mentioned low expresser plant #38 has a plant identification number that is greater than 32 - the number of transgenic plants analysed by ELISA. The cost of the ELISA kit was approximately \$540.00. Generation of the standard curve using *E. coli* derived rthGM-CSF decreased the number of sample wells that were available for determining the unknown rthGM-CSF concentrations in protein extracts. The hGM-CSF ELISA Kit expires within one month of first use and plants flower at different times. These constraints significantly consumed the limited financial resources that were available to analyse these transgenic plants and hence, only one set of values was obtained for each standard ELISA curve.

The seeds of all transformed tobacco plants produced rthGM-CSF up to 736 ng/mL of protein extract, and this indicated that every transgenic plant had the capacity to accumulate the recombinant cytokine. No rthGM-CSF was detected in the seeds of an NT plant (Table 1). The competent application of recombinant DNA technology, coupled with plant transformation technology, resulted in transgenic tobacco plants not producing rthGM-CSF in the leaves but explicitly in seeds. The specificity of the glutelin-1 promoter was confirmed when the leaves of transgenic tobacco plants were analysed via ELISA and no rthGM-CSF was observed. In a related work, the use of a different rice glutelin promoter resulted in accumulation of phaseolin protein up to a level of 4% of the total soluble seed proteins in transgenic rice (Zheng *et al.*, 1995). Similarly, ferritin expressed

under the control of a larger rice glutelin promoter, accumulated at the level of 0.01-0.3% of the total rice seed protein (Goto *et al.*, 1999). Therefore, the next stage of the project will be conducted in rice to determine if higher yields are possible.

The actual rthGM-CSF levels present in the seeds of transgenic plants are most likely higher because the re-extracted pellets were found to have significant amounts of rthGM-CSF protein (data not shown). In addition, it was found that the levels of rthGM-CSF protein varied for various lots of seeds collected at different DAP and or even from the same plant (data not shown). These results (Southern, independent genomic inserts, RT-PCR, and ELISA data), for the parent plant 14 and progeny plant F1A(14) are not too radically different. Overall, the data indicate the possibility of a predictable pattern of transgene expression in the progeny of a tobacco plant having approximately the same number of inserts.

The Western blot analysis was the most intricate experiment conducted. Initially, the transgenic crude seed protein extracts did not give any positive result when analysed via Western blotting experiments. A previous report showed hGM-CSF precipitated at 80% ammonium sulphate saturation (Wong *et al.*, 1985a). The precipitated fraction obtained between 65% and 95 % ammonium sulphate saturation resulted in higher yields of rthGM-CSF than the fraction obtained at 80% ammonium sulphate saturation (data not shown). The ammonium sulphate precipitation experiments should result in the removal of some plant proteins and does not result in a homogenous form of the plant synthesized rthGM-CSF. Therefore, ammonium sulphate precipitation experiments were conducted to obtain only partially purified rthGM-CSF from the transgenic seed extracts of two transgenic plants {14 and F1A(14)}. The appearance of distinct bands of 19, 21, and 50 kDa from the seed extracts of these two transgenic plants in Western blots further

confirmed the presence of the rthGM-CSF in the seeds of transgenic tobacco plants (Figure 8: lanes 5-6). In case of the seed-produced rthGM-CSF, the 19-21 kDa bands may be glycosylated although it remains to analyze the carbohydrate structure. The high-molecular weight band (~50 kDa) most likely represents another glycosylated form of the rthGM-CSF protein. Although the NT lane in the Western blot was devoid of bands in the 19-50 kDa size ranges, an approximate 55-kDa band did appear. This 55-kDa band also appeared in other lanes containing plant derived proteins and this may be due to non-specific binding of the rthGM-CSF antibodies.

These observations are parallel to reports on the production of rthGM-CSF in yeast and human Namalwa cells. For example, rthGM-CSF produced in transformed yeast cells ranged in size from 14.5 kDa to 50 kDa (Ernst *et al.*, 1987). Similarly, the Namalwa cell-derived rthGM-CSF showed heterogeneity in the molecular mass of the protein ranging from 16-35 kDa as determined by Western blot analysis (Okamoto *et al.*, 1990). In the hGM-CSF protein backbone, there are two potential N-linked glycosylation sites at Asn27 and Asn37 (Cantrell *et al.*, 1985; Lee *et al.*, 1985). In the Namalwa cell studies, it was proposed that the largest size of rthGM-CSF (25-35 kDa) had both sites N-linked glycosylated, the intermediate size had one site N-linked glycosylated and the smallest size molecules (16-18kDa) had neither site glycosylated (Okamoto *et al.*, 1990).

A number of factors such as the cellular environment, protein structure and molecular interactions can affect the state and efficiency of glycosylation. For example, the human and mouse GM-CSF proteins are 60% homologous but were differentially glycosylated during secretion by a yeast host (Ernst *et al.*, 1987). About 50% of the mouse GM-CSF was not glycosylated in yeast. Similarly, phaseolin, a plant storage protein, was

synthesized as a mixture of singly-glycosylated and fully-glycosylated proteins when expressed in yeast (Vitale *et al.*, 1993).

Because the signal peptides of seed storage proteins such as zeins, glutenins and legumins from different plant species such as maize, wheat and field beans were correctly processed in transgenic tobacco plants (Robert *et al.*, 1989; Bagga *et al.*, 1995; Fiedler and Conrad, 1995), it is quite likely that the rice glutelin signal peptide (24 amino acid residues) was also cleaved from the rthGM-CSF proteins. The exact cleavage of the signal peptide should release the mature rthGM-CSF in Gt1 plants, and mature rthGM-CSF protein with the eight glutelin amino acid sylogistic peptide on the N-terminal end in Gt3 plants (Ganz *et al.*, 1996). Ideally, N-terminal peptide sequencing can confirm the hypothesis for cleavage of the signal peptide.

Furthermore, the proper cleavage of the glutelin signal peptide should permit rthGM-CSF proteins to follow the secretory pathway and, through electron microscopy studies, it should be possible to find the exact subcellular locations of both the Gt1 and Gt3 recombinant factors. In a different targeting experiment in our lab, the Gt3 promoter and its associated signal sequence were used to target a human cytomegalovirus immunodominant glycoprotein B complex (HCMV IG-BC) to the seeds of tobacco plants (Tackaberry *et al.*, 1999). The HCMV IG-BC was found to localize in protein bodies of tobacco endosperm cells (E. Tackaberry; personal communication). In addition, the 15-kDa-zein protein from maize, after correct processing in the ER of transgenic tobacco plants, accumulated in novel ER-derived protein bodies in the endosperm cells (Bagga *et al.*, 1995).

Once a recombinant protein is produced, it will be of little use if it is non-functional. It is this very reason why the crude seed protein extracts from transgenic plants were analysed in biological assays. One key point here is the use of both Gt1 and Gt3 seed-derived extracts simultaneously in biological assay experiments. The TF-1 cell line is an hGM-CSF dependent cell line that was previously used in biological assays (Kitamura *et al.*, 1989). The results in Figure 10 indicated proliferation of the TF-1 cells was only possible when rthGM-CSF from *E. coli* or from transgenic (Gt1 and Gt3) seeds were added to the medium. This experiment showed that TF-1 cells required rthGM-CSF for their proliferation and that no TF-1 cell proliferative factor(s) were present in the NT seed extracts.

The Gt3 synthesized glutelin/hGM-CSF sylogistic peptide was used for comparison with the rthGM-CSF produced by Gt1 plants. The design of the "N-terminal bit" was to mask or disguise the foreign protein since storage proteins like glutelins are known to be stable in tobacco endosperm cells (Leisy *et al.*, 1990; Zheng *et al.*, 1993). Further, since glutelin is normally targeted into stable protein bodies, "N-tagged" rthGM-CSF with the sylogistic or antecedent peptide of glutelin was an attempt to increase its stability in an endosperm environment (Ganz *et al.*, 1996). Since the biological assay for both the Gt1 and Gt3 rthGM-CSF was done in the presence of seed proteins and there were approximate equivalent activities for both proteins, it is inconclusive whether the sylogistic approach, of making the foreign protein as more plant-like, is useful. However, both recombinant proteins were biologically active and "N-terminal" tagging did not affect rthGM-CSF biological activity.

In addition to the NT extracts not having any proliferative capacity on TF-1 cells, it appeared to diminished the biological activity of the *E. coli* derived rthGM-CSF on TF-1

cells (Figure 10). The decrease was so drastic that even the Gt1 rthGM-CSF extracts appeared to be better or equivalent to the mixed cocktail of *E. coli* derived rthGM-CSF and NT extract in supporting the growth of TF-1 cells. This shows that compound(s) in the NT tobacco seed extracts may reduce the biological activity of non-glycosylated *E. coli* derived rthGM-CSF. It is possible that some of the rthGM-CSF from *E. coli* was degraded by plant protease(s).

These observations would suggest that biological activity of tobacco seed-derived rthGM-CSF proteins is also similarly affected. However, the transgenic seed extracts (i.e. from Gt1 and Gt3 plants) are somewhat more effective in supporting the proliferation of TF-1 cells than the mixed cocktail of *E. coli* derived rthGM-CSF and NT extracts. This indicates that the seed-derived rthGM-CSF is more active in supporting the TF-1 cells and this may be due to an increased stability of glycosylated forms of seed-derived rthGM-CSF in the presence of other plant proteins.

Additional information on the effects of the NT extract and the rthGM-CSF protein from the progeny plant F1A(14) were obtained from the biological assay that continued on to 96 hours. A 20-fold stimulation of TF-1 cell proliferation was observed when the seed protein extracts from the F1A(14) plant was used in this biological assay (Figure 10). This stimulation of TF-1 cells was equivalent to proliferation of TF-1 cells obtained by *E. coli* derived rthGM-CSF. The effect of the NT seed extract on *E. coli* derived rthGM-CSF was significant to decrease the 20-fold proliferative effect to only 5-fold. This again suggests that the seed extract from an NT plant contained substance(s) that diminished the activity of the *E. coli* derived rthGM-CSF. The rthGM-CSF protein produced by monkey kidney COS-1 cells and the murine rGM-CSF produced in transgenic tobacco calli and plant tissues were reported to be biologically active as assessed by proliferation studies on

GM-CSF dependent leukemic cell lines (Luo *et al.*, 1995; Lee *et al.*, 1997). These reports support the experimental results being reported in this thesis for the rthGM-CSF expressed in Gt1 tobacco seeds.

As discussed earlier, the seed-derived rthGM-CSF appears to be glycosylated. Although plant N-glycosylation patterns might be somewhat different than humans (e.g. presence of xylose, fucose), the glycosylated tobacco rthGM-CSF might be better tolerated in human therapy than the *E. coli* derived rthGM-CSF (Schuh and Morrissey, 1999). Clinical trials showed that the glycosylated yeast-derived rthGM-CSF resulted in decreased adverse effects when compared to *E. coli* derived rthGM-CSF. Therefore, the carbohydrate structure of the seed-derived rthGM-CSF needs to be analysed, although, glycosylation does not seem to be essential for biological activity of hGM-CSF, either *in vivo* or *in vitro* (Kaushansky *et al.*, 1987; Moonen *et al.*, 1987; Burgess *et al.*, 1987). However, previous reports suggested that rthGM-CSF preparations expressed in yeast, bacteria or mammalian cells have variations in their pharmacokinetics, biological activity and immunogenicity (Dorr, 1993; Hovgaard *et al.*, 1993; Robison and Myers, 1993). These variations may lead to increased clinical toxicity of rthGM-CSF and thus may influence the clinical use of human GM-CSF. Further work on seed-derived rthGM-CSF is needed to determine the direction this project may take.

## **Conclusion**

In this research, the expression of rthGM-CSF was studied in the tobacco system that utilized the well-characterized glutelin-1 promoter. The glutelin-1 promoter and its associated signal sequence were used for direct deposition into the stable, low hydrated compartments of tobacco seeds. This resulted in a highly ordered accumulation of rthGM-CSF in the seed of transgenic tobacco plants. Further, the rthGM-CSF was found to be biologically active as assessed on TF-1 cells suggesting that the seed compartment is suitable for rthGM-CSF production. An antecedent peptide may not be necessary for recombinant cytokine accumulation in transgenic tobacco. Currently, other cDNAs are being expressed in several plant systems, as the initial transgenic rthGM-CSF produced in tobacco appeared as though it was correctly processed. These promising results should act as a foundation for further studies in edible plant systems (carrot, grapes, rice).

The experimental results suggest that it is feasible to produce biologically active rthGM-CSF, and probably other colony stimulating factors also, in the seeds of a higher plant. Further fine tuning of the promoter, signal sequences, proper codon choice (Sardana *et al.*, 1996; Cheng *et al.*, 1998) and host plant combinations (e.g. rice and other cereals) should make it possible to produce very high levels of recombinant proteins in the seeds of several plant species. This work has important implications for the use of edible plant seeds as future bioreactors and for the delivery of drugs.

## **Future directions**

Since successful expression of a cytokine in the seeds of tobacco plants was demonstrated when the seed-specific rice glutelin-1 promoter was used, most likely, the expression of the Gt1 construct in rice will similarly yield positive results. It is important that other analyses on the tobacco-produced rthGM-CSF be conducted (administration in laboratory animals, electron microscopy, carbohydrate moiety analyses and N and C terminal sequencing) to determine the full structural and functional nature of the recombinant protein. This study can be extended to produce other therapeutically useful cytokines, antigens and antibodies in plants using different targeting sequences: promoters, signal sequences and ER retention sequences. Studies to increase recombinant protein yield will need to fine-tune the cDNAs with plant-like codons and determine the most efficient copy number of transgenes needed to achieve maximum expression levels in other plant systems.

## REFERENCES

- Arakawa, T., Chong, D.K. and Langridge, W.H. 1998a. Efficacy of a food plant-based oral cholera toxin B subunit vaccine. *Nat. Biotechnol.* **16**: 292-297. Published erratum appears in *Nat. Biotechnol.* **16**: 478.
- Arakawa, T., Chong, D.K., Merritt, J.L. and Langridge, W.H. 1997. Expression of cholera toxin B subunit oligomers in transgenic potato plants. *Transgenic Res.* **6**: 403-413.
- Arakawa, T., Yu, J., Daniel, K.X. C., Hough, J., Engen, P.C. and Langridge, H.R. 1998b. A plant-based cholera toxin B subunit-insulin fusion protein protects against the development of autoimmune diabetes. *Nat. Biotechnol.* **16**: 934-938.
- Au, L.C., Liu, T.J., Shen, H.D., Choo, K.B. and Wang, S.Y. 1996. Secretory production of bioactive recombinant human granulocyte-macrophage colony-stimulating factor by a baculovirus expression system. *J. Biotechnol.* **51**: 107-113.
- Bagga, S., Adams, H., Kemp, J.D. and Sengupta-Gopalan, C. 1995. Accumulation of 15-kilodalton zein in novel protein bodies in transgenic tobacco. *Plant Physiol.* **107**: 13-23.
- Bourgaize, D., Jewell, T. R. and Buiser, R. G. 2000. Genetic engineering: tools and techniques. Chapter 8, p. 138-169. In Roybal, M. (ed.) *Biotechnology: Demystifying the concepts*. Addison Wesley Longman, Inc., San Francisco.
- Bradford, M.M. 1976. Rapid and quantitative method for quantification of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* **72**: 248-252.
- Brizzi, M.F., Aronica, M.G., Rosso, A., Bagnara, G.P., Yarden, Y. and Pegoraro, L. 1996. Granulocyte macrophage colony-stimulating factor stimulates JAK2 signaling pathway and rapidly activates p93<sup>+</sup>, STAT1 p91, and STAT3 p92 in polymorphonuclear leukocytes. *J. Biol. Chem.* **271**: 3562-3567.
- Brizzi, M.F., Rossi, P.R., Rosso, A., Avanzi, G.C. and Pegoraro, L. 1995. Transcriptional and post-transcriptional regulation of granulocyte-macrophage colony-stimulating factor production in human growth factor dependent M-07e cells. *Br. J. Haematol.* **90**: 258-265.
- Burgess, A.W., Begley, C.G., Johnson, G.R., Lopez, A.F., Williamson, D.J., Mermod, J.J., Simpson, R.J., Schmitz, A. and DeLamar, J.F. 1987. Purification and properties of bacterially synthesized human granulocyte-macrophage colony stimulating factor. *Blood* **69**: 43-51.

- Cabanes-Macheteau, M., Fichette-Laine, A.C., Loutelier-Bourhis, C., Lange, C., Vine, N.D., Ma, J.K., Lerouge, P. and Faye, L. 1999. N-Glycosylation of a mouse IgG expressed in transgenic tobacco plants. *Glycobiology* **9**: 365-372.
- Cantrell, M.A., Anderson, D., Cerretti, D.P., Price, V., McKereghan, K., Tushinski, R.J., Mochizuki, D.Y., Larsen, A., Grabstein, K., Gillis, S., *et al.* 1985. Cloning, sequence, and expression of a human granulocyte/macrophage colony-stimulating factor. *Proc. Natl. Acad. Sci. USA* **82**: 6250-6254.
- Cheng, X.Y., Sardana, R., Kaplan, H. and Altosaar, I. 1998. *Agrobacterium*-transformed rice plants expressing synthetic CryIA(b) and CryIA(c) genes are highly toxic to striped stem borer and yellow stem borer. *Proc. Natl. Acad. Sci. USA* **95**(6): 2767-2772.
- Chiou, C.J. and Wu, M.C. 1990. Expression of human granulocyte-macrophage colony-stimulating factor gene in insect cells by a baculovirus vector. *FEBS Lett.* **259**: 249-253.
- Chrispeels, M. J. and Faye, L. 1996. The production of recombinant glycoproteins with defined non-immunogenic glycans. Chapter 2, p. 99-113. In Owen, M. R. L. and Pen, J. (eds.) *Transgenic Plants: A Production System for Industrial and Pharmaceutical Proteins*. Wiley, New York.
- Conrad, U., Fiedler, U. 1998. Compartment-specific accumulation of recombinant immunoglobulins in plant cells: an essential tool for antibody production and immunomodulation of physiological functions and pathogen activity. *Plant Mol. Biol.* **38**: 101-109.
- Coscarella, A., Carloni, C., Liddi, R., Mauro, S., Novelli, S., Mele, A., Masella, B., Valtieri, M. and De Santis, R. 1997. Production of recombinant human GM-CSF-EPO hybrid proteins: in vitro biological characterization. *Eur. J. Haematol.* **59**: 238-246.
- Cramer, C. L., Weissenborn, D.L., Oishi, K.K., Grabau, E.A., Bennett, S., Ponce, E., Grabowski, G.A. and Radin, D.N. 1996. Bioproduction of human enzymes in transgenic tobacco. *Ann. N.Y. Acad. Sci.* **792**: 62-71.
- Crystal, R.G., McElvaney, N.G., Rosenfield, M.A., Chu, C.S., Mastrangeli, A., Hay, J.G., Brody, S.I., Jaffe, H.A., Eissa, N.T. and Danel, C. 1994. Administration of an adenovirus containing the human *CFTR* cDNA to the respiratory tract of individuals with cystic fibrosis. *Nat. Gen.* **8**:42-51.
- Datla, R.S.S., Hammerlindl, J.K., Panchuk, B., Pelcher, L.E., and Keller, W. 1992. Modified binary plant transformation vectors with the wild-type gene encoding NPTII. *Gene* **122**: 383-384.

- D'Andrea, R. J. and Gonda, T. J. 2000. A model for assembly and activation of the GM-CSF, IL-3 and IL-5 receptors: insights from activated mutants of the common beta-subunit. *Exp. Hematol.* **28**:231-43.
- Dorr, R.T. 1993. Clinical properties of yeast-derived versus *Escherichia coli*-derived granulocyte-macrophage colony-stimulating factor. *Clin. Ther.* **15**(1): 19-29.
- Ernst, J.F., Mermod, J.-J., DeLamarter, J.F., Mattaliano, R.J. and Moonen, P. 1987. O-glycosylation and novel processing events during secretion of  $\alpha$ -factor/GM-CSF fusions by *Saccharomyces cerevisiae*. *Bio/Technology* **5**: 831-834.
- Ernst, J.F., Mermod, J.J. and Richman, L.H. 1992. Site-specific O-glycosylation of human granulocyte/macrophage colony-stimulating factor secreted by yeast and animal cells. *Eur. J. Biochem.* **203**: 663-667.
- Fiedler, U. and Conrad, U. 1995. High-level production and long-term storage of engineered antibodies in transgenic tobacco seeds. *Bio/Technology* **13**: 1090-1093.
- Fiedler, U., Phillips, J., Artsaenko, O. and Conrad, U. 1997. Optimization of scFv antibody production in transgenic plants. *Immunotechnology* **3**: 205-216.
- Foley, R., Ellis, R., Walker, I., Wan, Y., Carter, R., Boyle, M., Braciak, T., Addison, C., Graham, F. and Gauldie, J. 1997. Intramarrow cytokine gene transfer by adenoviral vectors in dogs. *Hum. Gene Ther.* **20**: 545-553.
- Gasson, J. 1991. Molecular physiology of granulocyte-macrophage colony-stimulating factor. *Blood* **77**: 1131-1145.
- Ganz, P.R., Dudani, A.K., Tackaberry, E.S., Sardana, R., Sauder, C., Cheng, X. and Altosaar, I. 1995. Plants as factories for producing human blood proteins. *Can. Soc. Transfusion Med. Bull.* **7**: 112-121.
- Ganz, P. R., Dudani, A.K., Tackaberry, E.S., Sardana, R., Sauder, C., Cheng, X. and Altosaar, I. 1996. Expression of human blood proteins in transgenic plants: the cytokine GM-CSF as a model protein. Chapter 7, p. 281-297. In Owen, M. R. L. and Pen, J. (eds.) *Transgenic Plants: A Production System for Pharmaceutical and Industrial Proteins*. John Wiley and Sons Ltd., NY.
- Gearing, D.P., King, J.A., Gough, N.M. and Nicola, N.A. 1989. Expression cloning of a receptor for human granulocyte-macrophage colony-stimulating factor. *EMBO J.* **8**: 3667-3676.
- Goto, F., Yoshihara, T., Shigemoto, N., Toki, S. and Takaiwa, F. 1999. Iron fortification of rice seed by the soybean ferritin gene. *Nat. Biotechnol.* **17**: 282-286.

- Gough, N.M., Gough, J., Metcalf, D., Kelson, A., Grail, D., Nicola, N.A., Burgess, A.W. and Dunn, A.R. 1984. Molecular cloning of cDNA encoding a murine haematopoietic growth regulator, granulocyte-macrophage colony stimulating factor. *Nature* **309**: 763-767.
- Haq, T.A., Mason, H.S., Clements, J.D. and Arntzen, C.J. 1995. Oral immunization with a recombinant bacterial antigen produced in transgenic plants. *Science* **268**: 714-716.
- Harris, E. L. V. 1989. Chapter 3: Concentration of the extract. p. 154-157. *In Protein purification methods: a practical approach*. Oxford University Press. Edited by Harris, E. L. V. and Angal, S.
- Hayashida, K., Kitamura, T., Gorman, D.M. *et al.* 1990. Molecular cloning of a second subunit of the receptor for human granulocyte-macrophage colony-stimulating factor (GM-CSF): reconstitution of a high-affinity GM-CSF receptor. *Proc. Natl. Acad. Sci. USA* **87**: 9655-9659.
- Herbers, K., Wilke, I., and Sonnewald, U. 1995. A thermostable xylanase from *Clostridium thermocellum* expressed at high levels in the apoplast of transgenic tobacco has no detrimental effects and is easily purified. *Bio/Technology* **13**: 63-66.
- Hercus, T.R., Cambareri, B., Dottore, M., Woodcock, J., Bagley, C.J., Vadas, M.A., Shannon, M.F. and Lopez, A.F. 1994. Identification of residues in the first and fourth helices of human granulocyte-macrophage colony-stimulating factor involved in biologic activity and in binding to the alpha- and beta-chains of its receptor. *Blood* **83**: 3500-3508.
- Hiatt, A., Cafferkey, R. and Bowdish, K. 1989. Production of antibodies in transgenic plants. *Nature* **342**: 76-78.
- Higuchi, W. and Fukazawa, C. 1987. A rice glutelin and a soybean glycinin have evolved from a common ancestral gene. *Gene* **3**: 245-253.
- Higuchi, M., Oh-eda, M., Kuboniwa, H., Tomonoh, K., Shimonaka, Y. and Ochi, N. 1992. Role of sugar chains in the expression of the biological activity of human erythropoietin. *J. Biol. Chem.* **267**: 7703-7709.
- Horsch, R.B., Fry, J.E., Hoffmann, N.L., Eichholtz, D., Rogers, S.G. and Fraley, R.T. 1985. A simple and general method for transferring genes into plants. *Science* **227**: 1229-1231.
- Horvath, H., Huang, J., Wong, O., Kohl, E., Okita, T., Kannangara, C.G. and von Wettstein, D. 2000. The production of recombinant proteins in transgenic barley grains. *Proc. Natl. Acad. Sci. USA* **97**: 1914-1919.

- Hovgaard, D., Mortensen, B.T., Schifter, S. and Nissen, N.I. 1993. Comparative pharmacokinetics of single-dose administration of mammalian and bacterially derived recombinant human granulocyte-macrophage colony-stimulating factor. *Eur. J. Haematol.* **50**(1): 32-36.
- Inumaru, S., Kokuho, T., Denham, S., Denyer, M.S., Momotani, E., Kitamura, S., Corteyn, A., Brookes, S., Parkhouse, R.M. and Takamatsu, H. 1998. Expression of biologically active recombinant porcine GM-CSF by baculovirus gene expression system. *Immunol. Cell Biol.* **76**: 195-201.
- Kaushansky, K., O'Hara, P.J., Hart, C.E., Forstrom, J.W., and Hagen, F.S. 1987. Role of carbohydrate in the function of human granulocyte-macrophage colony-stimulating factor. *Biochemistry* **26**: 4861-4867.
- Kitamura, T., Tange, T., Terasawa, T., Chiba, S., Kuwaki, T., Miyagawa, K., Piao, Y.-F., Miyazono, K., Urabe, A. and Takaku F. 1989. Establishment and characterization of a unique human cell line that proliferates dependently on GM-CSF, IL-3, or erythropoietin. *J. Cellular Physiol.* **140**: 323-334.
- Kitamura, T., Sato, N., Arai, K. *et al.* 1991. Expression cloning of the human IL-3 receptor cDNA reveals a shared  $\beta$  subunit for the human IL-3 and GM-CSF receptors. *Cell* **66**: 1165-1174.
- Kleinerman, E.S., Knowles, R.D., Lachman, L.B. and Gutterman, J.U. 1988. Effect of recombinant granulocyte/macrophage colony-stimulating factor on human monocyte activity *in vitro* and following intravenous administration. *Cancer Res.* **48**: 2604-2609.
- Klingemann, H., Shepherd, J.D., Eaves, C.J. and Eaves, A.C. 1991. The role of erythropoietin and other growth factors in transfusion medicine. *Transfus. Med. Rev.* **5**: 33-47.
- Kuhlemeier, C. 1992. Transcriptional and post-transcriptional regulation of gene expression in plants. *Plant Mol. Biol.* **19**: 1-14.
- Lee, F., Yokota, T., Otsuka, T., Gemmell, L., Larson, N., Luh, J., Arai, K. and Rennick, D. 1985. Isolation of cDNA for a human granulocyte-macrophage colony-stimulating factor by functional expression in mammalian cells. *Proc. Natl. Acad. Sci. USA* **82**: 4360-4364.
- Lee, J.S., Choi, S.J., Kang, H.S., Oh, W.G., Cho, K.H., Kwon, T.H., Kim, D.H., Jang, Y.S. and Yang, M.S. 1997. Establishment of a transgenic tobacco cell suspension culture system for producing murine granulocyte-macrophage colony stimulating factor. *Mol. Cells* **7**: 783-787.

- Leisy, D.J., Hnilo, J., Zhao, Y. and Okita, T.W. 1990. Expression of rice glutelin promoter in transgenic tobacco. *Plant Mol. Biol.* **14**: 41-50.
- Leong, S.R., Flagg, G.M., Lawman, M.J. and Gray, P.W. 1989. Cloning and expression of the cDNA for bovine granulocyte-macrophage colony-stimulating factor. *Vet. Immunol. Immunopathol.* **21**: 261-278.
- Lindemann, A., Riedel, D., Oster, W., Meuer, S.C., Blohm, D., Mertelsmann, R.H. and Herrmann, F. 1988. *J. Immunol.* **140**: 837-840.
- Lindsey, K. 1996. Plant transformation systems. Chapter 1, p. 5-25. In Owen, M. R. L. and Pen, J. (eds.) *Transgenic Plants: A Production System for Industrial and Pharmaceutical Proteins*. Wiley, New York.
- Luo, F., Zhou, S.Z., Cooper, S., Munshi, N.C., Boswell, H.S., Broxmeyer, H.E. and Srivastava, A. 1995. Adeno-associated virus 2-mediated gene transfer and functional expression of the human granulocyte-macrophage colony-stimulating factor. *Exp. Hematol.* **23**: 1261-1267.
- Maliszewski, C.R., Schoenborn, M.A., Cerretti, D.P., Wignall, J.M., Picha, K.S., Cosman, D., Tushinski, R.J., Gillis, S. and Baker, P.E. 1988. Bovine GM-CSF: molecular cloning and biological activity of the recombinant protein. *Mol. Immunol.* **25**: 843-850.
- Mason, H.S., Lam, D. M-K., and Arntzen, C.J. 1992. Expression of hepatitis B surface antigen in transgenic plants. *Proc. Natl. Acad. Sci. USA* **89**: 11745-11749.
- Mason, H.S., Ball, J.M., Shi, J.J., Jiang, X., Estes, M.K. and Arntzen, C.J. 1996. Expression of Norwalk virus capsid protein in transgenic tobacco and potato and its oral immunogenicity in mice. *Proc. Natl. Acad. Sci. USA* **93**: 5335-5340.
- Mason, H.S., Haq, T.A., Clements, J.D., and Arntzen, C.J. 1998. Edible vaccine protects mice against *Escherichia coli* heat-labile enterotoxin (LT): potatoes expressing a synthetic LT-B gene. *Vaccine* **16**: 1336-1343.
- Matsumoto, S., Ikura, K., Ueda, M. and Sasaki, R. 1995. Characterization of human glycoprotein (erythropoietin) produced in cultured tobacco cells. *Plant Mol. Biol.* **27**: 1163-1172.
- McGarvey, P.B., Hammond, J., Dienelt, M.M., Hooper, D.C., Fu, Z.F., Dietzschold, B., Koprowski, H. and Michaels, F.H. 1995. Expression of the rabies virus glycoprotein in transgenic tomatoes. *Bio/Technology* **13**: 1484-1487.
- Metcalf, D. 1991a. Control of granulocytes and macrophages: Molecular, cellular, and clinical aspects. *Science* **254**: 529-533.

- Metcalf, D. 1991b. The Florey Lecture, the colony-stimulating factors: discovery to clinical use. *Phil. Trans. R. Soc. Lond. B* **333**: 147-173.
- Modelska, A., Dietzschold, B., Sleysh, N., Fu, Z.F., Steplewski, K., Hooper, D.C., Koprowski, H. and Yusibov, V. 1998. Immunization against rabies with plant-derived antigen. *Proc. Natl. Acad. Sci. USA* **95**: 2481-2485.
- Moonen, P., Mermoud, J.J., Ernst, J.F., Hirschi M. and DeLamarter, J.F. 1987. Increased biological activity of deglycosylated recombinant human granulocyte/macrophage colony-stimulating factor produced by yeast or animal cells. *Proc. Natl. Acad. Sci. USA* **84**: 4428-4431.
- Muneta, Y., Mori, Y., Shimoji, Y. and Yokomizo, Y. 2000. Porcine interleukin 18: cloning, characterization of the cDNA and expression with the baculovirus system. *Cytokine* **12**: 566-572.
- Nicola, N.A. 1989. Hemopoietic cell growth factors and their receptors. *Ann. Rev. Biochem.* **3**: 45-59.
- Okamoto, M., Nakayama, C., Nakai, M., and Yanagi, H. 1990. Amplification and high-level expression for human granulocyte-macrophage colony-stimulating factor in human lymphoblastoid Namalwa cells. *Bio/Technology* **8**: 550-553.
- Okita, T.W., Hwang, Y.S., Hnilo, J., Kim, W.T., Aryan, A.P., Larson, R. and Krishnan, H.B. 1989. Structure and expression of the rice glutelin multigene family. *J. Biol. Chem.* **264**: 12573-12581.
- Okita, T. W. 1996. Compartmentation of proteins in the endomembrane system of plant cells. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **47**: 327-350.
- Palacpac, N.Q., Yoshida, S., Sakai, H., Kimura, Y., Fujiyama, K., Yoshida, T., Seki, T. 1999. Stable expression of human beta 1,4-galactosyltransferase in plant cells modifies N-linked glycosylation patterns. *Proc. Natl. Acad. Sci. USA* **96**: 4692-4697.
- Polotskaya, A., Zhao, Y., Lilly, M.L. and Kraft, A.S. 1993. A critical role for the cytoplasmic domain of the granulocyte-macrophage colony-stimulating factor alpha-receptor in mediating cell growth. *Cell. Growth Differ.* **4**: 523-531.
- Polotskaya, A., Zhao, Y., Lilly, M.B. and Kraft, A.S. 1994. Mapping the intracytoplasmic regions of the alpha granulocyte-macrophage colony-stimulating factor receptor necessary for cell growth regulation. *J. Biol. Chem.* **269**: 14607-14613.
- Quesniaux, V.J.F. and Jones, T.C. 1998. Chapter 22: Granulocyte-macrophage colony-stimulating factor. *In The Cytokine Handbook*, 3rd ed. Academic Press Limited. Edited by Thomson Angus.

- Rajotte, D., Cadieux, C., Haman, A., Wilkes, B.C., Clark, S.C., Hercus, T., Woodcock, J.A., Lopez, A. and Hoang, T. 1997. Crucial role of the residue R280 at the F'-G' loop of the human granulocyte/macrophage colony-stimulating factor receptor alpha chain for ligand recognition. *J. Exp. Med.* **185**: 1939-1950.
- Robb, L., Drinkwater, C.C., Metcalf, D. *et al.* 1995. Hematopoietic and lung abnormalities in mice with a null mutation of the common  $\beta$  subunit of the receptors for granulocyte-macrophage colony-stimulating factor and interleukins 3 and 5. *Proc. Natl. Acad. Sci. USA* **92**: 9565-9569.
- Robert, L.S., Nozzolillo, C. and Altosaar, I. 1985a. Homology between legumin-like polypeptides from cereals and pea. *Biochem. J.* **226**:847-52.
- Robert, L.S., Nozzolillo, C. and Altosaar, I. 1985b. Homology between rice glutelin and oat 12S globulin. *Biochim. Biophys. Acta.* **829**: 19-26.
- Robert, L.S., Thompson, R.D. and Flavell, R.B. 1989. Tissue-specific expression of a wheat high molecular weight glutenin gene in transgenic tobacco. *Plant Cell* **1**: 569-578.
- Robison, R.L. and Myers, L.A. 1993. Preclinical safety assessment of recombinant human GM-CSF in rhesus monkeys. *Int. Rev. Exp. Pathol.* **34A**: 149-172.
- Ronco, L.V., Doyle, S.E., Raines, M., Park, L.S. and Gasson, J.C. 1995. Conserved amino acids in the human granulocyte-macrophage colony-stimulating factor receptor-binding subunit essential for tyrosine phosphorylation and proliferation. *J. Immunol.* **154**: 3444-3453.
- Ruth, J.H., Esnault, S., Jarzembowski, J.A. and Malter, J.S. 1999. Calcium ionophore upregulation of AUUUA-specific binding protein activity is contemporaneous with granulocyte macrophage colony-stimulating factor messenger RNA stabilization in AML14.3D10 cells. *J. Respir. Cell. Mol. Biol.* **21**: 621-628.
- Sakamoto, K.M., Fraser, J.K., Lee, H.J., Lehman, E. and Gasson, J.C. 1994. Granulocyte-macrophage colony-stimulating factor and interleukin-3 signaling pathways converge on the CREB-binding site in the human egr-1 promoter. *Mol. Cell. Biol.* **14**: 5975-5985.
- Santoli, D., Clark, S.C., Kreider, B.L., Maslin, P.A. and Rovera, G. 1988. Amplification of IL-2-driven T cell proliferation by recombinant human IL-3 and granulocyte-macrophage colony-stimulating factor. *J. Immunol.* **141**: 519-521.
- Sambrook, J., Fritsch, E.F. and Maniatis, T. 1989. Molecular cloning: A Laboratory Manual, 2nd ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

- Sardana, R., Dukiandjiev, S., Giband, M., Cheng, X., Cowan, K., Sauder, C. and Altosaar, I. 1996. Construction and rapid testing of synthetic and modified toxin gene sequences CryIA(b & c) by expression in maize endosperm culture. *Plant Cell Reports* **15**: 677-681.
- Schouten, A., Roosien, J., van Engelen, F.A., de Jong, G.A., Borst-Vreessen, A.W., Zilverentant, J.F., Bosch, D., Stiekema, W.J., Gommers, F.J., Schots, A. and Bakker, J. 1996. The C-terminal KDEL sequence increases the expression level of a single-chain antibody designed to be targeted to both the cytosol and the secretory pathway in transgenic tobacco. *Plant Mol. Biol.* **30**: 781-793.
- Schuh, J.C. and Morrissey, P.J. 1999. Development of a recombinant growth factor and fusion protein: lessons from GM-CSF. *Toxicol. Pathol.* **27**: 72-77.
- Shen, Y., Baker, E., Callen, D.F., Sutherland, G.R., Willson, T.A., Rakar, S. and Gough, N.M. 1992. Localization of the human GM-CSF receptor beta chain gene (CSF2RB) to chromosome 22q12.2-->q13.1. *Cytogenet. Cell. Genet.* **61**: 175-177.
- Shi, X., Qin, J., Zhu, J. and Zhu, D. 1996. Expression of biologically active human granulocyte-macrophage colony-stimulating factor in the silkworm (*Bombyx mori*). *Biotechnol. Appl. Biochem.* **24**: 245-249.
- Shyur, L.F., Wen, T.N. and Chen, C.S. 1992. cDNA cloning and gene expression of the major prolamins of rice. *Plant Mol. Biol.* **20**: 323-632.
- Sieff, C.A. 1987. Hematopoietic growth factors. *J. Clin. Invest.* **19**: 1549-1557.
- Sijmons, P.C., Dekker, B.M.M., Schrammeijer, B., Verwoerd T.C., van den Elzen, P.J.M., and Hoekema, A. 1990. Production of correctly processed human serum albumin in transgenic plants. *Bio/Technology* **8**: 217-221.
- Smith, L.T., Hohaus, S., Gonzalez, D.A., Dziennis, S.E. and Tenen, D.G. 1996. PU.1 (Spi-1) and C/EBP alpha regulate the granulocyte colony-stimulating factor receptor promoter in myeloid cells. *Blood* **88**: 1234-1247.
- Spielholz, C., Haeney, M.L., Morrison, M.E., Houghthon, A.N., Vera, J.C. and Golde, D.W. 1995. Granulocyte-macrophage colony-stimulating factor signals for increased glucose uptake in human melanoma cells. *Blood* **85**:973-980.
- Stoger, E., Vaquero, C., Torres, E., Sack, M., Nicholson, L., Drossard, J., Williams, S., Keen, D., Perrin, Y., Christou, P. and Fischer, R. 2000. Cereal crops as viable production and storage systems for pharmaceutical scFv antibodies. *Plant. Mol. Biol.* **42**:583-590.
- Sullivan, R., Fredette, J.P., Socinski, M., Elias, A., Antman, K., Schnipper, L. and Griffin, J.D. 1989. Enhancement of superoxide anion release by granulocytes harvested

- from patients receiving granulocyte-macrophage colony-stimulating factor. *Br. J. Hematol.* **71**: 475-484.
- Tackaberry, E., Dudani, A.K., Prior, F., Tocchi, M., Sardana, R., Altosaar, I. and Ganz, P.R. 1999. Development of biopharmaceuticals in plant expression systems: cloning, expression and immunological reactivity of human cytomegalovirus glycoprotein B (UL55) in seeds of transgenic tobacco. *Vaccine* **17**: 3020-3029.
- Tacket, C.O., Mason, H.S., Losonsky, G., Estes, M.K., Levine, M.M. and Arntzen, C.J. 2000. Human immune responses to a novel norwalk virus vaccine delivered in transgenic potatoes. *J. Infect. Dis.* **182**: 302-305.
- Takaiwa, F., Kikuchi, S., and Oono, K. 1989. The complete nucleotide sequence of new type cDNA coding for rice storage protein glutelin. *Nucl. Acids Res.* **17**: 32-89.
- Takaiwa, F., Oono, K. and Kato, A. 1991. Analysis of the 5' flanking region responsible for the endosperm-specific expression of a rice glutelin chimeric gene in transgenic tobacco. *Plant Mol. Biol.* **16**: 49-58.
- Takaiwa, F., Yamanouchi, U., Yoshihara, T., Washida, H., Tanabe, F., Kato, A. and Yamada, K. 1996. Characterization of common cis-regulatory elements responsible for the endosperm-specific expression of members of the rice glutelin multigene family. *Plant Mol. Biol.* **30**: 1207-1221.
- Tavladoraki, P., Benvenuto, E., Trinca, S., De Martinis, D., Cattaneo, A. and Galeffi, P. 1993. Transgenic plants expressing a functional single-chain Fv antibody are specifically protected from virus attack. *Nature* **366**: 469-472.
- Thanavala, Y., Yang, Y.F., Lyons, P., Mason, H.S. and Arntzen, C. 1995. Immunogenicity of transgenic plant-derived hepatitis B surface antigen. *Proc. Natl. Acad. Sci. USA* **92**: 3358-3361.
- Torres, E., Vaquero, C., Nicholson, L., Sack, M., Stoger, E., Drossard, J., Christou, P., Fischer, R. and Perrin, Y. 1999. Rice cell culture as an alternative production system for functional diagnostic and therapeutic antibodies. *Transgenic Res.* **8**: 441-449.
- Tsuboi, A., Muramatsu, M., Tsutsumi, A., Arai, K., Arai, N. 1994. Calcineurin activates transcription from the GM-CSF promoter in synergy with either protein kinase C or NF-kappa B/AP-1 in T cells. *Biochem. Biophys. Res. Commun.* **199**: 1064-72.
- Tsuda, E., Kawanishi, G., Ueda, M., Masuda, S. and Sasaki, R. 1990. The role of carbohydrate in recombinant human erythropoietin. *Eur. J. Biochem.* **188**: 405-411.

- Turpen, T.H., Reindl, S.J., Charoenvit, Y., Hoffman, S.L., Fallarme, V. and Grill, L.K. 1995. Malarial epitopes expressed on the surface of recombinant tobacco mosaic virus. *Bio/Technology* **13**: 53-57.
- van Engelen, F.A., Schouten, A., Molthoff, J.W., Roosien, J., Salinas, J., Dirkse, W.G., Schots, A., Bakker, J., Gommers, F.J., Jongsma, M.A., *et al.* 1994. Coordinate expression of antibody subunit genes yields high levels of functional antibodies in roots of transgenic tobacco. *Plant Mol. Biol.* **26**: 1701-1710.
- Van Rooijen, G.J.H. and Moloney, M.M. 1995. Plant seed oil-bodies as carriers for foreign proteins. *Bio/Technology* **13**: 72-77.
- Vandekerckhove, J., Van Damme, J., Van Lijsebettens, M., Botterman, J., De Block, M., Vandewiele, M., De Clercq, A., Leemans, J., Van Montagu, M. and Krebbers, E. 1989. Enkephalins produced in transgenic plants using modified 2S seed storage proteins. *Bio/Technology* **7**: 929-932.
- Verwoerd, T.C., Dekker, B.M.M. and Hoekema, A. 1989. A small scale procedure for the rapid isolation of plant mRNAs. *Nucl. Acids Res.* **17**: 2362.
- Vitale, A., Ceriotti, A. and Denecke, J. 1993. The role of endoplasmic reticulum in protein synthesis, modification and intracellular transport. *J. Exp. Botany* **44**: 1417-1444.
- Walden, R. and Wingender, R. 1995. Gene-transfer and plant-regeneration techniques. *TIBTECH* **13**: 324-337.
- Watanabe, S., Itoh, T. and Arai, K. 1996. JAK2 is essential for activation of c-fos and c-myc promoters and cell proliferation through the human granulocyte-macrophage colony-stimulating factor receptor in BA/F3 cells. *J. Biol. Chem.* **271**: 12681-12686.
- Weiss, M., Yokoyama, C., Shikama, Y., Naugle, C., Druker, B. and Sieff, C.A. 1993. Human granulocyte-macrophage colony-stimulating factor receptor signal transduction requires the proximal cytoplasmic domains of the alpha and beta subunits. *Blood* **82**: 3298-306.
- Wong, G.G., Witek, J.S., Temple, P.A. *et al.* 1985a. Human GM-CSF: molecular cloning of the complementary DNA and purification of the natural and recombinant proteins. *Science* **228**: 810-815.
- Wong, G.G., Witek, J.S., Temple, P.A., *et al.* 1985b. Isolation of cDNAs encoding human and gibbon GM-CSF. *Prog. Clin. Biol. Res.* **191**: 351-366.
- Woodcock, J.M., McClure, B.J., Stomski, F.C., Elliott, M.J., Bagley, C.J. and Lopez, A.F. 1997. The human granulocyte-macrophage colony-stimulating factor (GM-CSF)

receptor exists as a preformed receptor complex that can be activated by GM-CSF, interleukin-3, or interleukin-5. *Blood* **90**: 3005-3017.

- Yamagata, H. and Tanaka, K. 1986. The site of synthesis and accumulation of rice storage proteins. *Plant Cell. Physiol.* **27**: 135-145.
- Yusibov, V., Modelska, A., Steplewski, K., Agadjanyan, M., Weiner, D., Hooper, D.C. and Koprowski, H. 1997. Antigens produced in plants by infection with chimeric plant viruses immunize against rabies virus and HIV-1. *Proc. Natl. Acad. Sci. USA* **94**: 5784-5788.
- Zeitlin, L., Olmsted, S.B., Moench, T.R., Co, M.S., Martinell, B.J., Paradkar, V.M., Russell, D.R., Queen, C., Cone, R.A. and Whaley, K. 1998. A humanized monoclonal antibody produced in transgenic plants for immunoprotection of the vagina against genital herpes. *Nat. Biotechnol.* **16**: 1361-1364.
- Zhao, Y., Leisy, D. J., and Okita, T. W. 1994. Tissue-specific expression and temporal regulation of the rice glutelin Gt3 gene are conferred by at least two spatially separated cis-regulatory elements. *Plant Mol. Bio.* **25**: 429-436.
- Zhao, Y., Okita, T.W. 1995. Interactions of the glutelin Gt3 5' flanking regulatory regions with rice nuclear proteins. *Plant Cell. Physiol.* **36**: 1657-1667
- Zheng, Z., Kawagoe, Y., Xiao, S., Li, Z., Okita, T. Hau, T.L., Lin, A. and Murai, N. 1993. 5' distal and proximal cis-acting regulator elements are required for developmental control of a rice seed storage protein glutelin gene. *Plant J.* **4**: 357-366.
- Zheng, Z.W., Sumi, K., Tanaka, K., and Murai, N. 1995. The bean seed storage protein beta-phaseolin is synthesized, processed, and accumulated in the vacuolar type-II protein bodies of transgenic rice endosperm. *Plant Physiol.* **109**: 777-786.