PROMOTERS FOR SUGARCANE TRANSFORMATION: 
ISOLATION OF SPECIFIC SEQUENCES AND EVALUATION 
OF rolC

by

Sarita Groenewald

Submitted in fulfilment of the academic 
requirements for the degree of 
Philosophiae Doctor 
in the 
Department of Botany, 
University of Natal

Supervisor: Prof. F.C. Botha

Pietermaritzburg
1997
The experimental work described in this thesis was carried out in the Biotechnology Department of the South African Sugar Association Experiment Station (SASEX), Mount Edgecombe, from January 1994 to December 1997, under the supervision of Prof. F.C.Botha.

These studies represent original work by the author and have not otherwise been submitted in any form to another university. Where use was made of the work of others, it has been duly acknowledged in the text.
ACKNOWLEDGEMENTS

- My sincere thanks go to my supervisor, Prof. Frikkie Botha, for his advice, encouragement and support through the duration of this study. His enthusiasm and inquisitiveness have made this study an exciting experience.
- My thanks and appreciation are also extended to the following:
  - Hennie Groenewald for his advice and assistance with the construction of expression vectors and photography;
  - Dr. Barbara Huckett for her advice on plant molecular biology and Sandy Snyman for her help and advice on sugarcane tissue culture;
  - Sharon McFarlane (Plant Pathology Department, SASEX), James Wesley-Smith (Electron microscopy unit, UND), Edith Elliot (Biochemistry Department, UNP) and Vihitha Beharee (Microbiology Department, UNP) for their assistance with the microscopy;
  - Peter Hare (Botany Department, UNP) for the subcloning of the mGFP4 reporter gene into pBluescript(SK+);
  - Chris Davies (CSIRO, Australia) for advice on northern blots;
  - Julie Barnes for the careful and critical reading of this thesis, and for her friendship and enthusiasm;
  - My colleagues in the Biotechnology Department (SASEX) for their friendship, support and encouragement;
  - My family and friends for their love and support through the years;
  - The South African Sugar Association and the Foundation for Research Development for financial support.
- A special thanks go to my husband, Hennie, for his love, patience and support during this study.
- I am most grateful for the privilege of studying God's creation at molecular level.
ABSTRACT

Increasing the sucrose yield and the disease resistance of plants are two major objectives of the transgenic sugarcane plant programme in South Africa. The sugarcane culm has thus been identified as one of the main target areas for transgene expression. A shortage of reliable promoter elements as well as patent limitations have necessitated the isolation of promoters that are preferentially expressed in the sugarcane culm. In the present study two different approaches were followed to isolate such promoters, and the bacterial promoter, rolC, was evaluated for tissue-specific expression in sugarcane.

Differential display is a non-directed technique that was used to identify genes that are differentially expressed in the mature sugarcane culm. The original method was modified, and four putative culm-preferential fragments were isolated. Sequence and hybridisation analyses revealed that these fragments were false positives, and could therefore not be used to obtain a culm-specific promoter.

Activity of the Agrobacterium rolC promoter was evaluated by analysing expression patterns of two reporter genes in the mature culm of transgenic sugarcane plants. Nucleic acid analyses indicated that the foreign DNA was incorporated into the sugarcane genome, and that mRNA transcripts were produced. Histochemical analysis was done to visualise rolC-driven GUS and GFP expression in the mature sugarcane culm. In both cases the reporter gene expression was restricted to the vascular bundles and specifically to the phloem.

A directed approach was followed to isolate the gene and subsequently the promoter of the β-subunit of pyrophosphate-dependent phosphofructokinase (PFP-β). An incomplete cDNA clone was obtained from a mature culm cDNA library, and was used for the screening of a sugarcane genomic library. Two clones containing different parts of the PFP-β gene were isolated. A Deletion Factory™ system was used to analyse the clone containing the 5' end of the gene. The first five exons and 1747 bp of the 5' flanking region of the gene were sequenced. Preliminary activity
analysis of the promoter region was done by constructing two expression vectors, and analysing transient GUS expression in sugarcane callus. Results indicated that the promoter is capable of driving foreign gene expression in callus. Transient expression levels were lower than that of the maize \textit{Ubi-1} promoter. Further analysis of the 5' flanking region will be done to establish whether \textit{cis}-acting elements outside the analysed area have an influence on the activity of the promoter.
## CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preface</td>
<td>ii</td>
</tr>
<tr>
<td>Acknowledgements</td>
<td>iii</td>
</tr>
<tr>
<td>Abstract</td>
<td>iv</td>
</tr>
<tr>
<td>List of figures</td>
<td>xii</td>
</tr>
<tr>
<td>List of tables</td>
<td>xiv</td>
</tr>
<tr>
<td>Abbreviations</td>
<td>xv</td>
</tr>
<tr>
<td><strong>Chapter 1</strong> Introduction</td>
<td>1</td>
</tr>
<tr>
<td><strong>Chapter 2</strong> Literature Review</td>
<td>6</td>
</tr>
<tr>
<td>2.1 Introduction</td>
<td>6</td>
</tr>
<tr>
<td>2.2 Plant Promoters</td>
<td>6</td>
</tr>
<tr>
<td>2.2.1 Constitutive promoters</td>
<td>9</td>
</tr>
<tr>
<td>2.2.2 Inducible promoters</td>
<td>10</td>
</tr>
<tr>
<td>2.2.2.1 Temperature-inducible promoters:</td>
<td></td>
</tr>
<tr>
<td>Heat-shock promoters</td>
<td>10</td>
</tr>
<tr>
<td>2.2.2.2 Temperature-inducible promoters:</td>
<td></td>
</tr>
<tr>
<td>Cold-induced promoters</td>
<td>13</td>
</tr>
<tr>
<td>2.2.2.3 Hormone-inducible promoters</td>
<td>13</td>
</tr>
<tr>
<td>2.2.2.4 Pathogen-inducible promoters</td>
<td>16</td>
</tr>
<tr>
<td>i) PR-1 Genes</td>
<td>17</td>
</tr>
<tr>
<td>ii) PR-2 and PR-3 Genes</td>
<td>17</td>
</tr>
<tr>
<td>iii) PR-5 Genes</td>
<td>19</td>
</tr>
<tr>
<td>iv) PR-6 Genes</td>
<td>19</td>
</tr>
<tr>
<td>v) PR-10 Genes</td>
<td>19</td>
</tr>
<tr>
<td>2.2.2.5 Light-inducible promoters</td>
<td>20</td>
</tr>
<tr>
<td>2.2.3 Tissue-specific expression</td>
<td>22</td>
</tr>
<tr>
<td>2.2.4 Organelle Promoters</td>
<td>24</td>
</tr>
<tr>
<td>2.3 Isolation and analysis of plant promoters</td>
<td>25</td>
</tr>
</tbody>
</table>
| 2.3.1 Non-directed methods for the identification and
isolation of useful promoters 25
2.3.1.1 Differential display of mRNA by PCR 27
i) RNA template for RT-PCR 28
ii) Primers used for RT-PCR 29
iii) Display of amplified cDNA fragments 31
iv) Isolation of differentially expressed cDNA fragments 32
2.3.2 Directed techniques for the isolation of useful promoters 33
2.3.3 Isolation of tissue-specific promoters from sugarcane 34
2.4 Plant transformation 36
2.4.1 Transformation via particle bombardment 37
2.4.2 Vectors for plant transformation 39
  2.4.2.1 Selectable marker genes 42
  2.4.2.2 5' Regulatory sequences in plant expression vectors 42
i) The rolC promoter of Agrobacterium rhizogenes 44
2.4.2.3 Reporter genes 47
i) β-glucuronidase (GUS) 48
ii) Green fluorescent protein (GFP) 49
2.4.3 Transformation of sugarcane 52
2.5 Pyrophosphate-dependent phosphofructokinase (PFP) 56
  2.5.1 PFP in sugarcane 59

Chapter 3 Use of a differential display technique to isolate differentially expressed gene sequences from sugarcane culm tissue 61
3.1 Abstract 61
3.2 Introduction 62
3.3 Materials and methods 65
  3.3.1 RNA preparation 65
  3.3.2 Reverse transcription 65
  3.3.3 Amplification of cDNA 66
3.3.4. Preparation of molecular weight markers for agarose gel electrophoresis 66
3.3.5 Isolation and reamplification of possible tissue-specific fragments 67
3.3.6 Characterisation of isolated fragments by RT-PCR Southern blotting 68
3.3.7 Cloning of cDNA fragments 68
3.3.8 Sequence determination 69
3.3.9 Genomic Southern blot analysis 69
3.3.10 Northern blot analysis of isolated cDNA fragments 70
3.3.11 Characterisation of the culm-preferential fragment, SA11, by specific amplification 71
3.3.12 Characterisation of the culm-preferential fragment, SA11, by specific RT-PCR 71

3.4 Results 72
3.4.1 Optimisation of PCR template concentration 72
3.4.2 Differential display: the comparison of random amplified fragments from different sugarcane tissues 74
3.4.3 Characterisation of isolated fragments using RT-PCR Southern blots 75
3.4.4 Sequence determination of the cloned fragments 76
3.4.5 Genomic Southern blot analysis of culm-preferential fragments 81
3.4.6 Northern blot analysis of culm-preferential fragments 82
3.4.7 Characterisation of the culm-preferential fragment, SA11 82

3.5 Discussion 85
3.6 Conclusions 89
Chapter 4 Expression of reporter genes under control of the rolC promoter in sugarcane

4.1 Abstract 90
4.2 Introduction 90
4.3 Materials and methods 93
4.3.1 Construction of expression vectors 93
4.3.2 Sugarcane tissue culture 94
4.3.3 Sugarcane transformation 94
4.3.4 Characterisation of transformed plants: DNA analyses 95
4.3.5 Characterisation of transformed plants: RNA analysis 96
4.3.6 Characterisation of transformed plants: Histochemical analysis 96
4.4 Results 97
4.4.1 Construction of the expression vectors, pGUS800 and pGFP800 97
4.4.2 Sugarcane transformation 99
4.4.3 Characterisation of transformed plants: DNA analyses 99
4.4.4 Characterisation of transformed plants: RNA analysis 100
4.4.5 Characterisation of transformed plants: Histochemical analysis 101
4.5 Discussion 105
4.6 Conclusions 108

Chapter 5 Isolation of the pyrophosphate-dependent phosphofructokinase β-subunit (PFP-β) promoter 109

5.1 Abstract 109
5.2 Introduction 110
5.3 Materials and methods 112
5.3.1 Degenerate and specific primer design 112
5.3.2 Amplification of PFP fragments from cDNA libraries 112
5.3.3 Cloning of cDNA fragments 113
5.3.4 Sequence determination of cDNA fragments 113
5.3.5 Screening cDNA and genomic DNA libraries 113
5.3.6 Southern blot analysis 115
5.3.7 Northern blot analysis 115
5.3.8 Isolation of Lambda-DNA 115
5.3.9 Characterisation of positive genomic library clones 116
5.3.10 Construction of a Deletion Factory™ 117
5.3.11 Sequence determination of cloned inserts 119
5.3.12 Isolation of the 5' flanking region of sugarcane PFP-β 119
5.3.13 Construction of expression vectors with the GUS reporter gene 120
5.3.14 Transient GUS expression in sugarcane callus 120

5.4 Results 121
5.4.1 Amplification of PFP-β from cDNA libraries 121
5.4.2 Screening the sugarcane mature culm cDNA library 122
5.4.3 Southern and northern blot analysis 122
5.4.4 Screening of a sugarcane genomic library and characterisation of positive clones 123
5.4.5 Construction of a Deletion Factory™ 125
5.4.6 Comparative sequence analysis of the sugarcane, castor bean and potato PFP-β coding sequences 126
5.4.7 Sequence analysis of the sugarcane PFP-β promoter 127
5.4.8 Transient GUS expression analysis in sugarcane 129

5.5 Discussion 131
5.6 Conclusions 134

Chapter 6 Conclusions 135
Appendix I: Nucleotide sequence of the first five exons of the PFP-β gene from sugarcane
LIST OF FIGURES

Figure 3.1 Molecular weight markers used in this study. 67
Figure 3.2 Amplification of cDNA using the oligo(dT)$_{15}$ and a random decamer primer. 72
Figure 3.3 Optimisation of the PCR template concentration for differential display. 73
Figure 3.4 Optimisation of the PCR template concentration for differential display. 74
Figure 3.5 Amplification profiles of four sugarcane tissues using different random primers. 75
Figure 3.6 Identification and characterisation of culm-preferential fragment 2A08. 77
Figure 3.7 Identification and characterisation of culm-preferential fragment SA11. 77
Figure 3.8 Identification and characterisation of culm-preferential fragment 7C16. 78
Figure 3.9 Identification and characterisation of culm-preferential fragment 7F18. 78
Figure 3.10 Complete nucleotide sequence of fragment 2A08. 79
Figure 3.11 Complete nucleotide sequence of fragment SA11. 79
Figure 3.12 Complete nucleotide sequence of fragment 7C16. 80
Figure 3.13 Partial nucleotide sequence of fragment 7F18. 80
Figure 3.14 Specific amplification of fragment SA11 from cDNA and genomic DNA. 83
Figure 3.15 Determining the origin of fragment SA11. 83
Figure 3.16 Dependence of the formation of fragment SA11 on reverse transcription. 84
Figure 4.1 Expression of reporter genes in transformed sugarcane callus. 95
Figure 4.2 Restriction analysis of the reporter gene expression vectors, pGUS800 and pGFP800. 98
Figure 4.3 Schematic representation of the two reporter gene expression vectors, pGUS800 and pGFP800. 98
Figure 4.4 PCR amplification of reporter genes from transformed plants. 99
Figure 4.5 Southern blot analysis of transformed plants to confirm the presence of the reporter genes. 100
Figure 4.6 RT-PCR of transformed plant RNA to confirm the presence of the reporter gene transcripts. 101
Figure 4.7 Histochemical analysis of GUS expression in sugarcane cells. 102
Figure 4.8 Visualisation of GFP expression in phloem cells using different filter sets and microscopes. 103
Figure 4.9 Visualisation of GFP expression in sugarcane cells. 104
Figure 5.1 Amplification of PFP-\(\beta\) cDNA from leafroll and mature culm cDNA libraries.

Figure 5.2 Southern blot analysis of PFP-\(\beta\) in sugarcane and other plants.

Figure 5.3 Northern blot analysis of PFP-\(\beta\) in sugarcane leafroll.

Figure 5.4 Characterisation of positive genomic library clones by restriction and hybridisation analyses.

Figure 5.5 Deletion Factory: Counterclockwise deletion subclones (K).

Figure 5.6 Sequence of 1 747 bp of the 5' flanking region of sugarcane PFP-\(\beta\).

Figure 5.7 Schematic representation of the expression vectors, pGUS1072 and pGUS1355.

Figure 5.8 Transient GUS expression in sugarcane callus using different promoter constructs.
LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Cis-regulatory regions identified in plant inducible promoters.</td>
<td>11</td>
</tr>
<tr>
<td>3.1</td>
<td>Summary of results obtained from sequence similarity BLAST searches of the culm-preferential fragments.</td>
<td>81</td>
</tr>
<tr>
<td>3.2</td>
<td>Characterisation of culm-preferential fragments by genomic Southern blot analysis</td>
<td>82</td>
</tr>
<tr>
<td>4.1</td>
<td>Restriction analysis of the reporter gene expression vectors, pGUS800 and pGFP800.</td>
<td>97</td>
</tr>
<tr>
<td>5.1</td>
<td>Oligonucleotide primers designed for the amplification of PFP-β cDNA and 5’ flanking region.</td>
<td>112</td>
</tr>
<tr>
<td>5.2</td>
<td>Restriction sites (A) and promoter motifs (B) in the promoter region of PFP-β.</td>
<td>128</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
<td></td>
</tr>
<tr>
<td>--------------</td>
<td>-------------</td>
<td></td>
</tr>
<tr>
<td>2,4-D</td>
<td>2,4-dichloro-phenoxycetic acid</td>
<td></td>
</tr>
<tr>
<td>35S</td>
<td>Cauliflower mosaic virus 35S promoter</td>
<td></td>
</tr>
<tr>
<td>4-MUG</td>
<td>4-methylumbelliferyl-β-glucuronide</td>
<td></td>
</tr>
<tr>
<td>5' UTR</td>
<td>5' untranslated region</td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>clock-wise deletion subclones</td>
<td></td>
</tr>
<tr>
<td>aa</td>
<td>amino acid</td>
<td></td>
</tr>
<tr>
<td>ABA</td>
<td>abscisic acid</td>
<td></td>
</tr>
<tr>
<td>ABRE</td>
<td>abscisic acid response element</td>
<td></td>
</tr>
<tr>
<td>Act1</td>
<td>rice actin-1 promoter</td>
<td></td>
</tr>
<tr>
<td>Adh1</td>
<td>maize alcohol dehydrogenase-1 promoter</td>
<td></td>
</tr>
<tr>
<td>Arg</td>
<td>arginine</td>
<td></td>
</tr>
<tr>
<td>Asp</td>
<td>asparagine</td>
<td></td>
</tr>
<tr>
<td>ATP</td>
<td>adenosine 5'-triphosphate</td>
<td></td>
</tr>
<tr>
<td>BAC</td>
<td>bacterial artificial chromosome</td>
<td></td>
</tr>
<tr>
<td>BLAST</td>
<td>basic local alignment search tool</td>
<td></td>
</tr>
<tr>
<td>bp</td>
<td>base pair</td>
<td></td>
</tr>
<tr>
<td>BP</td>
<td>band-pass</td>
<td></td>
</tr>
<tr>
<td>BSA</td>
<td>bovine serum albumin</td>
<td></td>
</tr>
<tr>
<td>cab</td>
<td>chlorophyll a/b binding protein</td>
<td></td>
</tr>
<tr>
<td>CaMV</td>
<td>cauliflower mosaic virus</td>
<td></td>
</tr>
<tr>
<td>cat</td>
<td>chloramphenicol acyltransferase</td>
<td></td>
</tr>
<tr>
<td>cDNA</td>
<td>complementary deoxyribonucleic acid</td>
<td></td>
</tr>
<tr>
<td>CLP</td>
<td>chlorallactophenol</td>
<td></td>
</tr>
<tr>
<td>dbEST</td>
<td>expressed sequence tag database</td>
<td></td>
</tr>
<tr>
<td>DEPC</td>
<td>diethyl pyrocarbonate</td>
<td></td>
</tr>
<tr>
<td>DNA</td>
<td>deoxyribonucleic acid</td>
<td></td>
</tr>
<tr>
<td>dNTP</td>
<td>deoxynucleoside triphosphate</td>
<td></td>
</tr>
<tr>
<td>DRE</td>
<td>dehydration response element</td>
<td></td>
</tr>
<tr>
<td>dT</td>
<td>deoxithymidine</td>
<td></td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
<td></td>
</tr>
<tr>
<td>--------------</td>
<td>-------------</td>
<td></td>
</tr>
<tr>
<td>EDD</td>
<td>enhanced differential display</td>
<td></td>
</tr>
<tr>
<td>EDTA</td>
<td>ethylenediaminetetraacetate</td>
<td></td>
</tr>
<tr>
<td>EMBL</td>
<td>European molecular biology laboratory</td>
<td></td>
</tr>
<tr>
<td>ERE</td>
<td>elicitor-responsive element</td>
<td></td>
</tr>
<tr>
<td>EST</td>
<td>expressed sequence tag</td>
<td></td>
</tr>
<tr>
<td>FITC</td>
<td>fluorescein isothiocyanate</td>
<td></td>
</tr>
<tr>
<td>Fru-1,6-P₂</td>
<td>fructose 1,6-bisphosphate</td>
<td></td>
</tr>
<tr>
<td>Fru-2,6-P₂</td>
<td>fructose 2,6-bisphosphate</td>
<td></td>
</tr>
<tr>
<td>Fru-6-P</td>
<td>fructose 6-phosphate</td>
<td></td>
</tr>
<tr>
<td>g</td>
<td>gravitational acceleration (9.806 m s⁻¹)</td>
<td></td>
</tr>
<tr>
<td>G418</td>
<td>geneticin</td>
<td></td>
</tr>
<tr>
<td>GA</td>
<td>giberellic acid</td>
<td></td>
</tr>
<tr>
<td>GFP</td>
<td>green fluorescent protein</td>
<td></td>
</tr>
<tr>
<td>Gly</td>
<td>glycine</td>
<td></td>
</tr>
<tr>
<td>GUS</td>
<td>β-glucuronidase</td>
<td></td>
</tr>
<tr>
<td>HDC</td>
<td>high definition color</td>
<td></td>
</tr>
<tr>
<td>HSE</td>
<td>heat shock element</td>
<td></td>
</tr>
<tr>
<td>HSF</td>
<td>heat shock transcription factor</td>
<td></td>
</tr>
<tr>
<td>iPCR</td>
<td>inverse polymerase chain reaction</td>
<td></td>
</tr>
<tr>
<td>K</td>
<td>Counterclockwise deletion subclones</td>
<td></td>
</tr>
<tr>
<td>kb</td>
<td>kilobase pair</td>
<td></td>
</tr>
<tr>
<td>lac</td>
<td>β-galactosidase</td>
<td></td>
</tr>
<tr>
<td>LP</td>
<td>long-pass</td>
<td></td>
</tr>
<tr>
<td>LRP</td>
<td>light-responsive promoter</td>
<td></td>
</tr>
<tr>
<td>luc</td>
<td>firefly luciferase</td>
<td></td>
</tr>
<tr>
<td>mas</td>
<td>mannopine synthase</td>
<td></td>
</tr>
<tr>
<td>MAST</td>
<td>magnet-assisted subtraction technique</td>
<td></td>
</tr>
<tr>
<td>MJ</td>
<td>methyl jasmonate</td>
<td></td>
</tr>
<tr>
<td>MOPS</td>
<td>3-[[N-morpholino]-propanesulphonic acid</td>
<td></td>
</tr>
<tr>
<td>MRC</td>
<td>Medical Research Council</td>
<td></td>
</tr>
<tr>
<td>mRNA</td>
<td>messenger ribonucleic acid</td>
<td></td>
</tr>
<tr>
<td>Term</td>
<td>Description</td>
<td></td>
</tr>
<tr>
<td>------------</td>
<td>-------------------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>MS</td>
<td>Murashige and Skoog medium</td>
<td></td>
</tr>
<tr>
<td>M-MuLV</td>
<td>Moloney Murine Leukemia Virus</td>
<td></td>
</tr>
<tr>
<td>nos</td>
<td>nopaline synthase</td>
<td></td>
</tr>
<tr>
<td>NPT</td>
<td>neomycin phosphotransferase</td>
<td></td>
</tr>
<tr>
<td>NR</td>
<td>non-redundant</td>
<td></td>
</tr>
<tr>
<td>OCS</td>
<td>octapine synthase</td>
<td></td>
</tr>
<tr>
<td>OD</td>
<td>optical density</td>
<td></td>
</tr>
<tr>
<td>ORF</td>
<td>open reading frame</td>
<td></td>
</tr>
<tr>
<td>ori</td>
<td>origin of replication</td>
<td></td>
</tr>
<tr>
<td>PAT</td>
<td>phosphinothricin acyltransferase</td>
<td></td>
</tr>
<tr>
<td>PCR</td>
<td>polymerase chain reaction</td>
<td></td>
</tr>
<tr>
<td>PFK</td>
<td>phosphofructokinase</td>
<td></td>
</tr>
<tr>
<td>PFP</td>
<td>pyrophosphate-dependent phosphofructokinase</td>
<td></td>
</tr>
<tr>
<td>pfu</td>
<td>plaque forming units</td>
<td></td>
</tr>
<tr>
<td>PG</td>
<td>polygalacturonase</td>
<td></td>
</tr>
<tr>
<td>Pi</td>
<td>inorganic phosphate</td>
<td></td>
</tr>
<tr>
<td>PIG</td>
<td>particle inflow gun</td>
<td></td>
</tr>
<tr>
<td>PLRV</td>
<td>potato leafroll virus</td>
<td></td>
</tr>
<tr>
<td>PPI</td>
<td>pyrophosphate</td>
<td></td>
</tr>
<tr>
<td>PR</td>
<td>pathogenesis-related</td>
<td></td>
</tr>
<tr>
<td>PTT</td>
<td>phosphinothricin</td>
<td></td>
</tr>
<tr>
<td>Pu</td>
<td>purine</td>
<td></td>
</tr>
<tr>
<td>PVP</td>
<td>polyvinylpyrrolidone</td>
<td></td>
</tr>
<tr>
<td>R</td>
<td>pUC18/M13 reverse primer</td>
<td></td>
</tr>
<tr>
<td>RAPD</td>
<td>random amplification of polymorphic DNAs</td>
<td></td>
</tr>
<tr>
<td>rbcS</td>
<td>ribulose 1,5-bisphosphate carboxylase-oxygenase</td>
<td></td>
</tr>
<tr>
<td>Ri</td>
<td>root-inducing</td>
<td></td>
</tr>
<tr>
<td>RNA</td>
<td>ribonucleic acid</td>
<td></td>
</tr>
<tr>
<td>RNase</td>
<td>ribonuclease</td>
<td></td>
</tr>
<tr>
<td>rRNA</td>
<td>ribosomal RNA</td>
<td></td>
</tr>
<tr>
<td>RT</td>
<td>reverse transcription</td>
<td></td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
<td></td>
</tr>
<tr>
<td>--------------</td>
<td>-------------</td>
<td></td>
</tr>
<tr>
<td>RT-PCR</td>
<td>reverse transcription polymerase chain reaction</td>
<td></td>
</tr>
<tr>
<td>SA</td>
<td>salicylic acid</td>
<td></td>
</tr>
<tr>
<td>SAR</td>
<td>systemic acquired resistance</td>
<td></td>
</tr>
<tr>
<td>SASEX</td>
<td>South African Sugar Association Experiment Station</td>
<td></td>
</tr>
<tr>
<td>SBDS</td>
<td>Southern blot differential screening</td>
<td></td>
</tr>
<tr>
<td>SDS</td>
<td>sodium dodecyl sulphate</td>
<td></td>
</tr>
<tr>
<td>Ser</td>
<td>serine</td>
<td></td>
</tr>
<tr>
<td>Sh1</td>
<td><em>Shrunken-1</em> (maize sucrose synthase1 gene)</td>
<td></td>
</tr>
<tr>
<td>spp.</td>
<td>species</td>
<td></td>
</tr>
<tr>
<td>TL-DNA</td>
<td>transfer DNA of <em>Agrobacterium</em></td>
<td></td>
</tr>
<tr>
<td>Ti</td>
<td>tumor-inducing</td>
<td></td>
</tr>
<tr>
<td>Tm</td>
<td>temperature, melting</td>
<td></td>
</tr>
<tr>
<td>TMV</td>
<td>tobacco mosaic virus</td>
<td></td>
</tr>
<tr>
<td>tRNA</td>
<td>transfer RNA</td>
<td></td>
</tr>
<tr>
<td>Tris</td>
<td>Tris(hydroxymethyl)-aminomethane</td>
<td></td>
</tr>
<tr>
<td>Tyr</td>
<td>tyrosine</td>
<td></td>
</tr>
<tr>
<td>Ubi-1</td>
<td>maize polyubiquitin-1 promoter</td>
<td></td>
</tr>
<tr>
<td>UDPGlc</td>
<td>uridine 5'-diphosphate glucose</td>
<td></td>
</tr>
<tr>
<td>uidA</td>
<td>β-glucuronidase gene</td>
<td></td>
</tr>
<tr>
<td>UV</td>
<td>ultraviolet</td>
<td></td>
</tr>
<tr>
<td>vir</td>
<td>virulence gene</td>
<td></td>
</tr>
<tr>
<td>X-Gal</td>
<td>5-bromo-4-chloro-3-indolyl-β-D-galactoside</td>
<td></td>
</tr>
<tr>
<td>X-Gluc</td>
<td>5-bromo-4-chloro-3-indolyl-β-glucuronic acid</td>
<td></td>
</tr>
</tbody>
</table>
Spectacular progress has been made in the genetic manipulation of plants over the past two decades. Molecular techniques have now been applied to over 120 species in at least 35 families, including most major economical crops, vegetables, ornamental and medicinal plants (Birch, 1997). The refinement of techniques to introduce and express diverse foreign genes in plants and the development of sophisticated techniques to recover intact plants after transformation are the two major factors responsible for this progress. Transformation of sugarcane is also well established (Bower and Birch, 1992), and the genetic manipulation of this crop is now possible.

In South Africa the sugarcane breeders are faced with two major concerns, the first being the increase of sucrose yield and the second being resistance to the numerous pathogens and pests that cause major crop losses. Sucrose is the third most important agricultural export product in South Africa and a valuable source of revenue with a net income of approximately 3 600 million rand per season. An average annual harvest of 20 million tons of cane reaches the mills, yielding 2.3 million tons of sucrose. Improvement of sucrose yield is thus an important objective of the South African sugar industry for it to maintain international competitiveness. Over the past two decades, however, there has been little improvement in sucrose yield through conventional plant breeding, despite increased efforts in this regard. Although this is partly due to the difficulty of selecting for the desired phenotype in field trials, it may also be an indication that the genetic resources available in the parental germplasm have been substantially exploited already. If this is the case, genetic engineering could play a major role in the future.

The biotechnology programme at the South African Sugar Association Experiment Station (SASEX, Mount Edgecombe, Kwazulu Natal) is aimed at the improvement of the varietal potential of sugarcane, and is closely linked to the plant breeding
programme. One of the main focal points of the sugarcane biotechnology programme at SASEX is genetic engineering. The present study forms part of the transgenic plant programme.

To achieve effective transfer of foreign DNA into sugarcane cells, appropriate genetic constructs need to be made to facilitate integration and expression of the transgene. A typical genetic construct will contain a promoter, a transgene and a terminating signal. Several common promoters and genes that are active in different plants are currently available. There is, however, a commercial limitation to the use of many of these elements. Because the development of a promoter or resistance gene is effectively an invention, it is eligible for patent protection. Patents have already been issued on most established or promising genes and promoters, which has necessitated a renewed effort to isolate novel regulatory elements from plants. Another phenomenon that has been the cause of major concern in transgenic plant programmes is transgene silencing (Finnegan and McElroy, 1994). The mechanism of silencing is not clear, but it can occur at either transcriptional or post-transcriptional level, and from one generation to another, or somatically within a single plant (Jorgensen, 1995). The genetic material is not lost from the genome, but is simply no longer expressed. The use of stable promoters is thus a prerequisite for the successful production of transgenic plants.

The use of molecular technology to increase product yield through the manipulation of key enzymes requires strict control of transgene expression in certain tissue types or at specific developmental stages. This is especially true when attempts are being made to alter metabolism. In sugarcane, sucrose metabolism should be altered in the culm, preferably in the storage parenchyma, and not in the leaves or other tissues. Changes in sucrose metabolism in the whole plant will most probably have a negative affect on growth and yield.

The shortage of reliable regulatory elements to drive expression of transgenes in specific areas of the plant as well as patent considerations have made it necessary
to isolate specific promoters from sugarcane, and to investigate the performance of other possibly useful elements in sugarcane cells. Since sucrose is accumulated in the culm and some pathogens gain entry to the plant through the buds, the culm has become one of the major target areas for transgene expression in sugarcane and is thus the focus area from which possible promoters are isolated.

The aims of the present study was firstly to isolate promoter(s) that will drive stable transgene expression in the sugarcane culm, and secondly, to evaluate the use of the bacterial promoter, rolC, for tissue-specific expression in sugarcane.

Two distinct groups of techniques are available for the isolation of differentially expressed genes and these include non-directed and directed techniques. An overview of these techniques are discussed in Chapter 2. Both of these approaches have been followed in the present study.

The differential display technique is a non-directed technique that was developed to identify and characterise differentially expressed genes by detecting individual mRNA species, and to subsequently isolate and analyse the cDNA (Liang and Pardee, 1992). Polymerase chain reaction (PCR) is used to amplify cDNA fragments of different tissues after which amplification profiles are compared. The technique has been used to isolate differentially expressed genes in a number of experimental systems and a modified version of the technique was evaluated for the identification of culm-specific genes in sugarcane. Differential display analysis of four different sugarcane tissues, and refinements of the technique for use in sugarcane are discussed in Chapter 3.

Regulatory elements of certain genes in Agrobacterium tumefaciens and A.rhizogenes are not expressed in Agrobacterium, but become active on transfer to plant cells. One of these promoters, rolC, has been shown to drive phloem-specific gene expression in transgenic tobacco (Schmülling et al., 1989), rice (Matsuki et al., 1989) and hybrid aspen (Nilsson et al., 1996). The rolC promoter was also found to
be activated by sucrose in phloem cells of transgenic tobacco plants (Yokoyama et al., 1994). Phloem-specific transgene expression in sugarcane will be advantageous in cases where resistance to phloem-feeders such as aphids and leafhoppers that carry viruses is conferred or where viral resistance genes are introduced into the plant. A sucrose-activated promoter is also of great benefit for studying sucrose metabolism in the sugarcane plant. In addition, phloem-specific gene expression may impose a decreased metabolic load on the plant in comparison to constitutive expression, and this may be crucial for the stable expression of transgenes.

The tissue-specificity of the Agrobacterium rolC promoter was investigated for the first time in sugarcane culm tissue. Transformation of sugarcane using two different reporter gene constructs, and the observed rolC expression patterns in mature sugarcane culm tissue are discussed in Chapter 4.

Another approach for the isolation of promoter elements is based on prior knowledge of the expression pattern of a particular gene or the sequence of the protein. The expression patterns of key enzymes involved in carbohydrate metabolism in sugarcane have been studied (Botha et al., 1996). Although the expression (total activity per internode) of several of the enzymes increases with culm maturity one of these, the pyrophosphate-dependent phosphofructokinase (PFP), is of particular interest. PFP activity is negatively correlated with sucrose content in all analysed sugarcane varieties, and is apparently preferentially expressed in the peripheral areas of the culm (1996, unpublished1). These properties suggest that the PFP promoter might be an ideal candidate for the enrichment of transgene expression in the mature culm. The PFP enzyme is a heterotetramer with two α and two β-subunits, and the β-subunit has been shown to be the catalytic subunit (Yan and Tao, 1984; Carlisle et al., 1990).

The gene sequence of PFP-β has been determined in castor bean and potato, and

---

1F.C.Botha, Institute for Plant Biotechnology, Private Bag X1, Matieland, 7602, South Africa
partially in rice. This information was used in a directed approach to isolate the gene and subsequently the promoter from sugarcane. Since this is the first reported investigation of the activity of the PFP-β promoter, two different constructs were used to evaluate its expression in sugarcane cells. Activity analyses of these two constructs containing the PFP-β promoter in sugarcane are discussed in Chapter 5.
CHAPTER 2
LITERATURE REVIEW

2.1 INTRODUCTION
The present study was aimed at the identification and isolation of potential promoter elements for sugarcane transformation. Comprehensive reviews on plant promoters (Guilfoyle, 1997) and their use in transformation studies have been written in the past five years (An and Kim, 1993; Christou, 1996; Birch, 1997). This chapter is therefore not an attempt to cover all aspects of plant promoters, but is rather an overview of those issues that are relevant to the current work. For more detailed information the reader is referred to these review papers.

As is evident from the available literature, two distinct approaches can be used to identify and isolate new promoters. The first is independent of prior knowledge of a specific gene and is directed towards a group rather than an individual gene. The second is aimed at a specific gene and requires some information of the gene or gene product (protein). Both these approaches have been used in the current investigation.

However, due to the very limited information on sugarcane gene sequences and gene expression patterns, options in the second approach are limited. In previous work in the South African sugar industry, evidence was found that the expression of the pyrophosphate-dependent phosphofructokinase (PFP) is high in culm tissue and apparently increases in more mature tissues. This makes the PFP promoter an interesting candidate for culm directed gene expression. Some background information on PFP is therefore included in this chapter.

2.2 PLANT PROMOTERS
The regulation of gene expression in plants is controlled at both the transcriptional and post-transcriptional level. Multiple levels of control exist, including splicing, 3' end formation, stability, and translation of the mRNA as well as post-translational events
such as protein stability and modification (Sullivan and Green, 1993).

Transcription takes place in three stages, namely initiation, elongation and termination. Initiation begins with the binding of RNA polymerase to the promoter region of each gene, typically 40-200 nucleotides upstream of the translation initiation codon, ATG.

Plant genes can be divided into three classes which are defined by the type of promoter. Each class of gene promoter is recognised by a different polymerase enzyme: ribosomal RNA (rRNA) is transcribed by RNA polymerase I, messenger RNA (mRNA) by RNA polymerase II, and transfer RNA (tRNA) and other small RNAs by RNA polymerase III (An and Kim, 1993). In addition to RNA polymerases, trans-acting proteins or transcription factors are needed for transcription initiation. Elongation occurs by movement of the RNA polymerase along the DNA, extending the RNA chain until RNA synthesis is terminated at the transcription terminator region where the transcription complex dissociates (An and Kim, 1993).

Eukaryotic promoters are composed of many cis-acting control motifs that are recognised by specific trans-acting factors. The main characteristics of promoter motifs are distance-dependent activation and the requirement for location 5' to the cap site (Thangue and Rigby, 1988), an unusual 5'→5' triphosphate linkage involving a 7-methylguanine residue which is added after transcription (Grierson and Covey, 1984).

Most promoters contain a TATA box (TATA(T/A)A(T/A)), an AT-rich sequence located 21 to 35 nucleotides upstream from the transcription start site (Breathnach and Chambon, 1981). This region is surrounded by GC-rich sequences which may play a role in its function (Lefebvre and Gellatly, 1997) and interacts with a TATA-binding protein which comprises the TFIID transcription factor complex necessary for accurate initiation of promoters transcribed by RNA polymerase II (Rosenthal, 1987; Mukumoto et al., 1993; Czamecka-Verner et al., 1994). Deletion of the TATA box
reduces promoter activity and results in multiple transcription start sites (An and Kim, 1993). In some promoters, such as those for the maize zein gene, two TATA boxes may be found as close as 10 nucleotides apart (Lefebvre and Gellatly, 1997).

Another control motif, the CCAAT box sequence, is located close to the -80 region of several, but not all plant promoters and is functional in both orientations (Thangue and Rigby, 1988). It has been established that a family of transcription factors recognise this motif and that factors binding to this sequence can co-operate with those binding to surrounding motifs (Thangue and Rigby, 1988).

A hexamer sequence, TGACGT, occurs in most constitutive promoters within a few hundred nucleotides from the transcription start site (An and Kim, 1993). These hexamer motifs are often found as repeats separated by six to eight nucleotides, and deletion analysis has indicated that these motifs are essential for the transcription activity of the cauliflower mosaic virus (CaMV) 35S, octapine synthase (ocs) and nopaline synthase (nos) promoters (An and Kim, 1993). Genes for the transcription factor that specifically interacts with the hexamer motifs have been isolated from both dicotyledonous and monocotyledonous plant species (Katagiri et al., 1989; Singh et al., 1990).

Motifs that are found in promoters of genes that are environmentally inducible include the G-box sequence (CCACGTGG) in various photosynthesis gene promoters (An and Kim, 1993) and the H-box sequence (CCTACC) which, together with the G-Box sequence, is essential for the expression of the bean chalcone synthase gene (Arias et al., 1993). AT-rich motifs appear to increase the amplitude of induction in heat-shock protein genes (Czarnecka-Verner et al., 1994) and may play a regulatory role in several other higher plant genes (Forde, 1994). Examples of genes that contain AT-rich motifs are nodulin genes, seed-specific genes like lectin, trypsin inhibitor and β-phaseolin, photoregulated genes and some heat shock genes (Forde, 1994).

Another group of cis-acting regulatory sequences are called enhancers. These share
common characteristics with promoters and they are involved in increasing the level of transcription. Enhancers are typically 100-200 base pairs in length and can function from a distance, sometimes several thousand base pairs, independently of orientation and either from the 5' or 3' end of a gene (Lefebvre and Gellatly, 1997). They can also activate transcription from heterologous promoters with exceptional positional flexibility (Hatzopoulos et al., 1988). These characteristics distinguish enhancers from other regulatory elements. In some cases, however, tandem copies of some promoter motifs can acquire enhancer properties (Thangue and Rigby, 1988). Enhancers from different eukaryotic organisms are very similar, for example many ribulose-1,5-bisphosphate carboxylase-oxygenase (rbcS) genes from dicots contain a sequence closely related to an animal enhancer, GTGTGG(T/C)(T/C)A(A/T)TA(T/A)G, at -140 from the transcriptional start site (Kuhlemeier et al., 1987). Enhancer sequences can also function as gene silencers (Lefebvre and Gellatly, 1997).

Promoters are classified in three groups, namely constitutive, inducible and tissue-specifically or developmentally regulated. Plant organelles also contain their own class of promoters and some of these play an essential role in the control of gene expression during photosynthesis and CO$_2$ fixation. The promoters that are chosen for transformation studies will vary according to the expression pattern that is preferred (Gruber and Crosby, 1993).

### 2.2.1 Constitutive promoters

Constitutive promoters are active during most stages of development and in most plant tissues. A gene is classified as constitutive when its expression remains constant on a per cell basis under all conditions examined (Kuhlemeier et al., 1987). This implies that there must be little variation of mRNA abundance or translational efficiency during development, in different organs or upon application of various endogenous or environmental stimuli (Kuhlemeier et al., 1987). Examples of constitutive promoters are the nuclear gene promoter for the $\beta$-subunit of the mitochondrial ATPase complex (Boutry and Chua, 1985), and the rice actin promoter...
2.2.2 Inducible promoters

There is a wide variety of inducible promoters, a few of which include temperature-inducible, hormone-inducible, pathogen-inducible and light-inducible promoters (Gruber and Crosby, 1993). Important cis-regulatory regions with consensus sequences identified in inducible promoters are summarised in Table 2.1.

2.2.2.1 Temperature-inducible promoters: Heat-shock promoters

The heat stress response in plants and other organisms comprises a highly complex but transient reprogramming of cellular activities necessary to protect the cells from extensive damage and to provide optimum conditions for recovery after the stress period (Scharf et al., 1994). Heat shock proteins are sorted into four major classes on the basis of their sizes. They are encoded by nuclear genes but are localised in different cell compartments, including the cytoplasm, chloroplast, mitochondria and endoplasmic reticulum (Czarnecka-Verner et al., 1994).

The thermo-inducible expression of heat shock proteins is attributed to the presence of controlling DNA sequences located in their promoters (for reviews see Czarnecka-Verner et al., 1994; Scharf et al., 1994). These conserved sequences were named heat shock elements, or HSEs (Scharf et al., 1994; Czarnecka-Verner et al., 1994). Single, overlapping or multiple copies of the HSEs can occur in the promoter, usually within 15 to 18 bp upstream from the TATA box (Czarnecka-Verner et al., 1994). The concise analysis of HSEs has resulted in a general structure which represents multimers of five nucleotide modules with alternating, palindromic orientation: 5'- (aGAAn)(nTTCt)(aGAAn)(nTTCt)-3' (Scharf et al., 1994). Three consecutive modules are evidently the minimum combination required for an active heat shock promoter. In appropriate configurations, the HSE confers both high selectivity of heat shock induction and high overall efficiency of transcription initiation (Scharf et al., 1994).

Most heat shock promoters contain TATA and CCAAT boxes. The requirement for
Table 2.1 Cis-regulatory regions identified in plant inducible promoters.

<table>
<thead>
<tr>
<th>GENE</th>
<th>INDUCER</th>
<th>REGULATORY REGION</th>
<th>ELEMENT/MOTIF</th>
<th>CONSENSUS SEQUENCE (5'-3')</th>
<th>REFERENCE</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Temperature- and water stress-inducible promoters</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Heat shock proteins (several)</td>
<td>High temperatures</td>
<td>1) (15-18 bp from TATA-box</td>
<td>TATA-box</td>
<td>1) (wGAAn)nTTC(T/G)2</td>
<td>Scharf et al., 1994</td>
</tr>
<tr>
<td>Potato α27</td>
<td>Low temperatures</td>
<td>-2000 to +1</td>
<td>CCAAT-box</td>
<td>2) (TATA(T/A)(T/A)</td>
<td>Lefebvre and Gellatly, 1997</td>
</tr>
<tr>
<td>Arabidopsis rd29a</td>
<td>Dehydration due to drought, low temperatures or salt stress</td>
<td>-100 region</td>
<td>G-box</td>
<td>3) (GACCG)GACG(T/G)2</td>
<td>Czamecka et al., 1989</td>
</tr>
<tr>
<td>Maize rab28</td>
<td>Water stress</td>
<td>-2000 to +1</td>
<td>ABA</td>
<td></td>
<td>Kirch et al., 1997</td>
</tr>
<tr>
<td><strong>Hormone-inducible promoters</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wheat embryo protein, Em</td>
<td>ABA or signals that cause changes in ABA levels</td>
<td>-94 to -85</td>
<td>ABRE</td>
<td>TTGCCGACACGTTGCCGCGCA</td>
<td>Quatrano et al., 1993</td>
</tr>
<tr>
<td>Barley HVA22</td>
<td>ABA</td>
<td>2) (64 to 56</td>
<td>ABRE3</td>
<td>1) GCCCACTACA</td>
<td>Shen &amp; Ho, 1995</td>
</tr>
<tr>
<td>Rab17</td>
<td>ABA</td>
<td>1) (119)</td>
<td>CE1</td>
<td>2) GCCCACTACA</td>
<td>Vierdell et al., 1990</td>
</tr>
<tr>
<td>Cae1727-45</td>
<td>ABA</td>
<td>2) (148)</td>
<td>CE1</td>
<td>1) GCCCACTACA</td>
<td>Michel et al., 1993</td>
</tr>
<tr>
<td>HVA1</td>
<td>ABA</td>
<td>1) (78)</td>
<td>CE1</td>
<td>2) GCCCACTACA</td>
<td>Straub et al., 1994</td>
</tr>
<tr>
<td>Lycopersicon esculentum LE25</td>
<td>ABA</td>
<td>1) (83)</td>
<td>CE1</td>
<td>1) GCCCACTACA</td>
<td>Cohen and Bray, 1992</td>
</tr>
<tr>
<td>Soybean GH3 &amp; SAUR</td>
<td>Auxin</td>
<td>-76 (GH3)</td>
<td>GCC-box (k2)</td>
<td>1) GCCCACTACA</td>
<td>Hagen, 1994</td>
</tr>
<tr>
<td>Tobacco chinase chin48</td>
<td>Ethylene</td>
<td>-503 to -358</td>
<td>ERE</td>
<td>1) TGGCGCGCCGCGCCGCGT</td>
<td>Shinshi et al., 1995</td>
</tr>
<tr>
<td>Tobacco prb-1b</td>
<td>Ethylene</td>
<td>-187</td>
<td>ERE</td>
<td>2) GCCCACTACA</td>
<td>Sonnensch, 1994</td>
</tr>
<tr>
<td>Bean chinase CH5B</td>
<td>Ethylene</td>
<td>-2130 to -355</td>
<td>ERE</td>
<td>1) GCCCACTACA</td>
<td>Roby et al., 1991</td>
</tr>
<tr>
<td>Lycopersicon esculentum L LeMTB</td>
<td>Ethylene</td>
<td>-76 (GH3)</td>
<td>GCC-box (k2)</td>
<td>1) GCCCACTACA</td>
<td>Whitelaw et al., 1997</td>
</tr>
<tr>
<td>PRRBfb</td>
<td>Ethylene</td>
<td>-213 to -141</td>
<td>ERE</td>
<td>1) TGGCGCGCCGCGCCGCGT</td>
<td>Sessa et al., 1995</td>
</tr>
<tr>
<td>Barley high pi α-amy lase</td>
<td>GA</td>
<td>-203</td>
<td>GARE</td>
<td>2) GCCCACTACA</td>
<td>Gubler &amp; Jacobsen, 1992</td>
</tr>
<tr>
<td>Cereal α-amy lase Amy23b</td>
<td>GA &amp; ABA</td>
<td>-2400 to +1</td>
<td>GARE</td>
<td>1) GCCCACTACA</td>
<td>Lanahan et al., 1992</td>
</tr>
<tr>
<td>Tomato lipoxygenases tomox A &amp; tomox B</td>
<td>Methyl jasmonate (in seedlings)</td>
<td>-2400 to +1</td>
<td>GARE</td>
<td>2) GCCCACTACA</td>
<td>Beauduin &amp; Rothstein, 1997</td>
</tr>
<tr>
<td>Pathogen-inducible promoters and promoters induced in defense responses</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---------------------------------------------------------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Tobacco class I B-1,3-glucanase</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Tobacco chitinase CHN50</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>PR-1a</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>PR-5 thaumatin-like protein</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Parsley PR2</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Parsley CCu5OMT</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Chitinase and glucanase</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Proteinase inhibitor IIK</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Asparagus AoPR1</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Soybean CHS genes</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| **Pathogen attack**                                         |
| **Elitor**                                                  |
| **SA and TMV infection**                                   |
| **Fungal elicitor**                                         |
| **Stress elicitor**                                         |
| **Bacteria, fungi, fungal elicitor, ethylene, SA, wounding and cytokinin** |

| **Wounding, elicitor, light**                               |
| **2 MU**                                                    |
| **G-box**                                                   |

| **Light-inducible promoters**                               |
| **Soybean CHS genes**                                       |
| **Peanr3c5-3A**                                             |
| **Lemma rbc5 SSUSB**                                        |
| **Arabidopsis FAD7**                                        |
| **Lemma gibba cab4S19**                                     |
| **Wheat cab1**                                              |
| **Peanr3c5880**                                             |
| **Arabidopsis cab140**                                      |
| **Maize cab1**                                              |

| **UV-light**                                                |
| **25 UV-light**                                             |
| **White/red light**                                         |
| **Light (tissue-specific)**                                 |
| **Phytochrome**                                             |

| **1)White light**                                           |
| **2)White-light**                                           |
| **3)White light**                                           |
| **1)White light**                                           |
| **2)White-light**                                           |
| **3)White-light**                                           |
| **Light** (tissue-specific)                                 |

| **-1452 to -1193**                                          |
| **1)-596 to -437**                                         |
| **2)-47 to -68**                                           |
| **3)-539 to -518**                                         |
| **-902 to -425**                                           |
| **-718 to -1364**                                          |
| **-52 to -168**                                            |
| **-139**                                                    |
| **-136 to -567**                                           |
| **-574**                                                    |
| **-175 to -134**                                           |
| **1)-1511 to -138**                                        |
| **2)-211 to -201**                                         |
| **3)-194 to -176**                                         |
| **-156 to -141**                                           |
| **2)-85 to -60**                                           |
| **3)-982**                                                  |
| **-154 to -138**                                           |
| **1)-1561 to -138**                                        |
| **2)-211 to -201**                                         |
| **3)-194 to -176**                                         |
| **-156 to -141**                                           |
| **2)-85 to -60**                                           |
| **1)GT-box I**                                             |
| **1)GT-box I**                                             |
| **2)GT-box I**                                             |
| **3)GA-1**                                                 |

| **Silencer**                                                |
| **AGC-box**                                                 |

| **3)GTCAAGAAAGTCAG**                                        |
| **3)GTCAAGAAAGTCAG**                                        |
| **1)AACGGTAAGGT**                                          |
| **2)AACCTGGCC**                                            |
| **CACGTGG**                                                |
| **Similar to ACCTGACC**                                    |

| **1)TACCTACCAA**                                           |
| **1)TACCTACCAA**                                           |
| **2)ACCTTGAGCCT**                                          |
| **3)CTGTAAGGG**                                            |
| **1)GATAAGGATGTAATCC**                                    |
| **2)GATCGAGCGATGGTAGCCCTAGAA**                             |
| **3)GATAAGGATGTAATCC**                                    |
| **2)GATCGAGCGATGGTAGCCCTAGAA**                             |
| **1)GTTATA**                                               |
| **2)GTTATA**                                               |
| **1)AAGCAATCCCAAGCC**                                      |
| **2)AACGCCAAGCCCAAGCC**                                    |
| **3)AACGCCAAGCCCAAGCC**                                    |
| **1)AAGCGTAAGT**                                           |
| **2)ACCTTGCC**                                             |
| **CACGTG**                                                 |
| **Similar to ACCTGACC**                                    |

| **TACCTACCAA**                                             |
| **1)TACCTACCAA**                                           |
| **2)ACCTTGAGCCT**                                          |
| **3)CTGTAAGGG**                                            |
| **1)GATAAGGATGTAATCC**                                    |
| **2)GATCGAGCGATGGTAGCCCTAGAA**                             |
| **3)GATAAGGATGTAATCC**                                    |
| **2)GATCGAGCGATGGTAGCCCTAGAA**                             |
| **1)GTTATA**                                               |
| **2)GTTATA**                                               |
| **1)AAGCAATCCCAAGCC**                                      |
| **2)AACGCCAAGCCCAAGCC**                                    |
| **3)AACGCCAAGCCCAAGCC**                                    |
| **1)AAGCGTAAGT**                                           |
| **2)ACCTTGCC**                                             |
| **CACGTG**                                                 |
| **Similar to ACCTGACC**                                    |

| **Vogel-Lange et al., 1994**                                |
| **Fukuda & Shinshi 1994**                                   |
| **Van de Rhee et al., 1990**                                |
| **Voskuil et al., 1990**                                    |
| **Gillman & Maren, 1997**                                   |
| **Sonenshie, 1994**                                         |
| **Palin et al., 1990**                                      |
| **Kim et al., 1992**                                        |
| **Wanner et al., 1993**                                     |
| **Wingender et al., 1990**                                  |

| **Soybean CHS genes**                                       |
| **Peanr3c5-3A**                                             |
| **Lemma rbc5 SSUSB**                                        |
| **Arabidopsis FAD7**                                        |
| **Lemma gibba cab4S19**                                     |
| **Wheat cab1**                                              |
| **Peanr3c5880**                                             |
| **Arabidopsis cab140**                                      |
| **Maize cab1**                                              |

| **-42**                                                     |
| **-151 to -138**                                            |
| **-211 to -201**                                            |
| **-194 to -176**                                           |
| **-156 to -141**                                           |
| **-85 to -60**                                             |
| **-982**                                                    |
| **-154 to -138**                                           |
| **-1561 to -138**                                          |
| **-211 to -201**                                           |
| **-194 to -176**                                           |
| **-156 to -141**                                           |
| **-85 to -60**                                             |
| **-134 to -106**                                           |
| **-142 to -114**                                           |
| **-82 to -54**                                             |
| **-89 to -61**                                             |

| **GT-box II**                                               |
| **2)GT-box II**                                             |
| **3)GA-1**                                                 |
| **1)KTAAGGATGTAATCC**                                       |
| **2)GATCGAGCGATGGTAGCCCTAGAA**                              |
| **1)KTAAGGATGTAATCC**                                       |
| **2)GATCGAGCGATGGTAGCCCTAGAA**                              |
| **1)GTATA**                                                |
| **2)GTATA**                                                |
| **1)GTTATA**                                               |
| **2)GTTATA**                                               |
| **1)AAGCAATCCCAAGCC**                                      |
| **2)AACGCCAAGCCCAAGCC**                                    |
| **3)AACGCCAAGCCCAAGCC**                                    |
| **1)AAGCGTAAGT**                                           |
| **2)ACCTTGCC**                                             |
| **CACGTG**                                                 |

| **1)AGC-box**                                              |

| **3)GTCAAGAAAGTCAG**                                        |
| **1)AACGGTAAGGT**                                          |
| **2)AACCTGGCC**                                            |
| **CACGTGG**                                                |
| **Similar to ACCTGACC**                                    |

| **1)TACCTACCAA**                                           |
| **1)TACCTACCAA**                                           |
| **2)ACCTTGAGCCT**                                          |
| **3)CTGTAAGGG**                                            |
| **1)GATAAGGATGTAATCC**                                    |
| **2)GATCGAGCGATGGTAGCCCTAGAA**                             |
| **3)GATAAGGATGTAATCC**                                    |
| **2)GATCGAGCGATGGTAGCCCTAGAA**                             |
| **1)GTTATA**                                               |
| **2)GTTATA**                                               |
| **1)AAGCAATCCCAAGCC**                                      |
| **2)AACGCCAAGCCCAAGCC**                                    |
| **3)AACGCCAAGCCCAAGCC**                                    |
these motifs together with the HSEs has been demonstrated by deletion analysis (Czarnecka et al., 1989). Promoters of several soybean genes encoding heat shock proteins contain numerous blocks of contiguous AT-rich sequences located upstream from the TATA box. They vary in length and are not fixed to a specific position in the 5' flanking region. They do not show strict sequence conservation but frequently exist as repeats of TAAT or ATTA motifs (Czarnecka-Verner et al., 1994).

Transcriptional activation of heat shock genes in response to elevated temperatures or chemical levels is mediated by the interaction of the trans-acting heat shock transcription factors (HSFs) with HSEs of the promoter. The genes for the HSFs have been cloned and characterised (Scharf et al., 1994).

2.2.2.2 Temperature-inducible promoters: Cold-induced promoters
Plants from tropical and subtropical regions show injury symptoms when exposed to non-freezing low temperatures below 10 - 15°C. The extent of chilling injury depends on the duration and severity of treatment (Yoshida and Kato, 1994). Plant response to low temperature leads to alteration of the cellular metabolism and is correlated with significant changes in gene expression (for reviews see Guy, 1990; Palva, 1994). Cold-regulated genes have been isolated and characterised in several plant species including alfalfa (Wolfraim et al., 1993), Arabidopsis (Hajela et al., 1990), spinach (Neven et al., 1993), tomato (Schaffner and Fischer, 1990), barley (Cattivelli and Bartels, 1990) and wheat (Houde et al., 1992). In the Arabidopsis rd29a promoter a dehydration-responsive element (DRE) was identified (TACCGACAT). Within 1 kb of the transcription start site another two conserved sequence elements were identified and six G-box motifs (CACGTG) were found to be spread over the whole promoter (Kirch et al., 1997).

2.2.2.3 Hormone-inducible promoters
Plant hormones are considered to be an important class of internal signal compounds which are easily transported to target cells and have been shown to be directly involved in transcriptional and translational control of gene activity in a number of
plant processes. Specific regulatory elements have been identified for hormone-responsiveness and several review papers have been written (Kuhlemeier et al., 1987; Quatrano et al., 1993; Giraudat et al., 1994; Barendse and Peeters, 1995; Gatz, 1997). A few of the important examples are discussed below.

The growth regulator abscisic acid (ABA) is known to induce the synthesis of specific proteins, among which are certain seed-storage proteins (Kuhlemeier et al., 1987). Examples of genes that are expressed due to ABA influence include the beta conglycinin α-subunit, lectins and agglutinins, and these may be involved in maintaining seed dormancy, in protecting seeds against dessication, and in preventing deleterious effects of hydrolytic enzymes (Kuhlemeier et al., 1987).

A protein that has been found to be ABA-inducible is the 10 kD soluble protein, Em, which accumulates in wheat embryos in the latter third of grain development (Williamson and Quatrano, 1988). When the embryos were removed from the grain in the first third of grain development, the Em protein did not accumulate and the embryos germinated precociously. In the presence of ABA, however, Em did accumulate in these embryos and normal germination occurred (Williamson and Quatrano, 1988). The Em gene is not expressed in response to auxin, cytokinin, salicylic acid (SA) or gibberellic acid (GA). However, environmental signals that are believed to be transduced, at least in part, by changes in ABA levels, e.g. desiccation, temperature extremes, salinity and wounding, can result in Em gene expression (Quatrano et al., 1993).

The Em promoter was functionally dissected and it was found that the 5' distal region consists, at least in part, of an AT-rich region that appears to be responsible for quantitative levels of expression through effects on transcription. The next region of regulatory sequences is the ABA response element (ABRE) which is responsible for the dramatic increase of expression in the presence of ABA. The consensus sequence of the ABRE is 20 bp in length with the sequence 5'-TTGCGGCACACGTGGGCGCGA-3'. A third component of the Em gene 5' flanking
region, the 5' untranslated region (5'UTR), also appears to be involved in the ultimate level of expression from the Em promoter at a post-transcriptional level (Quatrano et al., 1993).

The plant hormone gibberellic acid (GA) is produced upon the onset of germination and induces the synthesis of a number of enzymes involved in the breakdown of seed storage products, for example alpha-amylase (Kuhlemeier et al., 1987). Deletion analysis of the promoter of a high-pl α-amylase gene of barley has shown that the major GA- and ABA-responsive elements are located between -174 and +53. The results showed that both the TAACAAA- and the TATCCAC-boxes play an important role in GA-regulated expression (Gubler and Jacobsen, 1992).

Auxins and cytokinins also influence expression of genes including those that are involved in maintenance of apical dominance, cell elongation and xylem differentiation (Kuhlemeier et al., 1987). The promoter regions of two soybean auxin responsive genes, GH3 and SAUR, have been analysed (Hagen et al., 1994). The auxin-responsive elements were defined to a 76 bp region in the GH3 promoter and to a 30 bp region in the SAUR promoter (Hagen et al., 1994).

The plant responds to pathogen infection or other stress by producing the stress hormone ethylene and this is the signal for plants to activate the defense mechanism against pathogens. Inducible defense responses are mediated, at least in part, by ethylene-induced gene expression (Shinshi et al., 1995). The promoter of an ethylene-induced tobacco class I basic chitinase gene, Chn48, has been examined in order to identify ethylene-responsive elements (Shinshi et al., 1995). In transgenic tobacco plants a chimeric construct containing a 2 kb Chn48 promoter fused to the β-glucuronidase (GUS) reporter gene was induced by ethylene in leaf tissues. Deletion analysis indicated that an ethylene-responsive element was located between -503 and -358 relative to the transcription initiation site. This 146 bp sequence conferred ethylene-responsive GUS expression in either orientation upstream of the heterologous promoter, indicating that the sequence functions as a regulatory
enhancer. The ethylene-responsive region contains two copies of the GCC-box (TAAGAGCCG), which is conserved in ethylene-responsive defence genes (Shinshi et al., 1995).

2.2.2.4 Pathogen-inducible promoters

Plants are constantly exposed to environmental stress that could be detrimental to their survival. Apart from fluctuations in temperature, changes in nutrient composition and salinity of the soil, flooding, drought, etc., plants are continuously exposed to airborne and soilborne bacterial and fungal pathogens, viruses or viroids and insect pests (Somssich, 1994; Van de Rhee et al., 1994).

In order to defend themselves against these biotic stresses, plants possess preformed structural physical barriers, such as the cell wall, and have also developed inducible mechanisms which allow them to respond rapidly to external stimuli. This hypersensitive response causes the cells in the immediate vicinity of the infection site to die, making it impossible for the pathogen to spread through the plant. In nearly all of the systems studied to date, transcriptional activation of specific plant genes plays a key role in the defence of plants against pathogens (Dixon and Lamb, 1990). The proteins induced in this way are all involved in defending the plant against attack or preventing the spread of infection, and include enzymes of the phenylpropanoid pathway, enzymes involved in phytoalexin production, structural cell wall proteins, peroxidases, superoxide dismutases, proteinase inhibitors, thionins and the "pathogenesis-related" (PR) proteins (Van de Rhee et al., 1994). PR proteins have been defined as proteins coded for by the host plant but induced by various pathogens as well as under stress situations similar to those provoked by pathogens (Van Loon, 1990). These proteins may also play a role in enhancing the levels of resistance against secondary challenges by pathogens since they not only accumulate locally around the sites of infection, but also appear in the uninfected parts of the plant (Van de Rhee et al., 1994). This implies the phenomenon of induced resistance or systemic acquired resistance (SAR) (Ross, 1966). Other stress conditions that induce specific subsets of PR proteins include wounding and
ultraviolet (UV) light, chemical substances such as salicylic acid, the plant hormone ethylene and fungal elicitors (Van de Rhee et al., 1994).

PR proteins have been divided into groups according to sequence similarities and immunological cross-reactivities. Five PR protein groups have been defined from tobacco, namely PR-1 to PR-5. Each group has been shown to comprise both extracellular, acidic proteins and intracellular, basic counterparts (Van de Rhee et al., 1993). Proteins from PR-2 and PR-3 have been identified as being β-1,3-glucanases and chitinases, respectively, while PR-5 proteins were identified as being thaumatin-like proteins (Somssich, 1994). Glucans and chitins are major constituents of fungal cell walls. Other PR protein groups have been identified in potatoes, tomatoes, parsley, bean and peas (Somssich, 1994).

The regulatory elements involved in PR gene expression have been studied by fusing large fragments of the respective promoters to appropriate heterologous reporter genes and analysing the expression patterns in transgenic plants. Review papers on regulatory elements governing PR gene expression are available (Somssich, 1994; Van de Rhee et al., 1994) and some examples are discussed below.

i) PR-1 Genes
At least 689 bp of 5' flanking sequences of the pr-1a promoter are needed for significant induction of GUS activity by salicylic acid (SA) treatment and infection with tobacco mosaic virus (TMV) in transgenic tobacco. An upstream fragment of nucleotides -902 to -625 from the pr-1a promoter was shown to contain salicylic acid and TMV-responsive sequences (Van de Rhee et al., 1993). Deletion analysis of the tobacco basic PR-1 protein gene promoter, prb-1b, identified sequences between -141 and -213 to be required for ethylene inducibility. This region contains a G-box consensus core sequence, CACGTG, at position -187 (Somssich, 1994).

ii) PR-2 and PR-3 Genes
Gene expression of β-1,3-glucanase (PR-2) and chitinase (PR-3) appears to be
coordinately regulated in several plant species upon induction by bacteria, fungal elicitors, ethylene, SA, wounding, cytokinin and during development (Meins et al., 1992). One common motif, AGCCGCC (AGC box), has been found in nearly all chitinase and glucanase promoters analysed so far (Somssich, 1994). The developmental and stress-induced expression of a GUS construct containing 2 kb of a β-1,3-glucanase promoter from Nicotiana plumbaginifolia in transgenic tobacco has been studied (Castresana et al., 1990). In healthy plants, GUS expression was mainly detected in the roots and lower leaves and was tissue-specific. GUS activity was strongly induced by SA, and bacterial infection leading to the hypersensitive response of the transgenic plants resulted in increased GUS activity around the necrotic lesions (Castresana et al., 1990).

Functional analysis of the bean chitinase ch5b promoter in transgenic tobacco plants revealed a region between -195 and -422 which is important for ethylene responsiveness (Broglie et al., 1984). Although this region contains one AGC box at position -205 to -211, subsequent transient expression assays using bean protoplasts more precisely defined the ethylene responsive element to reside between -236 and -305 (Roby et al., 1991). Deletion of the AGC box had no effect on overall expression and it did not appear to function as an enhancer in the bean system. A construct of 1.7 kb of the bean chitinase promoter fused to the GUS gene in transgenic tobacco showed that the promoter is activated by infection with several pathogenic fungi (Roby et al., 1990). GUS expression was very strong around the necrotic regions, but was much reduced further from the infection site, indicating that promoter activity was not strongly induced systemically (Roby et al., 1990).

Transient expression studies of various tobacco chitinase chn50 promoter-GUS constructs in tobacco protoplasts identified a distal region between -345 and -788 containing sequences required for high level expression, whereas the region between -47 and -68 contains a silencer. Two 12 bp direct repeat sequences, which include one AGC box motif, are located at -698 and -637. The silencer region contains an inverted repeat located between -49 and -70 (Fukuda et al., 1991). The region
between positions -539 and -518 contains a direct repeat, GTCAGAAAAGTCAG. Methylation interference experiments revealed that the repeated sequences are core sequences for the binding of nuclear proteins, suggesting that this motif is a candidate for an elicitor-responsive element (ERE) (Fukuda and Shinshi, 1994).

iii) **PR-5 Genes**

Analysis of transgenic tobacco plants containing acidic PR-5 (thaumatin-like)-E2 promoter-GUS constructs of varying lengths identified a region between -718 and -1364 to be involved in TMV induction of gene expression (Albrecht et al., 1992). GUS was found both locally within TMV-infected leaves, and systemically within virus-free leaf material. Transcriptional activation of a chimeric construct containing 1 630 bp of the basic PR-5 protein, osmotin, promoter and GUS was observed in the presence of ABA and 100-200 mM NaCl. The promoter lacks any consensus ABREs (Nelson et al., 1992).

iv) **PR-6 Genes**

Promoter studies of the potato proteinase inhibitor II gene in transgenic tobacco uncovered a 421 bp sequence between -136 and -557 which is required for wound-induced expression. A 10 bp sequence within this region, AAGCGTAAGT, binds to a wound-inducible nuclear protein derived from tomato leaves. This motif lies adjacent to an 8 bp motif, ACCTTGCC, which is present in several elicitor- and light-responsive promoters (Palm et al., 1990). A methyl jasmonate (MJ) responsive element which coincides with the position of a G-box motif at -574 was also identified by promoter deletion analyses (Kim et al., 1992).

v) **PR-10 Genes**

In a parsley PR1 promoter an elicitor-inducible in vivo footprint (Galas and Schmitz, 1978) was found around position -240, suggesting that protein-DNA interactions important for elicitor-induced expression occur at this site. A constitutive footprint was located further downstream at position -130. An 11 bp motif was found in opposite orientation at the sites of both footprints. The sequences did not show homology to
sequences in other elicitor-responsive promoters (Meier et al., 1991).

Regulatory elements involved in elicitor-mediated induction of gene expression were studied for a second group of parsley PR genes (pr2). A 116 bp sequence from -52 to -168 was found to be necessary and efficient for responsiveness to fungal elicitor-mediated gene expression (Van de Locht et al., 1990).

A wound-responsive gene, AoPR1, related in sequence to the parsley PR1 and PR2 genes, has been isolated from asparagus. Transgenic tobacco plants containing a 982 bp aopr1 promoter-GUS construct showed strong GUS expression localised to wound sites and in the vicinity of necrotic lesions caused by infection with Botrytis cinerea (Warner et al., 1993). This promoter contains sequences very similar to the ACCTGACC motif in the parsley phenylalanine ammonia-lyase gene and to the H-box sequence (Warner et al., 1993).

2.2.2.5 Light-inducible promoters

Plants use light as a source of information about their environment as well as a source of energy for photosynthesis, and can sense both the quantity (fluence) and quality (wavelength) of light (Terzaghi and Cashmore, 1995). Light-induced responses require changes in both nuclear and chloroplast gene expression and provide the basis for much of plant development (Kuhlemeier et al., 1987). Many of the regulatory effects of light on the growth and differentiation of plants are mediated by the control of gene expression for numerous proteins, including those for a variety of enzymes that change in response to irradiation (Gilmartin et al., 1990). The induction of gene expression as well as the degree to which the genes are expressed are not only determined by the quality and quantity of light, but also by the developmental state of the cells intercepting the light (Fluhr and Chua, 1986). Transcriptional regulation became a major focal point after it was found that in vitro translation of mRNA from light and dark-grown plants yielded different sets of polypeptides (Apel and Kloppstech, 1978). Several review papers have been written on light-regulated gene expression (Tobin and Silverthome, 1985; Gilmartin et al.,...
The general paradigm adopted for light-regulated transcription is that upon light perception, photoreceptors generate signals that are transduced via intermediates to activate transcription factors bound to their cognate sequences within regulatory regions of light-regulated genes (Terzaghi and Cashmore, 1995). The most extensively studied light-responsive genes are those encoding the small subunit of ribulose-1,5-bisphosphate carboxylase-oxygenase (rbcS) and the chlorophyll a/b-binding proteins (cab) (Gilmartin et al., 1990). Numerous regulatory cis-acting sequences have been identified by a combination of deletion and mutagenesis studies, but relatively few have been studied in detail. The following generalisations were made in a review by Terzaghi and Cashmore (1995):

• Similar elements are found in many different promoters.
• Elements that may be the same are assembled differently, and even promoters of the same gene family can be quite different.
• No universal element is found in all light-regulated promoters.
• No element is sufficient to confer light-responsiveness by itself to a non-light-regulated minimal promoter.
• The same element can confer different expression characteristics in different contexts.
• Monocots may use different elements for light regulation than do dicots.
• Light regulation may involve turning genes off as much as turning them on and some light-regulatory sequences may repress rather than activate transcription.

Elements found repeatedly in light-regulated genes include the G-box, the GT-1 site, l-boxes, and AT-rich elements (Terzaghi and Cashmore, 1995).

GT-1 binding sites, or boxes II and III of the pea rbcS3a promoter, have the core sequence GGTAAA, but the transcription factor GT-1 will bind to many related sequences such that the consensus site is Gpu(T/A)AA(T/A). Similar sequences are
found in the promoters of light-regulated genes including \textit{rbc}s from many species, \textit{phya} from oats and rice, \textit{cab} from tobacco, \textit{rca} and \textit{peta} from spinach and \textit{chs15} from bean (Terzaghi and Cashmore, 1995). L-boxes, which are defined as GATAA, are regulatory elements that are functionally important in many light-regulated promoters of both monocots and dicots. Related GATA motifs are also found in many other promoters, some that are not regulated by light (Terzaghi and Cashmore, 1995). G-boxes are elements with the core CACGTG, and these are found in the promoters of many genes that respond to a variety of different stimuli. The H-box motif is essential for both light regulation and elicitor induction, and has the consensus sequence CCTACC.

AT-rich sequences that are bound by nuclear proteins have been found in numerous light-regulated promoters but have not been functionally characterised. The \textit{Lemna gibba} \textit{Lhcb} gene encoding a light-harvesting chlorophyll \textit{a/b} protein contains two 10 bp regions that are essential for phytochrome-responsiveness. One region contains a CCAAT motif (-125 to -134) and the other a GATA motif (-106 to -114). These motifs are highly conserved in \textit{Lhcb} genes as is shown in Table 2.1 (Kehoe \textit{et al.}, 1994).

GC-rich elements have been identified in many promoters and a protein binding to a 21-bp GC-rich motif from the \textit{cabe} promoter was identified in tobacco nuclear extracts. Although a specific role for the GC-rich elements has not been demonstrated conclusively in dicots, elements identified as important for light regulation in monocots, but that are not recognised by dicots, are often GC-rich (Schäffner and Sheen, 1991).

\subsection*{2.2.3 Tissue-specific expression}

Development of a multicellular organism is critically dependent on the proper temporal and spatial cues that regulate gene expression. Thus, in any one cell at a particular time, only a subset of the total genetic information is expressed. Typically, the regulatory elements that are responsible for directing specific transcription
initiation reside within the 5' promoter region of the gene (Lam, 1994). Numerous genes and associated promoters have been described which exhibit a wide range of tissue-specific or developmental expression patterns. Some include genes that are specifically expressed in pollen (Miyoshi et al., 1995), flower (Faktor et al., 1996), phloem (Yang and Russell, 1990; Rhode et al., 1995), root (Depater and Schilperoort, 1992) and seed (lida et al., 1995; Nunberg et al., 1995; Van der Geest and Hall, 1996).

The specific regions of promoters that are necessary for tissue-specific expression are identified by deletion analysis and reporter gene expression studies. This approach was used, for example, to identify the location within the proximal 295 bp of the β-phaseolin phas promoter of sequence motifs that have been implicated in directing seed-specific expression (Van der Geest and Hall, 1996). These include the legumin box (CATGCATG) which has been identified in other seed-specific promoters, for example soybean glycinin (lida et al., 1995), and endosperm boxes (TGTAAGGT and RTGAGTCAT) (Van der Geest and Hall, 1996).

Deletion analysis of the helianthinin promoter of sunflower revealed regionalised GUS expression in transgenic tobacco seeds. Deletions up to -74 resulted in GUS activity in the cotyledons extending through the shoot apical region but not into the mid zone. Additional sequences up to -116 did not alter the shoot/ cotyledon expression pattern. However, sequences between -116 and -321 extended the tissue range of GUS expression to include most of the mid zone of the seed. Staining of the cotyledons, shoot and entire root apical regions were obtained when the -739 deletion was examined (Nunberg et al., 1995).

Functional dissection of a bean chalcone synthase gene promoter (chs15) revealed that maximal floral and root-specific expression required sequence elements located on both sides of the TATA-box. Two adjacent sequence motifs located near the TATA-box, the G-box (CACGTG) and H-box (CCTACC(N),CT), were both essential for floral expression (Faktor et al., 1996).
One element in the promoter of the \textit{CHS1} gene of white mustard, Unit 1, was found to mediate organ-specific expression in seedlings and flowers. Unit 1 consists of two motifs, Box I (GTCCCTCCAACCTAACC) and Box II (CCACGTGGCC), and is located between -131 and -168 (Kaiser and Batschauer, 1995). Box II includes the G-box, as was observed in the \textit{chs15} promoter (Faktor et al., 1996).

The promoter of maize sucrose synthase-1 (Sh) directs phloem cell-specific expression of GUS in transgenic tobacco, but not in xylem parenchyma, vascular cambium or any other cell types of vegetative tissue. No specific sequence motifs were identified in this promoter (Yang and Russell, 1990).

2.2.4 Organelle Promoters

Plastids contain their own genome in the form of multiple, double-stranded, circular DNA molecules coding for the organellar rRNAs and tRNAs as well as for mRNAs specifying almost 100 plastid-localised proteins. The latter can be grouped into proteins involved in photosynthesis and CO\textsubscript{2} fixation, and those involved in plastid gene expression itself, such as ribosomal proteins, translation factors and components of the plastid transcription apparatus (Link, 1994). Despite the high coding capacity of the organelle genome, the majority of plastid proteins are nuclear-encoded and post-translationally imported from the cytoplasm, suggesting a balanced network of regulatory interactions within each plant cell (Link, 1994).

Regulation of plastid gene expression occurs at various levels, ranging from DNA modification and transcription, to RNA maturation, translation and post-translational modifications. By using sequencing and mapping techniques, it has become possible to locate numerous plastid genes and to identify their coding regions and regulatory sequences. It has been shown that -35/-10-like motifs (-35: TTGACA and -10: TATAAT) typical for many \textit{E.coli} promoters (Reznikoff et al., 1985) are present in front of the mapped RNA 5' end of a number of mustard chloroplast genes, for example the \textit{rm} operon, \textit{psbA}, \textit{tmK} and \textit{tmQ} (for a review see Link, 1994).
A putative blue light-responsive promoter (LRP) upstream of the psbD-psbC chloroplast operon has been identified in several plant species. It shows a slight resemblance to the -35/-10 consensus sequence and reveals GATA motifs reminiscent of light-responsive plant nuclear genes (Gilmartin et al., 1990). There are examples of genes that lack the -35-type element, but the TATA-box consensus sequence element (−10) has been shown to be required for faithful transcription in a chloroplast in vitro system (Eisermann et al., 1990).

2.3 ISOLATION AND ANALYSIS OF PLANT PROMOTERS
Several approaches have been used to isolate DNA fragments carrying regulatory elements from plants, with the choice of techniques to be used in a study depending on the particular circumstances.

Two distinct groups of techniques can be identified. The first group is used for the identification and isolation of promoters of genes that have not previously been identified as being useful. These techniques are not directed at any specific gene, but rather at a group of genes that are expressed in a certain tissue, at a specific developmental stage, or in response to specific stimuli. The second group of techniques is directed at specific genes, and is used when more information is available about the gene product. All these techniques involve the isolation of cDNA fragments that can be used to screen a genomic library and isolate the promoter sequence. Promoter tagging methods can also be used to detect and isolate regulatory regions directly from the plant genome (An and Kim, 1993).

2.3.1 Non-directed methods for the identification and isolation of useful promoters
Differential or "+/-" screening of cDNA libraries have been successfully used to identify and isolate cDNAs representing genes whose expression is altered by stress. The cDNAs identified in this manner do not represent genes with known functions. During differential screening replicates of a cDNA library are screened by hybridisation to the RNA from which the library was constructed. For example, cDNA
synthesised from RNA isolated from drought-stressed organs can be compared to cDNAs synthesised from RNA isolated from non-stressed organs. cDNAs that hybridise to the source of the drought-stress library and not to the non-stressed comparison are candidates for genes that are induced by stress (Bray, 1994). The differential screening technique has been used in the identification of cDNAs representing genes from maize (Gómez et al., 1988), pea (Guerrero and Mullet, 1988), rice (Mundy and Chua, 1988), Craterostigma plantagineum (Bartels et al., 1992), tomato (Cohen and Bray, 1990) and Arabidopsis (Yamaguchi-Shinozaki et al., 1992).

A PCR differential screening method for the isolation of clones from a cDNA library has also been developed (Luo et al., 1994). This method involves a primary differential screening step followed by a PCR screening step. From the primary screening step, impure pools of positive cDNA clones are obtained. Each pool of clones is then amplified directly using primers flanking the cloning site in the vector. The PCR products are separated on two duplicate agarose gels and blotted onto filter strips that are subjected to differential Southern hybridisation with different cDNA probes. Each pure positive cDNA band on the gel is identified and selected for further analysis (Luo et al., 1994).

Another variation of this technique is the use of subtractive hybridisation which is based on subtracting away the common transcripts between different cell types or tissues, leaving the specific transcripts (mRNA) for further analysis (Sargent, 1987). A "tracer" cDNA is hybridised to an excess of "driver" mRNA from another cell type. Unhybridised tracer cDNA represents an enriched population of sequences that are expressed in the one cell type only (Lee, et al., 1991). Subtraction can also be performed by hybridisation between mRNA of one cell type and the cDNA synthesised from mRNA from another cell type using an oligo(dT)30 primer that is covalently linked to Latex particles in an Eppendorf tube. The mRNA common to both cell types can be removed by a brief centrifugation (Hara et al., 1993).
An improvement of this subtraction approach has been the use of magnetic bead-assisted subtraction (Sharma et al., 1993). First strand cDNA is immobilised onto magnetic beads as a result of the oligo(dT)$_{25}$ which is attached to the beads (Aasheim et al., 1994). This subtraction is also known as "magnet-assisted subtraction technique" or MAST (Schraml et al., 1993). A combination of this technique and a Southern blot screening by means of PCR (MAST-PCR-SBDS) was recently used to identify cDNAs from genes that are differentially expressed during strawberry fruit ripening (Medina-Escobar et al., 1997).

The major limitations of all these techniques are that they can only be used to identify genes that are relatively highly expressed, they are time-consuming and often difficult to perform, and they can only be used to compare two different cell types at a time. The differential display technique (Liang and Pardee, 1992) circumvents some of these limitations and will be discussed in greater detail.

2.3.1.1 Differential display of mRNA by PCR
RNA fingerprinting by the differential display technique has been used to isolate differentially expressed genes in a number of experimental systems (for a review see McClelland et al., 1995). The technique was developed with the goal of identifying differentially expressed genes by detecting individual mRNA species that are altered in mammalian cells and then recovering and cloning the cDNA (Liang and Pardee, 1992; Liang et al., 1993).

This approach involves PCR amplification of cDNA fragments and polyacrylamide gel electrophoresis for the separation of amplified products. The basic principle is to reverse transcribe and systematically amplify the 3' termini of mRNAs with a set of anchored oligo(dT) primers and a random decamer. Four sets of degenerate anchored oligo(dT) primers, (T$_{12}$MN), where M can be G, A or C and N is G, A, T or C, are used to prime reverse transcription by binding to the start of the poly(A) tails of the mRNA. Each primer set is dictated by the 3' base (N), with degeneracy in the penultimate (M) position. The resulting cDNA population is PCR-amplified using the
degenerate primer set, an arbitrary decamer, and a radioactive nucleotide. The radioactively labelled PCR products that represent a subpopulation of mRNAs defined by the given primer set are separated in a denaturing polyacrylamide gel. By changing primer combinations, most of the mRNA species in the cell may be represented. Multiple different RNA samples are compared to allow the identification and isolation of differentially expressed genes (Liang et al., 1993). The procedure depends on the amplification of an mRNA subpopulation without prior knowledge of sequence information.

The differential display technique has been used with varying degrees of success. Big drawbacks of the technique are the high incidence of false-positives and the fact that the technique often shows a strong bias towards high copy number mRNAs (Bertioli et al., 1995). Several changes to the original method have been made and certain critical parameters have been identified since the technique was first developed. These parameters are discussed below and changes that have been made by other research groups are included.

i) RNA template for RT-PCR

One of the most important factors for successful differential display is to use DNA-free RNA samples (Liang et al., 1993). One way of determining whether an RNA sample contains contaminating DNA is to perform the PCR without reverse transcription. If the sample is free of DNA, the display should be totally dependent on the reverse transcription step (Liang et al., 1993).

Because the technique relies on the reverse transcription of mRNA from an oligo(dT) primer, both total RNA (Bauer et al., 1993; Zimmermann and Schultz, 1994; Zhao et al., 1995) or purified poly(A)mRNA (Liang et al., 1993; Sokolov and Prockop, 1994) can be used. The use of cytoplasmic rather than total RNA eliminates false positives due to nuclear RNA that is not transported to the cytoplasm (Sompayrac et al., 1995).

Several investigators optimised the RNA template concentration for RT-PCR. Welsh,
et al. (1992) varied the concentration of input RNA, and serial dilutions of the RNA ranging from 0.012 to 2.5 μg per 20 μl reaction before cDNA synthesis revealed a broad concentration optimum. They concluded that the method would be reproducible over a wide range of RNA concentrations (Welsh, et al., 1992). In another series of reactions that were done to determine the sensitivity limits of the differential display method in terms of input RNA concentration, total RNA was varied from 0.002 to 2 μg per 20 μl RT reaction (Liang et al., 1993). Results indicated that the amount of input RNA can be as low as 0.02 μg which is equivalent to RNA from about 200 cells, but that the banding pattern was much sparser when 0.002 μg RNA was used (Liang et al., 1993).

The possibility of using very small RNA samples is a major advantage of the differential display technique because in some cases only very small tissue samples can be obtained. An example of this was the analysis of differential expression during preimplantation embryogenesis in mouse embryos. Reverse transcription was successfully performed on RNA obtained from a number of embryos that corresponded to about 10 pg of poly(A)+mRNA (Zimmermann and Schultz, 1994).

ii) Primers used for RT-PCR

Twelve anchored oligo(dT) primers consisting of 11 or 12 T's plus two additional 3' bases were initially used for reverse transcription in twelve separate reactions (Liang and Pardee, 1992). The number of primers have since been reduced to four to improve the efficiency and specificity of the technique, each primer being degenerate at the 3' penultimate position (T_{12}MN where M may be G, A or C and N may be G, A, T or C) (Liang et al., 1993; Sompayrac et al., 1995).

Results of an investigation into the specificity of anchored oligo(dT) primers suggested that the two 3' bases, MN, are absolutely essential to provide the primers with specificity for anchoring to the beginning of the poly(A) tails. The penultimate base, M, can exhibit considerable degeneracy because it is the last base on the 3' end that provides most of the specificity (Liang et al., 1993). Because most of the
fragments obtained from reverse transcription using oligo(dT) primers are situated on the 3' end of the mRNAs, some groups found it more advantageous to use random hexamers for reverse transcription. This ensured that fragments with internal sequences of the mRNAs were obtained instead of the less-informative 3' sequences (Sokolov and Prockop, 1994). In another case an oligo(dT) \(_{12-18}\) primer with no anchoring bases was used for the reverse transcription, while anchored oligo(dT) primers were used in subsequent amplification reactions (Doss, 1996).

For the amplification of the synthesised cDNA fragments, random decamer primers were used in combination with the oligo(dT) primer that was used for reverse transcription. Generally, any decamer primer with a GC-content of 50% to 70% can be used as long as it does not contain palindromic sequences. Since it has been shown that the random decamer primers can have up to 4 bp mismatches with the original cDNA template, and that these mismatches are often clustered on the 5' end of the primers (Liang et al., 1993), the primers can be designed in such a way that the 3' sequences are maximally randomised while the 5' bases (up to 4 bases) are fixed.

Several other investigations were done involving the upstream PCR primers (Bauer et al., 1993; Sokolov and Prockop, 1994; Bertioli et al., 1995; Zhao et al., 1995), and one of the conclusions that was made was that under the low stringency PCR conditions that are used in the differential display technique, mismatches in the priming site can occur and lead to the amplification of false-positive fragments and irreproducible PCR displays (Bauer et al., 1993; Zhao et al., 1995). The use of longer primers and elevated annealing temperatures was therefore suggested to minimise false positives. In an enhanced differential display technique (EDD), the stringency was increased by increasing the length of the primers to 22 nucleotides and using low annealing temperatures for only the initial few PCR cycles. Under these conditions the primers behave in a similar way to the short 10 nucleotide primers. For the remaining PCR cycles, the annealing temperature was raised during which the products made in the initial cycles were efficiently and specifically replicated
Because random decamer primers are too short to prime a sequencing reaction, the fragments have to be cloned if the sequence is to be determined. The primers can be extended and used for direct sequencing of cDNA fragments (Wang and Feuerstein, 1995; Reeves et al., 1995) which will overcome the problem of some fragments being too short for efficient cloning or labelling (Sokolov and Prockop, 1994). Longer primers can also be designed to contain restriction enzyme sites which will simplify the cloning of the fragments (Linskens et al., 1995).

The number of primer combinations will depend on the application and the extent to which the cell types differ from each other. Liang and Pardee (1992) suggested that 20 random decamer primers must be used to obtain all possible mRNA sequences upstream of the 12 anchored oligo(dT) primers, because the decamer primers act like 6- to 7-mers under low stringency reaction conditions. Results obtained by Bauer et al. (1993) confirmed this and showed that PCR with decamer primers generates about 120 fragments after 40 cycles using an annealing temperature of 40°C. This corresponds to an effective primer length of 6 bases. In order to reach a confidence level of 0.95 for displaying any messenger as a cDNA fragment in a gel, one should use at least 25 upstream primers (Bauer et al., 1993).

iii) Display of amplified cDNA fragments
Initially, the differential display amplification products were separated in denaturing polyacrylamide gels which were dried after electrophoresis. Because a radioactive nucleotide was included in the amplification reaction mixture, the banding profile could be visualised by autoradiography (Liang and Pardee, 1992). The dried gel was oriented with the autoradiogram, and cDNA fragments of interest could be cut from the gel. Several variations of this method for the display of cDNA products have been developed. Non-denaturing polyacrylamide gels were used to reduce an observed artificial complexity of the banding pattern (Bauer et al., 1993; Bertioli et al., 1995), while silver staining of the polyacrylamide gels eliminated the need for radioactivity (Lohmann et al., 1995; Doss, 1996).
The separation of amplification products in agarose gels, followed by ethidium bromide staining, was used by Sokolov and Prockop (1994) and proved to be sensitive enough to detect differences in mRNA expression that appeared after infection of plant leaves with pathogens (Rompf and Kahl, 1997). Although the resolution of agarose gels is much lower than that of polyacrylamide gels and fewer amplified fragments can be resolved, this can be overcome by increasing the number of primer combinations used.

More recently, the use of digoxigenin-conjugated or biotinylated oligo(dT) primers and colorimetric or chemiluminescence-based detection of PCR products has been described (An et al., 1996). A fluorescent detection procedure, using a fluorescently labelled oligonucleotide has also been developed. This has the advantage that the amplification products can be analysed using an automated gene sequencer, which facilitates a high throughput of samples, easier data capture, and storage of data for future profile comparison (Smith et al., 1997). Several investigators have found that even with the availability of very sensitive detection methods, the differential display should be used as a qualitative rather than a quantitative technique (Liang et al., 1993).

iv) Isolation of differentially expressed cDNA fragments
After the banding profiles have been analysed, the differentially expressed sequences have to be isolated. Where denaturing polyacrylamide gel electrophoresis has been used for the separation of products, the gels are dried and after autoradiography the fragments of interest are excised, eluted from the gel, and reamplified prior to their analysis (Liang et al., 1993).

Several problems have occurred in the isolation of specific fragments, and many investigators have reported the isolation of more than one fragment from the same position in the gel (Welsh et al., 1992; Callard et al., 1994; Sokolov and Prockop, 1994; Hadman et al., 1995). After cloning an isolated fragment, sequencing of several different clones of the expected size revealed that the sequence of the insert
was heterogeneous (Callard et al., 1994). Hybridisation techniques have therefore been developed to identify the clones that contain the true differential cDNA fragment (Callard et al., 1994).

Gery and Lavi (1997) have also described a method to improve the efficiency of the isolation of differentially expressed fragments. Reamplification of a fragment can be done in the presence of an isotope, after which the reamplified product is run in parallel with the original PCR products to confirm that the correct fragment has been reamplified. Raising the annealing temperature of the reamplification reaction may also reduce contaminating background (Gery and Lavi, 1997).

2.3.2 Directed techniques for the isolation of useful promoters
When more information is available about the gene product to be studied, directed methods can be used to isolate the desired cDNA. The directed techniques include heterologous hybridisation, hybridisation with degenerate primers after partial amino acid sequence determination, PCR-based screening, and functional complementation to identify specific cDNAs (Bray, 1994).

If a similar cDNA has been isolated from a different species it can be used to screen a cDNA library from the desired species. This method has, for example, been successfully used in the identification of an osmotin-like gene from Atriplex (Casas et al., 1992). A salt- and drought-induced gene was also isolated from rice using a degenerate oligonucleotide (Claes et al., 1990). The expressed salt-induced protein was identified by two dimensional polyacrylamide gel electrophoresis, and the aminoterminal amino acid sequence of the tryptic peptide was determined after transfer from the 2-D gel onto nitrocellulose. An oligonucleotide was synthesised based on the amino acid sequence and was used to screen a cDNA library (Claes et al., 1990).

In cases where a number of related proteins have been identified, regions of conservation among these proteins can be used to identify a specific mRNA. PCR primers can be designed and a probe can be synthesised by PCR that can be used
to screen a cDNA library. This PCR-based method has been used to isolate a drought- and abscisic acid (ABA)-induced protein kinase from wheat (Anderberg and Walker-Simmons, 1992). Degenerate oligonucleotides corresponding to highly conserved regions in serine / threonine protein kinases were designed for PCR amplification. Consequently, the cDNA synthesised from poly(A+)mRNA isolated from dormant wheat embryos was able to be amplified (Anderberg and Walker-Simmons, 1992).

If a bacterial mutant is available, functional complementation of that mutant may be used to obtain the desired cDNA. Many plant genes can be isolated by their ability to complement mutations in bacteria and yeast (Gibson and Somerville, 1993). For example, each of eight different auxotrophic mutants of yeast were functionally complemented by a cDNA library from *Arabidopsis*. Two genes involved in higher plant proline biosynthesis have also been cloned by functional complementation of *E.coli* mutants (Delauney and Verma, 1990; Hu *et al.*, 1992). A phagemid library produced in Lambda Zap II (Stratagene) and excised *in vivo* with a helper phage was transformed into an *E.coli* proC mutant. This mutant has a defective Δ1-pyrroline-5-carboxylate reductase gene and is unable to grow in media lacking proline. The complemented mutant was able to grow on media lacking proline, indicating that the gene product functions in the proline biosynthetic pathway (Delauney and Verma, 1990).

### 2.3.3 Isolation of tissue-specific promoters from sugarcane

Several options for isolating gene regulatory sequences with desired patterns of expression in sugarcane have been discussed (Birch *et al.*, 1996). The first approach involves the use of characterised promoters from other plant species. These promoters can be used either directly to drive introduced gene expression, or indirectly as the basis for probes or primers used to isolate corresponding promoters from sugarcane. Limiting factors of this approach are the availability of suitable foreign promoters and in some cases, patent considerations (Birch *et al.*, 1996).
Introduction of promoterless reporter genes to tag sugarcane promoters is another possible technique to use (Koncz et al., 1989). The relatively low efficiency of promoter tagging by direct gene transfer into plants and difficulty in recovering the tagged promoters have limited the use of this approach. Differential hybridisation approaches to isolate tissue-specific cDNA fragments and then to use these to recover corresponding gene and promoter sequences from genomic libraries, have been used in other plant systems. Constraints include the requirement for large amounts of RNA, technical difficulties in cDNA library construction, and a bias toward strongly expressed genes (Watson and Margulies, 1993).

The differential display technique discussed previously can be used to isolate tissue-specific cDNA fragments from sugarcane. These fragments can then be used as a starting point for isolating corresponding gene promoters by PCR or hybridisation techniques (Liang et al., 1993). The complex polyploid genome of sugarcane may create additional difficulties in isolating the correct promoter sequence driving an observed useful pattern of expression in differential hybridisation and display techniques (Birch et al., 1996).

Preliminary work has been done to isolate wound- and root-specific promoters from sugarcane. A differential hybridisation approach was used to isolate a full-length root-specific cDNA clone. Northern hybridisation analysis indicated that the gene is expressed in roots under various stress conditions, but not in leaves, stems or healthy field-grown roots. A corresponding 770 bp putative root-specific promoter region was isolated from a genomic library and fused to both GUS and luc reporter genes. Transient expression analysis has been done and expression levels in transformed roots were low. Insertion of an upstream enhancer (OCS) increased overall expression levels, but not root-specificity. This promoter is currently being characterised in transgenic plants (Birch et al., 1996).

A wound-induced peroxidase cDNA fragment from buffel grass has been used to isolate the corresponding gene from a sugarcane genomic library. The putative
promoter region is currently being characterised in transgenic sugarcane plants (Birch et al., 1996).

Culm-specific cDNA fragments have been isolated from a sugarcane cDNA library by differential hybridisation and are being characterised (Potier and Birch, 1997a). Two fragments were specifically expressed in the meristematic region of the culm, one was expressed in the young culm only, and another fragment was fully culm-specific with its expression peaking in the top and in the mature parts of the culm (Potier and Birch, 1997a). These fragments have been sequenced enabling the identification of a likely functional role for some of them. Based on expression patterns, copy number and gene identity, some of the cDNA fragments have been selected for the recovery of the corresponding promoter sequence by an inverse PCR (iPCR) technique. One of the isolated promoters has been used to make a fusion construct to the GUS reporter gene, and sugarcane and tobacco have been transformed with this construct. Further analysis remains to be done (Potier and Birch, 1997b).

2.4 PLANT TRANSFORMATION

Genetic transformation can be defined as the transfer of foreign genetic material isolated from plants, viruses, bacteria or animals into a new genetic background, resulting in chromosomal (nuclear or organellar) integration and the stable heritability of that material through meiosis (Webb and Morris, 1992; Gruber and Crosby, 1993). This definition distinguishes transformation from transient expression, where genetic material is introduced through physical uptake approaches and may exhibit short-lived mitotic stability in cells (Gruber and Crosby, 1993).

In plants, successful genetic transformation requires the production of normal, fertile plant which express the newly inserted genes (Webb and Morris, 1992). Transformation can utilise gametophytic tissues, prior to fertilisation, or somatic cells, such as explants, cells and protoplasts, which can be stimulated to regenerate plants de novo in culture (Webb and Morris, 1992).
Conceptually there are two reasons for transforming a plant. Firstly, it is used as an experimental tool for plant physiology where the capacity to introduce and express or inactivate specific genes allows the elucidation of certain biological processes. Secondly, it is used as a practical tool to confer a trait to a plant that will improve its agricultural, horticultural or ornamental value (Greenberg and Glick, 1993; Birch, 1997). Some of the traits that have been “engineered” into plants by the addition of a single gene or a small cluster of genes include insecticidal activity, tolerance to viral or fungal infection, herbicide resistance, altered fruit ripening, improved nutritional quality of seed proteins and self-incompatibility (Gasser and Fraley, 1989; Gatehouse et al., 1992; Peferoen, 1992; Glick and Thompson, 1993; Smith et al., 1993; Birch, 1997).

There are a number of reviews of the development of transformation systems including Agrobacterium-mediated transformation, direct gene transfer into protoplasts and particle bombardment (Hooykaas and Schilperoort, 1992; Zupan and Zambryski, 1995; Birch and Franks, 1991; Christou, 1992), and their potential for practical application (Gatehouse et al., 1992; Setlow, 1997). The methods involved have also been discussed in detail (Gatehouse et al., 1992; Glick and Thompson, 1993). In recent reviews, key problems remaining in the development of plant transformation systems, key issues to be resolved in the practical application of these systems, and strategies to overcome these limitations have been addressed (Christou, 1996; Birch, 1997).

### 2.4.1 Transformation via particle bombardment

Monocotyledonous plants, including important grain crops, have been difficult to transform. One reason is that these plants are not susceptible to infection by Agrobacterium tumefaciens. Consequently, the delivery of DNA to the cells using particle bombardment has become the method of choice for the transformation of these plants (Greenberg and Glick, 1993; Morrish et al., 1993). With this technique, microprojectile particles (e.g. 1.2 µm tungsten or gold beads) coated with DNA are accelerated at high speeds into plant tissue (Christou, 1993). A number of the cells
usually express the introduced gene transiently, while some are capable of stably integrating the foreign DNA into the genome. This approach has the distinct advantage of applicability to any intact plant tissue or region of the plant, and has been used to transform organellar DNA (Svab et al., 1990).

Since the first report of transformation of living cells using high velocity microprojectiles (Klein et al., 1987), several crop species that were not previously accessible to transformation have been successfully transformed (Sautter et al., 1991). Crop plants that have been transformed using this technique include maize (Klein et al., 1988; Oard et al., 1990; Vain et al., 1993a; Rasmussen et al., 1994; Kemper et al., 1996), wheat (Wang et al., 1988; Lonsdale et al., 1990; Chibbar et al., 1991; Vasil et al., 1991), barley (Mendel et al., 1989; Koprek et al., 1996), cotton (Finer and McMullen, 1990), tobacco (Klein et al., 1988; Tomes et al., 1990; lida et al., 1991), soybean (Christou et al., 1990; McCabe et al., 1988; Sato et al., 1993), rice (Christou et al., 1991; Li et al., 1993) and sugarcane (Franks and Birch, 1991; Bower and Birch, 1992; Gambley et al., 1993). Microprojectile bombardment has also been used as an effective method to wound plant tissues to promote *Agrobacterium*-mediated transformation (Bidney et al., 1992).

Several different types of bombardment devices have been developed, including an electrically triggered discharge gun (McCabe and Christou, 1993), pneumatic particle guns (Kikkert, 1993; Oard, 1993), helium, nitrogen and carbon dioxide powered devices (Finer et al., 1992; Vain et al., 1993b) and a micro-targeting gun (Sautter et al., 1991). Most of these devices are based on the same basic concept for particle acceleration and delivery: a force provided by either an explosion or expansion of compressed gas propels a macrocarrier holding particles towards an immobile object which retains the macrocarrier but permits the particles to pass (Finer et al., 1992).

The particle inflow gun (PIG)(Finer et al., 1992) is based on the acceleration of DNA-coated tungsten particles using pressurised helium in combination with a partial vacuum. The particles are accelerated directly in a helium stream rather than being
supported by a macrocarrier (Finer et al., 1992).

There are a number of features of the PIG that make it easy to construct and operate. Expansion of compressed helium is used since helium is inert, leaves no residue, and gives reproducible acceleration. A timer relay-driven solenoid is used to release the helium and can be used at low helium pressure, which allows good particle penetration and less damage to the target tissue. The particles are accelerated directly in the helium stream and no macrocarriers are necessary, saving costs and time in between bombardments. A vacuum chamber is used to hold the target tissue. The vacuum reduces the aerodynamic drag on the particles, thereby increasing particle velocities, and lessening tissue damage by dispersion of the helium gas prior to impact (Finer et al., 1992).

Parameters that have to be optimised for the PIG are helium pressure, distance of the target tissue from the syringe filter, packed cell volume of target cells and volume of particle preparation to be used. Other particle accelerating devices have been developed with various modifications on the original PIG, and these are commercially available (Brown et al., 1994; Buckley et al., 1995).

2.4.2 Vectors for plant transformation

Once a gene of interest is isolated, it is introduced into a plant transformation vector. The vector DNA facilitates manipulation of the gene in Escherichia coli prior to plant transformation, as well as transfer of the gene to the host plant. An ideal vector contains a multiple cloning site, an antibiotic resistance gene allowing for selection in bacteria (e.g. ampicillin resistance), a broad-host bacterial origin of replication, and an antibiotic resistance gene for selection of the foreign DNA in transformed plants (e.g. kanamycin resistance) (Walden et al., 1990; Greenberg and Glick, 1993).

Agrobacterium plasmids have been exploited as natural vectors for biological delivery of foreign DNA to plants and this has become the most widespread transformation strategy in use (Gruber and Crosby, 1993). Agrobacterium tumefaciens carries the
tumour-inducing or Ti plasmid whereas *Agrobacterium rhizogenes* contains the root-
inducing or Ri plasmid. Two types of vector systems have been developed, namely
cointegrating and binary (Walden and Schell, 1990).

The cointegrating system features two independent plasmids: a Ti plasmid in
*Agrobacterium* and an intermediate vector in *E.coli*. Both plasmids have a region
of homology which undergoes recombination to form a large, cointegrated plasmid after
conjugation between *Agrobacterium* and *E.coli* (Van Haute *et al.*, 1983). Genes which
are to be introduced into plants are cloned and manipulated in *E.coli* and, after
recombination with the Ti plasmid in *Agrobacterium*, are situated between two T-DNA
border repeats. The *E.coli* plasmid has no origin of replication for maintenance in
*Agrobacterium* and is not retained without the recombination step (Shaw *et al.*, 1983).

The binary system features two plasmids which coexist autonomously in
*Agrobacterium* after conjugation, a shuttle plasmid and a Ti plasmid (Hoekema *et al.*, 1983). Shuttle plasmids contain one or both T-DNA border repeats as well as broad-
host-range replication and mobilisation functions. They encode plant selectable
markers, expression signals and polylinkers for the subcloning of foreign genes within
the two T-DNA borders. The *Agrobacterium* plasmid in a binary system is an
engineered Ti plasmid with plasmid *vir* genes, an *Agrobacterium* origin of replication
(*oriA*), but no T-DNA borders. After conjugation, the two plasmid partners coexist
autonomously with selective pressure in *Agrobacterium*. When *Agrobacterium* infects
a wounded plant, the *vir* genes on the Ti plasmid interact with the right border on the
shuttle plasmid in *trans* to transfer the T-DNA into the plant genome. These two latter
features define the binary vector strategy (Hoekema *et al.*, 1983).

The gene cassettes within the T-DNA region can be arranged for optimum transfer
and plant expression (Velten and Schell, 1985). Genes of interest should be situated
directly adjacent to the right border where integration is faithful, leaving the plant
selectable marker closer to the left border where incorrect integration has been
known to occur. Dual promoters can also be subcloned in opposite orientation in the
centre of the T-DNA, so that there is less influence on the expression of introduced genes from flanking plant genomic DNA (Klee and Rogers, 1989).

Methods for physical delivery of DNA into plants include the direct uptake of DNA into protoplasts using polyethylene glycol, electroporation or microprojectile bombardment delivery to tissues using a biolistic device. Physical vectors are usually 4 to 7 kb in size although vectors of up to 11 kb have been used without compromising transgene stability (Vasil et al., 1991). DNA conformation plays a role in the frequency of physical transformation and both linear as well as supercoiled plasmids are used as vectors (Gruber and Crosby, 1993). Plant expression cassettes in vectors include multiple selection markers (Tomes et al., 1990), tandem promoters (Hain et al., 1985) and transcription enhancing elements (Callis et al., 1987).

The expression of intron-containing genes can be strongly influenced by the sequence content and the presence of functional gene regulatory sequences in an intron (Gruber and Crosby, 1993). Several examples have been described where a plant intron may contain a position-dependent enhancer sequence, for example in the maize Adh1 and rice Act 1 structural genes (Mascarenhas et al., 1990; McElroy et al., 1990). The inclusion of an intron between the promoter of maize Sh1 and reporter gene in a chimeric construct has enhanced reporter gene expression in rice and maize protoplasts by 100-fold (Maas et al., 1991).

Organelle transformation through particle bombardment is also possible (Butow and Fox, 1990). The vectors used for these studies include genes providing resistance to antibiotics which only affect the chloroplast ribosome, such as streptomycin, spectinomycin, chloramphenicol and erythromycin to select chloroplast transformants (Haring and De Block, 1990). Chloroplast-specific replication sequences and chloroplast-driven selectable markers can also be included in an E.coli plasmid framework (Gruber and Crosby, 1993).
2.4.2.1 Selectable marker genes

Selection of transformed cells is a key factor in developing successful methods for genetic transformation (Webb and Morris, 1992). The choice of marker is dictated by the antimetabolite chosen as counterselection agent, which is used in preventing growth of untransformed cells. Other factors must be considered in choosing a marker gene, including that the expression of the marker gene should not metabolically disrupt the transformed host cell, expression of the marker gene must effectively protect the host cell from the growth-inhibitory properties of the selection agent, and exposure to the antimetabolite should have minimal effects in subsequent growth and development of the transformed plant (Gruber and Crosby, 1993).

The gene nptII, encoding neomycin phosphotransferase (NPT), is most commonly used as a selectable marker and has proven particularly effective in transformation experiments involving potato, tobacco and tomato. However, nptII has proven less successful with legumes and some monocot species (Hauptmann et al., 1988; Gruber and Crosby, 1993). The nptII gene product acts by enzymatically phosphorylating and inactivating aminoglycoside antibiotics such as kanamycin, neomycin and geneticin (G418).

Several selectable marker gene products function by detoxifying antimetabolites such as hygromycin (Waldron et al., 1985) and gentamycin (Hayford et al., 1988), while other markers encode an altered target which is insensitive to the inhibitor (Etzold et al., 1987). Marker genes encoding tolerance to herbicides have also been used. For example, the bar gene isolated from Streptomyces hygroscopicus codes for phosphinothricin acetyltransferase (PAT), which inactivates phosphinotricin (PTT), the active ingredient of Bialaphos and Basta herbicides and an irreversible inhibitor of glutamine synthase (Webb and Morris, 1992).

2.4.2.2 5' Regulatory sequences in plant expression vectors

The expression of a transgene is dependent on the presence of transcriptionally-efficient 5' regulatory regions, a 5' untranslated leader sequence, a translational start
sequence (AUG) with plant favourable context, and a 3' transcription termination/polyadenylation sequence (Gruber and Crosby, 1993). These and other determinants contribute to mRNA and protein production, stability and turnover and are all important components to consider in a strategy to develop a plant vector (Webb and Morris, 1992; Gruber and Crosby, 1993).

The choice of a promoter for plant transformation will depend on the circumstances under which the transgene should be expressed. Constitutive promoters will drive expression in all developmental stages in all tissues, whereas tissue-specific promoters will only be active in certain tissues.

The 5' regions of the maize polyubiquitin gene (Christensen et al., 1992) and the rice actin 1 gene (Zhang et al., 1991) have been used to drive constitutive gene expression in transgenic plant studies. Another commonly used constitutive promoter is the CaMV 35S promoter (Guilley et al., 1982).

Examples of tissue-specific promoters which have been used to drive expression in particular areas of the plant, include the maize sucrose synthase-I (Sh) promoter suitable for phloem-specific expression (Yang and Russell, 1990), the pea and bean vicilin and phytohemagglutinin promoters suitable for seed-specific expression, and the α-amylase promoter for driving expression in the aleurone of cereal grains (Christou, 1996).

Promoters which are induced by certain chemical (Gatz, 1997) or environmental stimuli such as light (Terzaghi and Cashmore, 1995) or heat (Czamecka-Verner et al., 1994) can also be used in transformation studies. In some cases the use of promoter cassettes will be necessary to enhance gene expression. Several combinations of promoters and terminator sequences have been tested in both dicotyledonous and monocotyledonous plants (Mitsuhara et al., 1996).

The large differences between regulatory elements from prokaryotes and eukaryotes
mean that sequences from bacteria may not function in plant cells. Plant viruses, which depend on the plant transcription and translation factors, have been used as sources of regulatory elements. The most common of these is the promoter of the 35S RNA gene of CaMV (Guilley et al., 1982). This promoter is active in all cell types, although a certain degree of cell- and tissue-specificity has been observed (Benfey and Chua, 1990; Lam, 1994).

Regulatory elements in certain genes of Agrobacterium tumefaciens and A.rhizogenes are not expressed in Agrobacterium, but become active on transfer to plant cells. The nos (nopaline synthase), ocs (octapine synthase) and mas (mannopine synthase) gene promoters belong in this category, and have been used successfully to direct expression of genes in plant cells (Webb and Morris, 1992). These promoters have been studied intensively (Singh et al., 1990; Kim et al., 1994). One example of a bacterial promoter that is being used to direct tissue-specific transgene expression in plants is the rolC promoter of A.rhizogenes. The expression pattern of this promoter in sugarcane was evaluated in the present study.

i) The rolC promoter of Agrobacterium rhizogenes

The rolA, B, C and D genes of A.rhizogenes are located on the TL-DNA of the Ri plasmid A4 and are involved in the pathogenesis of hairy root disease (Moore et al., 1979; White and Nester, 1980; Vilaine et al., 1987; Nilsson and Olsson, 1997). When expressed separately in transgenic tobacco plants, these genes established phenotypic alterations characteristic for each gene (Oono et al., 1987; Schmülling et al., 1988; Sinkar et al., 1988).

Transgenic tobacco plants expressing a single rolA gene from its endogenous promoter displayed severe wrinkling of leaves and the plants were stunted with extremely short internodes and small rounded leaves (Schmülling et al., 1988). The rolB gene displayed alterations in leaf and flower morphology and adventitious root formation on the stem in transgenic tobacco plants (Schmülling et al., 1988). Transgenic tobacco plants expressing rolC from its endogenous promoter had
increased ratio of leaf length to leaf width. The plants were also shorter, more highly branched, flowered earlier and had smaller flowers and a reduced pollen production compared to the wild type (Oono et al., 1987; Scmulling et al., 1988). When rolC was expressed from the strong 35S promoter, these characteristics were exaggerated, giving dwarfed plants with very short internodes and an increased number of lanceolate light-green leaves. Apical dominance was reduced, resulting in bushy plants (Scmulling et al., 1988; Nilsson et al., 1993). Because of the dominant pleiotropic nature of the rolC gene and the fact that it causes a reduction in the chlorophyll content in leaves of transgenic plants when placed under control of the CaMV 35S promoter, it has been used as a visible cell-autonomous marker in aspen (Populus) (Fladung and Ahuju, 1997). The expression of rolD in transgenic tobacco plants caused early flowering and explants showed earlier and enhanced organogenesis of flowers (Mauro et al., 1996).

Several groups have studied rol gene regulation by means of promoter fusions to the reporter gene uidA, and subsequent analysis of GUS activity in transgenic tobacco, hybrid aspen and rice plants (Matsuki et al., 1989; Scmulling et al., 1989; Sugaya et al., 1989; Maurel et al., 1990, Altamura et al., 1991, Capone et al., 1991, Guivarc'h et al., 1996; Nilsson et al., 1996). These studies showed that the rolD gene differs from the others, because it does not exhibit tissue-specific expression. Instead, it appears to be developmentally regulated, being expressed during the elongation and maturation phases of certain tissues. The other rol genes have a similar tissue-specific expression pattern, being mainly confined to root meristems and the phloem (Nilsson et al., 1996).

The intergenic region separating the rolB and rolC genes represents a bidirectional promoter (Scmulling et al., 1989). Detailed analysis showed that the rolB promoter seems to be generally expressed in phloem, phloem parenchyma and ray cells (Altamura et al., 1991; Nilsson et al., 1997) while the rolC promoter is specifically expressed in the phloem companion cells (Guivarc'h et al., 1996; Nilsson et al., 1996).
In expression studies of the rolC gene in transgenic tobacco plants, it was shown that the transcriptional level of ORF12 (rolC) differed greatly among organs, the order being roots > stems > leaves (Nakamura et al., 1988). This organ-related activation of the GUS gene under control of the rolC promoter was not observed in rice plants, although phloem-specific expression was observed (Matsuki et al., 1989). In hybrid aspen a cell-specific localisation of the rolC promoter was reported. Within the phloem at the shoot base GUS activity was always detected in companion cells and occasionally in ray and axial parenchyma, but never in sieve tubes (Nilsson et al., 1996). Expression was also highly specific in groups of pericycle cells prior to and during lateral root initiation (Nilsson et al., 1997).

The rolC promoter was found to be activated by sucrose in phloem cells of transgenic tobacco plants bearing the rolC promoter-uidA chimeric fusion gene (Yokoyama et al., 1994). The highest GUS activity was recorded when 400 mM sucrose was exogenously supplied to young seedlings, and a time course analysis showed that the activity was highest after 40 - 48 h incubation. The cis-acting, sucrose-responsive region of the rolC promoter was shown to be located between positions -135 and -94 (Yokoyama et al., 1994). An AT-rich palindromic sequence in this sucrose-responsive region was found to be homologous to the cis-acting, sucrose-responsive region of the class I patatin gene, suggesting that the AT-rich sequence may be important for sucrose-responsiveness (Yokoyama et al., 1994).

The site of initiation of transcription of the rolC gene is located 27 bp upstream from the first ATG codon. A CCAAT-box and TATA-box were found at -57 and -35, respectively (Kanaya et al., 1990). DNA sequences responsible for the phloem-specificity of the rolC promoter were found within the -153 region, whereas a minimum region needed for the expression in the seed embryo was located at position -120 (Sugaya and Uchimiya, 1992). In the same study an enhancer element with consensus sequence CAAACCAC was found between -298 and -230, while an AT-box (AATATTTTTATT) which is a consensus sequence in the promoters of light-regulated genes, was found at -77 to -66 (Kanaya et al., 1990; Sugaya and Uchimiya,
In transgenic carrot cells containing the rolC-GUS fusion gene, suppression of the rolC promoter was seen when embryos were cultured in the presence of 2,4-dichlorophenoxyacetic acid. It was assumed that the developmental switch leading to the onset of embryogenesis could be responsible for the activation of the rolC promoter (Fujii and Uchimiya, 1991).

The regulatory functions of the rolB and rolC genes were shown to be conserved in the homologous genes (Ngrol) of Nicotiana glauca in tobacco genetic tumors (Nagata et al., 1995). The rolC promoter has also been used in studies where phloem-specific expression of genes were required. In a study involving the investigation of the role of inorganic phosphate in phloem metabolism, the ppa gene from E.coli was expressed using the rolC promoter (Lerchl et al., 1995). In another study, the rolC promoter was used together with the phloem-specific Sh promoter of maize sucrose synthase-1 to engineer potato leafroll virus (PLRV)-resistant plants (Graham et al., 1997). The authors showed that the rolC promoter can be used to create PLRV-resistant line of potato and that it is likely to be effective for engineering resistance to other phloem-limited viruses. It was, however, not useful for engineering resistance to viruses that replicate in non-vascular tissues (Reimann-Phillip and Beachy, 1993).

2.4.2.3 Reporter genes

The expression of a foreign gene by transformed plant cells can best be assessed by determining the abundance or activity of its product. The coding sequences of genes for bacterial enzymes, which are easily assayed and whose activity is not normally found in plants, form the basis of many reporter genes (Webb and Morris, 1992).

Reporter genes are used for the assessment of promoter expression characteristics and the analysis of subcellular compartmentation studies. An ideal plant reporter gene should have a number of inherent characteristics, including that its product should be unique and should not be toxic to the host plant; the marker enzyme should exhibit a high degree of post-translational stability; a convenient and sensitive
enzyme assay for the reporter gene product should be available; and it should be
amenable to translational fusions with extraneous polypeptides while retaining
enzymatic activity (Jefferson et al., 1987).

Examples of reporter genes are the uidA gene product of E.coli (β-glucuronidase or
GUS) (Jefferson et al., 1987), the lacZ gene product of E.coli (β-galactosidase or lac)
(Herrera-Estrella et al., 1990), chloramphenicol acyltransferase (cat) (Boutry et al.,
1987), neomycin phosphotransferase (nptII) (McDonnell et al., 1987) and luciferase
(luc) (Riggs and Chrispeels, 1987). Recently the jellyfish Aequorea victoria green
fluorescent protein (GFP) has become a popular reporter gene due to its non-
invasive detection method (Hu and Cheng, 1995). The two reporter genes that are
used in this study are GUS and GFP, and these are discussed below.

i) β-glucuronidase (GUS)
The E.coli β-Glucuronidase (GUS, EC 3.2.1.31) encoded by the uidA locus is a
hydrolase that catalyses the cleavage of a wide variety of β-glucuronides, many of
which are available as spectrophotometric, fluorometric and histochemical substrates
(Jefferson, 1987). The GUS gene has been cloned and sequenced, and encodes a
stable enzyme that has desirable properties for the construction and analysis of gene
fusions (Jefferson et al., 1987). The blue staining indicative of GUS activity, using 5-
bromo-4-chloro-3-indolyl-β-glucuronic acid (X-Gluc) as substrate in transgenic plants,
is usually easy to localise and gives visually striking results (Beeckman and Engler,
1994). It is not the product of the glucuronidase action on X-Gluc that is colored;
instead the indoxyl derivative produced must undergo an oxidative dimerisation to
form an insoluble and highly coloured indigo dye. The dimerisation is stimulated by
atmospheric oxygen and can be greatly enhanced by using an oxidation catalyst
such as potassium ferricyanide/ferrocyanide mixture. The relative degree of staining
may not necessarily reflect the concentrations of β-Glucuronidase (Draper et al.,
1988). The GUS fusion system has successfully been used in numerous studies
involving the development of transformation protocols and the analysis of promoter
activity.
In many cases, as in whole-mount preparations and thick hand-cut sections, the pigments present in most plant tissues can partially or even totally mask the sites of GUS activity and consequently hinder the detection of it (Beeckman and Engler, 1994). A technique to clear all pigments, and even the brown phenolics that are formed due to oxidative polymerisation, has been developed. The clearing agent contains a 2:1:1 mixture of chloral hydrate, lactic acid and phenol, and the tissue is cleared after staining in X-Gluc solution (Beeckman and Engler, 1994).

The commonly used histochemical protocols for the detection of GUS activity are destructive, as they involve fixation, vacuum infiltration and prolonged incubation of the plant material in solutions containing X-Gluc, detergents, and organic solvents. A non-destructive assay system for roots using short-term incubations in X-Gluc and the spraying of 4-methylumbelliferyl-β-glucuronide (4-MUG) on the leaves was developed to follow both the quantitative and qualitative expression of chimeric GUS genes in most plant tissues without destroying the plants (Martin et al., 1992).

ii) Green fluorescent protein (GFP)

Green fluorescent protein from the Pacific jellyfish (cnidarian) Aequorea victoria is a 27 kD monomer that absorbs violet/blue light ($\lambda_{max}$=395 nm with a minor peak at 470nm) and fluoresces green ($\lambda_{max}$=508 nm) (Hu and Cheng, 1995). The GFP chromophore is formed through intramolecular polypeptide cyclisation and oxidation (Heim et al., 1994) and is derived from the primary amino acid sequence of the protein by post-translational modification (Cody et al., 1993).

The structure and fluorescence mechanism (Youvan and Michel-Beyerle, 1996) as well as the physical and chemical properties of GFP (Yang et al., 1996) have been investigated and it was found that the full length GFP, consisting of 238 amino acids, is required for fluorescence. The minimal chromophore responsible for light emission can, however, be assigned to a hexapeptide within the protein. This region contains a Ser-Tyr-Gly trimer at amino acid positions 65 to 67, which cyclises and yields a fluorophore that emits light (Cody et al., 1993). The overall shape of the protein...
consists of eleven strands of β-sheet forming the shape of a cylinder, capped with α-helices on the top and bottom. The α-helices form a scaffold for the chromophore, which is located near the geometric centre of the protein. This structure represents a protein fold nicknamed the β-can (Yang et al., 1996).

The GFP is rather thermosensitive with the yield of fluorescently active to total GFP protein decreasing at temperatures greater than 30°C (Yang et al., 1996). Besides light, no exogenous substrate or cofactor is required for fluorescence and GFP is extremely resistant to photobleaching, allowing detection over extended periods of time (Chalfie et al., 1994). It maintains fluorescence even after prolonged incubation in strong denaturing agents such as 6 M guanidine HCl, 8 M urea or 1% (w/v) sodium dodecyl sulphate (SDS) (Bokman and Ward, 1981). GFP also has a broad range of pH stability, retaining conformation from pH 5.5 - 12.0 (Bokman and Ward, 1981).

GFP is expressed in both prokaryotic as well as eukaryotic cells (Kain et al., 1995), and has been used as a reporter gene in a variety of systems including E.coli, Caenorhabditis elegans (Chalfie et al., 1994), mosquitoes (Higgs et al., 1996), Drosophila melanogaster (Wang and Hazelrigg, 1994; Brand, 1995), mammalian cells (Moss and Rosenthal, 1994; Pines, 1995; Garamszegi et al., 1997), Dictyostelium (Hodgkinson, 1995), fungi (Flach et al., 1994) and plants (Haseloff and Amos, 1995, Baulcombe et al., 1995; Niedz et al., 1995; Köhler et al., 1997; Nagatani et al., 1997). In a recent review on the applications of green fluorescent protein in plants, this reporter gene was described as the first truly in vivo reporter system useful in whole plants (Leffel et al., 1997)

There are two major uses for GFP in plants: monitoring gene expression and protein localisation at high resolution, and providing an easily scored genetic marker in living plants (Haseloff and Amos, 1995). To employ GFP successfully in plants, the GFP apoprotein must be produced in suitable amounts in the plant cells, it must undergo efficient post-translational cyclisation and oxidation to produce the mature GFP, and the fluorescent protein may need to be suitably targeted within the cell to allow
efficient post-translational processing, safe accumulation to high levels, or to facilitate
detection of expressing cells (Haseloff and Amos, 1995). It proved to be difficult to
regenerate fertile Arabidopsis plants from the brightest GFP transformants, indicating
the possibility that high levels of GFP are mildly toxic or interfere with regeneration,
perhaps due to the fluorescent or catalytic properties of the protein (Haseloff and
Amos, 1995).

Several modified versions of GFP have been developed to improve fluorescence. In
Arabidopsis thaliana the expression of the wild-type GFP cDNA was curtailed by
aberrant splicing, with an 84 nucleotide intron being excised efficiently from within the
GFP coding sequence between nucleotides 400 and 483 (Haseloff et al., 1997). A
modified version of the GFP has been constructed with altered codon usage, to
mutate the cryptic splice sites and to decrease the AU content of the mRNA (Haseloff
et al., 1997). This modified GFP was named mGFP4 and has been shown to result
in increased fluorescence when compared directly to the wild-type GFP in
Arabidopsis (Haseloff et al., 1997) and conifer tissues (Tian et al., 1997). Similar
modifications were made to the GFP coding sequence and this modified protein,
synGFP, also exhibited improved fluorescence in plants (Rouwendal et al., 1997).
The wild-type GFP has, however, been successfully expressed in Citrus sinensis
protoplasts (Niedz et al., 1995) and in maize protoplasts (Hu and Cheng, 1995),
suggesting that the cryptic intron may not be recognised with equal efficiency in
different plant species or perhaps during transient expression studies (Haseloff and
Amos, 1995).

A synthetic GFP was also constructed with improved codon usage and by recursive
cycles of DNA shuffling, a whole cell fluorescence signal that was 45-fold greater
than the standard was obtained (Crameri et al., 1996). In another investigation, a
synthetic GFP gene was constructed to improve GFP expression in plants (Pang et
al., 1996). Replacement of the serine at position 65 with a threonine (S65Tpgfp) or
a cysteine (S65Cpgfp) yielded 100- to 120-fold brighter fluorescence than wild-type
GFP upon excitation with 490 nm light (Pang et al., 1996; Reichel et al., 1996).
Incorporation of a plant intron into the coding region yielded an additional 1.4-fold improvement (Pang et al., 1996). When the tyrosine at position 66 was exchanged with a histidine (Y66H), a blue fluorescence was observed (Reichel et al., 1996).

Several practical problems present themselves in visualising stably transformed plants for GFP and engineering a transgenic tracking system. The question whether GFP fluorescence is stable across generations was addressed by Leffel and coworkers (1997). Over 200 independent lines of GFP-containing transgenic tobacco was produced and the analysis of approximately 20 of these showed that the transgene was stably expressed. Fourth generation plants were still fluorescent (Leffel et al., 1997). Protein analysis was conducted on transgenic plants to quantify GFP expression and to determine how much GFP must be produced for it to be unambiguously visualised in whole plants. The results indicated that approximately 0.1% mGFP4 expression is sufficient to yield highly fluorescent plants, and that fluorescence is positively correlated with protein expression (P>0.05)(Leffel et al., 1997). A third important aspect that was looked at was the relationship between light quality and light quantity that plants use for growth and fluorescence. They found that irradiance can cause an apparent switch on/switch off of fluorescence. Under prolonged periods of cloudiness, previously highly fluorescent plants were scored as low-to-no fluorescence (Leffel et al., 1997). Field experiments will have to be performed to test whether GFP decreases plant growth, seed yield and photosynthesis.

### 2.4.3 Transformation of sugarcane

Sugarcane (Saccharum spp. hybrids, Gramineae) was the first grass species to be grown in tissue culture (Sun et al., 1993). Attempts to transform sugarcane using Agrobacterium tumefaciens have not been successful so far. Stably transformed sugarcane callus has been regenerated following electroporation and polyethylene glycol (PEG) treatment of protoplasts (Chen et al., 1987) but the regeneration of plants from protoplasts has been extremely difficult (Chen et al., 1988). The method of choice for the transformation of sugarcane is particle bombardment, and
suspension cultures (Franks and Birch, 1991; Sun et al., 1993), embryogenic callus (Bower and Birch, 1992) and meristems (Gambley et al., 1993) have been used as target tissues. The progress in molecular biology, and advances in the transformation of sugarcane have recently been reviewed (Birch et al., 1996; Moore, 1996).

The success of microprojectile bombardment-mediated transformation of sugarcane callus is variety dependent, and the most critical parameter is the ability of the variety to produce embryogenic callus (Bower and Birch, 1992). Twelve sugarcane varieties grown commercially in South Africa were compared for both their ability to form white, embryogenic callus and their transformation potential (Snyman et al., 1996). The plasmid pDP687 (Pioneer Hi-Bred International), which contains two anthocyanin transcriptional activators, each under control of a double CaMV 35S promoter, was used to monitor transient expression after particle bombardment. Varieties N11, N21, N22, N23 and N24 produced little or no white, embryogenic callus while NCo310, NCo376, N12, N16 and N20 produced the highest proportion of white callus. Of these, NCo310 gave the best results. The highest transient anthocyanin production was observed in varieties NCo310 and Nco376 (Snyman et al., 1996).

The relative expression strengths obtained from various promoter regions used in the transformation of sugarcane cells have been determined (Rathus et al., 1993; Gallo-Meagher and Irvine, 1993). The cultivar NCo310 was used in a study to determine the effects of tissue type and promoter strength on transient GUS expression following particle bombardment (Gallo-Meagher and Irvine, 1993). Four promoters were tested in this study and the constructs including the GUS reporter gene were as follows: 1) pAHC27 (Christensen et al., 1992) containing the promoter, first exon and first intron of the maize ubiquitin 1 gene (Ubi-1); 2) pEmuGN (Last et al., 1991) containing the recombinant Emu promoter consisting of six copies of the anaerobic responsive element of the maize Adh-1 gene, four copies of the Agrobacterium tumefaciens octopine synthase gene enhancer element, and a truncated maize Adh1 promoter plus first intron; 3) pAct1-D (Zhang et al., 1991) containing the promoter, first exon, first intron and a portion of the second exon of the rice actin 1 gene (Act1)
in pBlueScript and 4) pBI221, a derivative of pUC19 containing the cauliflower mosaic virus 35S promoter (35S) (Jefferson et al., 1987). The physical parameters used in this study had previously been optimised for the transformation of sugarcane. These included helium pressure regulator setting, target tissue distance, microprojectile type and size, DNA precipitation method and DNA load per bombardment. The results indicated that the Ubi-1-GUS construct (pAHC27) gave the highest mean number of GUS foci which was approximately 2, 4 and 12 times higher than Emu-GUS (pEmuGN), Actl-GUS (pAct1-D) and 35S-GUS (pBI221) respectively. The size and intensity of the blue foci were also different with each construct, and when fluorometric GUS analysis was done, the same expression pattern was observed (Gallo-Meagher and Irvine, 1993).

The effects of promoter, intron and enhancer elements on transient gene expression in sugarcane and carrot protoplasts have been investigated (Rathus et al., 1993). Several constructs were made containing the nopaline synthase promoter (nos), the CaMV 35S promoter (35S), the ocs enhancer, and intron 1 from maize Adh1, arranged upstream of the GUS-coding region. Levels of expression in electroporated sugarcane and carrot protoplasts were compared. nos and 35S are weak but approximately equivalent strength promoters in sugarcane and result in 100-fold to 500-fold lower expression in sugarcane than in carrot. When two 35S promoters were used back to back, they resulted in almost doubled expression in sugarcane, but caused no change in carrot. The inclusion of a 180 bp fragment including the ocs enhancer immediately upstream of 35S increased expression 5-fold in sugarcane and 2-fold in carrot. Inclusion of intron 1 from maize Adh1 between the promoter and coding region decreased expression from both nos and 35S in sugarcane. The Emu promoter, with multiple enhancers close to the core promoter region, is more strongly enhanced relative to 35S in sugarcane (400-fold) than in cereals (10- to 50-fold). This promoter is equivalent in strength to 35S in carrot (Rathus et al., 1993).

A lot of emphasis is currently being placed on the identification of gene regulatory elements which provide tissue and developmental specificity in sugarcane. Because
the first few sugarcane promoters have been silenced upon reintroduction by
transformation, it is very important to extend the range of promoters tested for activity
with heterologous genes following transformation (Potier and Birch, 1997a). These
regulatory elements will be utilised in the transformation of sugarcane with the aims
of improving crop traits and breeding higher yielding or better quality varieties (Moore,
1996).

In South Africa the major reasons for the production of transgenic sugarcane include
disease and pest resistance and the increase of sucrose yield. Damage caused to
cane plants by pests and diseases have a big effect on cane yield, with subsequent
economical implications. The industry is also looking toward the manipulation of
sucrose metabolism in order to increase sucrose production in the plant.

A major area of the sugarcane plant where transgene expression is required is the
culm. This is not only the area where the stalkborer, Eldana saccharina, enters the
plant and lives, but is also the region of the plant where sucrose is stored. Regulation
of specific expression of transgenes in the culm will require culm-specific or culm-
preferential promoter elements. The possible value of using regulatory elements
which direct expression specifically to the vascular bundles has become apparent
from a study of the anatomy of the sugarcane culm (Jacobsen et al., 1992).

Sugarcane culm is composed of storage parenchyma tissue permeated by numerous
vascular bundles. These bundles are surrounded by a sheath of two or more layers
of thick-walled, lignified sclerenchyma cells (Jacobsen et al., 1992). About 1 500
vascular bundles are scattered throughout sugarcane internodes. The bundles on the
periphery of the culm are both smaller and more densely packed than those closer
to the centre so that 50% of the total number are within the outer 1 mm and 75% are
within the outer 3 mm of the culm (Jacobsen et al., 1992).

The complicated arrangement of vascular bundles in the monocotyledonous stem,
and especially in the Gramineae, is associated with the position of the various
bundles of each leaf in the stem (Fahn, 1990). Three types of leaf trace bundles are present in the culm. The first type is represented by those bundles derived from midribs and the lateral veins. These penetrate deep into the interior of the stem and then pass to the periphery in lower nodes. The second type is represented by leaf trace bundles derived from smaller lateral veins which, immediately on entering the culm, occupy the position of the outermost peripheral bundles with which they fuse. The third type is represented by bundles that are very thin and which soon fade out within the cortex close to the level of the node. In the Gramineae there are special horizontal bundles at the nodes which connect the leaf bundles together (Fahn, 1990).

A consequence of the arrangement of the sugarcane vascular system is that products of transgenes that are expressed only in the vascular bundles will be concentrated on the periphery of the culm and only in the “veins” of the leaves. This is particularly advantageous for targeting aphids and leafhoppers, which are phloem feeders, viruses that replicate in the vascular bundles, and stalkborers. Because of the tissue-specificity, the expression of a foreign gene will not take place at the expense of other processes in the plant.

Genes that are expressed in a tissue-specific manner in sugarcane have to be identified. One candidate gene was identified in a study of sucrose metabolism where the enzyme activity patterns in the sink tissues of the sugarcane plant were investigated (Whittaker and Botha, 1996). This gene, pyrophosphate-dependent phosphofructokinase (PFP), has not previously been isolated from sugarcane.

2.5 PYROPHOSPHATE-DEPENDENT PHOSPHOFRUCTOKINASE (PFP)
Two enzymes in the cytosol of plant cells catalyse the phosphorylation of fructose 6-phosphate to fructose 1,6-bisphosphate. The first of these is ATP-dependent phosphofructokinase (PFK, EC 2.7.1.11) which utilises ATP, catalyses an irreversible reaction, and is found in the majority of organisms (for a review see Dennis and Greyson, 1987). The second, PFP (EC 2.7.1.90), catalyses the reversible
interconversion between fructose 6-phosphate (Fru-6-P) and fructose 1,6-bisphosphate (Fru-1,6-P₂), using pyrophosphate (PPi) as a phosphoryl donor in the forward reaction and inorganic phosphate (Pi) as a phosphoryl acceptor in the reverse direction:

\[
\text{Fru-6-P + PPi} \rightarrow \text{Fru-1,6-P₂ + Pi}
\]

A pH optimum of 7.5 - 8.0 and magnesium is essential for the interconversion reaction. PFP activity has been investigated in a number of plant species, for example pineapple (Carnal and Black, 1979), *Citrellis lanatus* (Botha and Botha, 1991), *Phaseolus vulgaris* (Botha and Small, 1987), *Brassica nigra* (Duff et al., 1989), potatoes (Kruger and Dennis, 1987), soybean (Macdonald and Preiss, 1986), carrots (Wong et al., 1990) and wheat (Yan and Tao, 1984), as well as in the specialised microorganisms *Entamoeba histolytica* (Reeves et al., 1976) and *Propionibacterium shermanii* (O'Brien et al., 1975).

In most of the plants examined, PFP is composed of two immunologically unrelated subunits (Kruger and Dennis, 1987), namely the larger α-subunit (65 - 68 kD) and the smaller β-subunit (60 - 64 kD). In wheat tissue the PFP protein is present as two isoforms, a heterotetramer (α₂β₂) and a homodimer (β₂) (Yan and Tao, 1984). In carrot (Wong et al., 1988) and *Citrellis lanatus* (Botha and Botha, 1991), isoforms were detected with either an α- and β-subunit or only a β-subunit composition. The amount of α- and β-subunit also varies in different plant tissues (Botha and Botha, 1991) and in some cases, correlates with the phosphate status of the tissue (Theodorou and Plaxton, 1992). Although there is an approximate 60% homology between the deduced amino acid sequence of the two subunits, they are immunologically distinct (Carlisle et al., 1990).

The sequences of cDNA clones indicate that the β-subunit is catalytic, since it has putative Fru-6-P/Fru-1,6-P₂ binding sites (Yan and Tao, 1984) and the catalytic residues, Asp-127 and Arg-171, were shown to be conserved in this subunit (Carlisle et al., 1990). The α-subunit may be involved in the regulation of the enzyme by binding the allosteric activator fructose 2,6-bisphosphate (Fru-2,6-P₂)(Cheng and
Tao, 1990). Fru-2,6-P₂ activates PFP in both the forward (glycolytic) and reverse (gluconeogenic) directions by increasing $V_{\text{max}}$ and increasing the affinity for Fru-6-P, Fru-1,6-P₂ and PPI. Fru-2,6-P₂ also relieves the inhibition of PFP by high levels of PPI (Sabularse and Anderson, 1981).

Although the kinetic properties of PFP have been studied in detail, the physiological role of the enzyme is still unknown. PFP is suggested to be involved in both the regulation of sucrose accumulation and the biosynthetic activity in plant cells (Dennis and Greyson, 1987; Huber and Akazawa, 1986; Nakamura et al., 1992). Possible functions that have been suggested for PFP include the regulation of cytosolic PPI (ap Rees et al., 1985), maintenance of the equilibrium of the hexose/triose phosphate pools (Dennis and Greyson, 1987), and an adenylate bypass for glycolysis during periods of phosphate limitation or starvation (Duff et al., 1989).

The regulation of the cytosolic PPI levels by PFP is consistent with the involvement of the enzyme during sucrose cleavage (Dancer and ap Rees, 1989). In certain tissues engaged in starch biosynthesis, PFP is proposed to generate PPI as an energy source to stimulate the PPI-dependent sucrose cleavage pathway by sucrose synthase and uridine 5'-diphosphate glucose (UDPGlc) pyrophosphorylase (Dancer and ap Rees, 1989). PFP and sucrose synthase activity have been suggested to play a role in controlling sink strength (Black et al., 1995), which can be defined as the ability of the sink organ to import assimilate (Ho et al., 1989). Sucrose may act as a signal molecule to induce expression of both sucrose synthase and PFP (Black et al., 1995). It has been concluded that PFP is an adaptive enzyme with kinetic and molecular properties that are responsive to a variety of environmental, developmental and tissue-specific cues (Plaxton, 1996).

Full-length cDNA clones for the α- and β-subunits of PFP have been isolated from a cDNA expression library derived from potato tuber mRNA (Carlisle et al., 1990). The nucleotide sequences of the subunits were determined and when compared, it was shown that the subunits were related with approximately 40% of the amino acids
being identical. It was suggested that the genes for these polypeptides may have been formed by gene duplication (Carlisle et al., 1990). However, the genes for the two subunits are complex and different with the α- and β-genes containing 19 and 16 exons, respectively, and only two introns showing alignment (Todd et al., 1995). Near full-length cDNA clones for both subunits from a developing castor seed endosperm cDNA library have also been isolated and these were similar in sequence to those described for potato (Blakeley et al., 1992).

2.5.1 PFP in sugarcane

Sucrose accumulation in sugarcane is suggested to be primarily regulated within the translocation system and/or the level of the sink (Moore, 1995). Significant differences in the levels of sucrose and fiber occur among the Saccharum species. In sugarcane sink tissue there is likely to be substantial competition for incoming sucrose for storage and biosynthetic activity, including structural polysaccharide (fiber) synthesis (Whittaker and Botha, 1996).

In a recent study to investigate whether PFP is linked to sucrose accumulation and fiber content in sugarcane, the activity patterns of PFP were measured in the top ten internodes of five different cane varieties (Whittaker and Botha, 1996). PFP activity increased from internode 3, reaching optimal activity at internode 6 and 7, after which the enzyme declined to internode 10. An inverse correlation between peak PFP activity/expression levels and the prevailing sucrose content of internodal tissue, and a positive relationship between PFP activity and the insoluble matter content was observed (Whittaker and Botha, 1996). Differences in PFP activity both within the developing sugarcane stalk and between different varieties, were reflected by PFP protein concentrations.

To investigate the distribution of sucrose, fiber and PFP activity within an internode, representative sections of an internode were cut into three concentric rings (each constituting 30% of the total area), from the core to the periphery. PFP activity correlated positively with the increase in fiber deposition from the core to the
periphery. There was an inverse trend in sucrose accumulation, which correlated negatively with PFP activity (Whittaker and Botha, 1996).
CHAPTER 3
USE OF A DIFFERENTIAL DISPLAY TECHNIQUE TO ISOLATE DIFFERENTIALLY EXPRESSED GENE SEQUENCES FROM SUGARCANE CULM TISSUE

3.1 ABSTRACT

An mRNA differential display technique was evaluated for the identification and isolation of differentially expressed gene sequences in the sugarcane culm. RNA was isolated from leafroll, leaf, mature culm and young culm tissues and reverse transcribed to cDNA. The two original methods developed by Liang and Pardee (1992) and Welsh et al. (1992) were combined and modified in this study. After reverse transcription polymerase chain reaction (RT-PCR), amplification products were separated in agarose gels and the fragments that appeared to be unique to the culm tissues were isolated and reamplified. An RT-PCR Southern blot technique was used to verify the tissue specificity of isolated fragments, and to determine whether those fragments were members of multigene families or if they were present in several isoforms. Fragments that were possibly culm-preferential were cloned and their sequences determined.

A series of 120 random decamer primers was used to obtain 1,767 fragments, resulting in an average of 15 fragments per primer. Thirty five (2%) of these fragments were possible culm-specific sequences. Twenty one of the fragments were characterised using a hybridisation technique and four of these were thus identified as putative culm-preferential rather than culm-specific fragments. The complete sequences of three of the four fragments and a partial sequence of the other one were determined, and no significant homology to known differentially expressed sequences in the international databases were found. One of the fragments, SA11, was analysed further by RT-PCR using longer sequence-specific primers. Amplification products obtained from cDNA and genomic DNA templates using these primers were of identical size. Results of a series of control reactions showed that the
synthesis of the SA11 fragment from RNA was reverse transcription-dependent, and was not a product of genomic DNA contamination. Although the fragment appeared to be culm-preferential, it could not be used to obtain a culm-specific promoter. The other three fragments were not analysed further because, as in the case of SA11, no conclusive evidence that these fragments were indeed culm-preferential could be obtained from either the northern blots or the sequencing analysis. Further modifications and refinements of the mRNA differential display technique have to be made before it can be routinely applied to sugarcane.

3.2 INTRODUCTION
The development of techniques to manipulate specific DNA fragments has led to advances in the ability to selectively alter the genetic make-up of plants and other organisms. Physiological function can be investigated by altering gene expression (ap Rees, 1995), and genes that can confer resistance to pests (Peferoen, 1992), diseases (Sturtevant and Beachy, 1993; Shah et al., 1995) or environmental conditions (Vierling and Kimpel, 1992) can be introduced. In addition, crop quality can also be improved for specific purposes, e.g. taste and processing characteristics (Hiatt, 1993), color (Fraley, 1992; Mol et al., 1995), shelf life (Oeller et al., 1991), etc. Many useful modifications involve the alteration of the amount of desirable or undesirable compounds by up- or down-regulating the gene expression. For example, the insertion of an antisense form of the polygalacturonase (PG) gene into tomato fruit, led to a reduction in the PG levels and an improvement of the post-harvest fruit quality, due largely to enhanced firmness and increased resistance to post-harvest fungal disease (Hiatt, 1993; Smith et al., 1993). Plants can also be used for the production of novel compounds (for reviews see Groenewald et al., 1995; Goddijn and Pen, 1995) or for the overproduction or bioconversion of natural compounds in the plant (Pen et al., 1993).

The most important aspect of the genetic manipulation of plants, is the isolation of relevant genes and regulatory elements (Fraley and Schell, 1992; Gibson and Somerville, 1993). By using appropriate promoter elements, the desired response of
the transgenes in the recipient plant to environmental stimuli, or the synthesis of the product in a specific tissue can be ensured. Furthermore, a strong promoter is the prerequisite of a high transcription frequency (Walden and Wingender, 1995).

For some applications the constitutive expression of transgenes will be required, while in other cases the use of an inducible or tissue-specific promoter sequence would be more advantageous. Targeted gene expression is currently being used in studies involving the elucidation of specific metabolic pathways. For example, the phloem-specific expression of the \textit{E.coli} inorganic pyrophosphatase gene led to the removal of pyrophosphate (PPI) from the phloem and the elucidation of the role of PPI in phloem metabolism (Lerchl \textit{et al.}, 1995).

The first step in the search for tissue-specific promoters is the identification and isolation of differentially expressed genes. General approaches that are followed to identify and isolate gene sequences that are differentially expressed, i.e. in a regionally, temporally or environmentally specific way, include "+/-" screening and subtraction approaches (Sargent, 1987). These methods have several disadvantages which have been discussed in Chapter 2.

The differential display technique was developed to identify and characterise differentially expressed genes by detecting individual mRNA species, and then isolating and analysing the cDNA (Liang and Pardee, 1992; Liang \textit{et al.}, 1993). The method uses the polymerase chain reaction (PCR) to amplify even rare cDNAs, making the identification of differentially expressed genes of low abundance possible. This method can replace subtractive hybridisation techniques because it is less time consuming, and allows the simultaneous handling of numerous samples. Comparison of RNA samples from different cells allows the identification and cloning of differentially expressed genes. A similar technique utilising a single randomly selected primer at low stringency for first and second strand cDNA synthesis, was also developed (Welsh \textit{et al.}, 1992).
The differential display technique was originally developed for the detection of altered gene expression in mammalian cell populations (Liang and Pardee, 1992). Numerous refinements and optimisations have since been made (Liang et al., 1993; Bauer et al., 1993; Callard et al., 1994; Hadman et al., 1995; Liu and Shearn, 1995; Zhao et al., 1995; Doss, 1996), and these have been discussed in Chapter 2. The method has been used in different studies with various degrees of success. Applications range from the analysis of gene expression in pre-implantation mouse embryos (Zimmermann and Schultz, 1994) to the isolation of senescence-specific genes from Arabidopsis plants (Callard et al., 1994).

In sugarcane, the culm is the major area of sucrose accumulation, and it is also the point of entry for the stalkborer, Eldana saccharina. The culm is therefore one of the main target areas for transgene expression in both borer resistance programmes as well as metabolic manipulation studies in South Africa. After positive results were obtained by other groups using the differential display technique, this method was investigated as a tool for the isolation of culm-specific genes from sugarcane. The aim was to simplify the method and yet still identify as many culm-specific cDNA sequences as possible, and then to use them to isolate the genes and subsequently the promoters.

The working hypothesis was as follows: the use of a simple oligo(dT)$_{15}$ primer instead of a group of anchored dT-primers to prime reverse transcription, will reduce the number of reverse transcription reactions from four to one, but will still target the poly(A)-tails of the mRNA. Secondly, since no sequence information of sugarcane genes is available, the cDNA can be amplified using the reverse transcription primer together with random primers. Thirdly, the separation of the amplification products in agarose gels instead of denaturing polyacrylamide gels can make the method easier and faster. Although less fragments can be resolved in agarose gels, more primer combinations can be used to overcome this drawback. The use of this method should thus allow the rapid screening of cDNA populations in sugarcane with large numbers of primers.
3.3 MATERIALS AND METHODS

3.3.1 RNA Preparation

Total RNA was isolated from mature (6 to 12 month old) field-grown sugarcane plants (variety N19) using a modified acid guanidinium thiocyanate-phenol-chloroform extraction method (Chomczynski and Sacchi, 1987). In sugarcane, the internode attached to the leaf with the uppermost visible dewlap was defined as internode no. 1, according to the system of Kuijper (Van Dillewijn, 1952). All solutions were prepared with water that had been treated with 0.1% (v/v) diethyl pyrocarbonate (DEPC). Glassware was autoclaved and baked at 200°C prior to use, and plastic ware was used directly from sterile unopened packs. Five grams of tissue from leafroll, leaf, young culm (internode 2) or mature culm (internode 7) was ground in liquid nitrogen and homogenised in 5 ml denaturing solution (4 M guanidinium thiocyanate; 25 mM sodium citrate, pH7.0; 0.5% (w/v) sarkosyl; 100 mM β-mercaptoethanol). After the addition of 0.5 ml of 2 M sodium acetate (pH 4.0), 5 ml phenol and 1 ml chloroform:isoamylalcohol (24:1), the suspension was shaken vigorously and left on ice for 15 min. After centrifugation at 10 000 g for 20 min at 4°C, the aqueous phase was removed and the RNA was precipitated with an equal volume of isopropanol at -20°C for at least 1 h. RNA was collected by centrifugation at 10 000 g for 20 min at 0°C. The pellet was dissolved in 1.5 ml denaturing solution and the precipitation was repeated, after which the resulting pellet was washed with 70% ethanol (room temperature) and dried under vacuum. The RNA was resuspended in 250 µl DEPC-treated water at 65°C for 15 min. Isolated RNA was quantified spectrophotometrically and the integrity was verified by examining the 18S and 28S ribosomal fragments after separation in a formaldehyde-containing 1.5% (w/v) agarose gel. RNA preparations were stored at -80°C.

3.3.2 Reverse transcription

cDNA was synthesised in 20 µl volumes containing 1 µg total RNA, 20 pmol oligo(dT)$_{15}$ primer, 10 units placental ribonuclease inhibitor, 0.1 mM of each deoxynucleoside triphosphate (dNTP) and 4 units M-MuLV Reverse Transcriptase (Boehringer Mannheim). The reaction mixture was incubated at 37°C for 1 h.
3.3.3 Amplification of cDNA

Prior to amplification the cDNA was purified through a QIAquick-spin PCR purification column (QIAGEN). Single random decamer primers were used for PCR (Operon Technologies). The PCR reaction mixture contained 5 μl of cDNA template, 10 mM Tris-Cl, pH 8.3; 10 mM KCl; 3.8 mM MgCl₂; 0.2 μM random primer; 0.1 mM of each dNTP and 1 unit of AmpliTaq Stoffel fragment (Perkin Elmer) in a volume of 15 μl. The samples were overlaid with 30 μl mineral oil and subjected to 41 cycles of PCR using the following cycling parameters: 1 cycle of 94°C for 3 min, 35°C for 1 min and 72°C for 2 min with a 2.4°C s⁻¹ ramp; 40 cycles of 94°C for 1 min, 35°C for 1 min and 72°C for 2 min using a 2.4°C s⁻¹ ramp. The last cycle was followed by a 5 min extension at 72°C. PCR amplification products were separated in 2% (w/v) agarose gels in 1x TAE buffer (40 mM Tris-acetate; 1 mM EDTA, pH 8.0). After staining with ethidium bromide (0.01 mg ml⁻¹), the gels were photographed and the banding patterns analysed for the presence of tissue-specific fragments.

3.3.4 Preparation of molecular weight markers for agarose gel electrophoresis

The numbers of the following molecular weight markers are those numbers that were assigned by Boehringer Mannheim.

**Marker 2:** 20 μg of bacteriophage lambda DNA was treated with 20 units of the restriction endonuclease Hind III for 3 h at 37°C in a 100 μl reaction volume. A volume of 200 μl TE buffer (10 mM Tris-Cl; 1 mM EDTA, pH 8.0) was added to the restricted product, together with 30 μl loading buffer (40% (v/v) glycerol; 1 mM EDTA; 0.02% (w/v) bromophenol blue; 0.02% (w/v) xylene cyanol). Ten microlitres of this mixture was loaded per lane of the agarose gels. The sizes of the fragments obtained after electrophoresis in a 0.8% (w/v) agarose gel are shown in Figure 3.1(A).

**Marker 3:** 20 μg of bacteriophage lambda DNA was treated with 20 units of each of the restriction endonucleases EcoRI and Hind III for 3 h at 37°C in a 100 μl reaction volume. The mixture was then treated the same as Marker 2. The sizes of the fragments obtained after electrophoresis in a 0.8% (w/v) agarose gel are shown in Figure 3.1(B).

**Marker 5:** 20 μg of pBR322 plasmid DNA was treated with the restriction
endonuclease Hae III for 3 h at 37°C in a reaction volume of 100 µl. The mixture was then treated the same as Marker 2. The sizes of the fragments obtained after electrophoresis in a 2% (w/v) agarose gel are shown in Figure 3.1(C).

![Figure 3.1](image)

**Figure 3.1** Molecular weight markers used in this study. A.) Marker 2. Fragments obtained after lambda DNA was treated with Hind III and separated in a 0.8% (w/v) agarose gel. B.) Marker 3. Fragments obtained after lambda DNA was treated with EcoRI and Hind III and separated in a 0.8% (w/v) agarose gel. C.) Marker 5. Fragments obtained after pBR322 DNA was treated with Hae III and separated in a 2% (w/v) agarose gel. Sizes of fragments are indicated in base pairs (bp).

3.3.5 **Isolation and reamplification of possible tissue-specific fragments**

Fragments that appeared to be unique to the culm tissue were punched out of the gel using a Pasteur pipette, and were placed in 100 µl TE buffer. Samples were incubated at 65°C for 10 min, vortexed and stored at 4°C. For reamplification of the fragments, 5 µl of this solution was used. PCR conditions were as described before (3.3.3) and after reamplification, the size of the product was verified by loading 1 µl of the reaction mixture in a 2% (w/v) agarose gel. In cases where more than one reamplification product was observed, the fragment of expected size was isolated from the gel using the QIAquick Gel Extraction Kit (QIAGEN). These isolated fragments were stored at -20°C.
3.3.6 Characterisation of isolated fragments by RT-PCR Southern blotting
The amplification products in the differential display gels showing putative tissue-specific fragments were transferred to a nylon membrane (MagnaGraph, MSI). A downward capillary blotting technique as described by Chomczynski and Mackey (1994) was used. After overnight transfer the damp membrane was exposed to ultraviolet (UV) light for 2.5 min at 120 mJ cm\(^{-1}\) to link the DNA to the nylon. The membrane was then probed with the specific isolated fragment, which was radioactively labelled using a random primer labelling kit (Prime-It II; Stratagene) in the presence of [\(\alpha\text{-}^{32}\text{P}\)]dCTP (Amersham; 3000 Ci mmol\(^{-1}\)). Prehybridisation and hybridisation were performed in a solution containing 5x SSC (150 mM NaCl; 15 mM tri-sodium citrate, pH 6.8), 50% (v/v) formamide, 10% (w/v) dextran sulphate, 0.5% (w/v) SDS, 5x Denhardt's solution (0.02% (w/v) of each bovine serum albumin (BSA), polyvinylpyrrolidone (PVP) and ficoll), and 100 \(\mu\text{g}\) ml\(^{-1}\) sheared herring sperm DNA. Prehybridisation of 4 h at 55°C was followed by overnight hybridisation at 55°C after the addition of the probe that was denatured by boiling for 5 min. The membrane was washed twice at room temperature for 20 min and once at 50°C for 20 min in a solution containing 2x SSC and 0.1% (w/v) SDS. This was followed by a 15 min wash at 50°C in 0.2x SSC and 0.1% (w/v) SDS. The membrane was exposed to X-ray film for 24 to 48 h.

3.3.7 Cloning of cDNA fragments
Because of the presence of the same primer on both ends of the fragment, sequencing could not be done directly using the random decamer primer. The fragments were therefore cloned using the pCR-Script (SK+) cloning kit (Promega). The manufacturer's instructions were followed, except for the use of \(E.\text{coli}\) JM83 cells instead of \(Epicurian\text{ coli}\) supercompetent cells for transformation. Clones were identified using blue/white selection on X-Gal-containing Luria plates (1% (w/v) NaCl; 1% (w/v) tryptone; 0.5% (w/v) yeast extract; 1.5% (w/v) agar). Insert sizes were determined by PCR amplification of the inserted fragments using the vector-specific T7 promoter and pUC18/M13 reverse primers.
3.3.8 Sequence determination

The fmol™ DNA sequencing system (Promega) was used together with the vector-specific T7 promoter primer and the M13/pUC18 reverse primer. A cycle sequencing protocol using direct incorporation of [α-^35S]dATP (Amersham) was followed, after which the sequencing products were separated in 6% (w/v) denaturing polyacrylamide gels. The gels were dried and exposed to X-ray film for 24 to 48 h. Sequences were read manually from the film.

3.3.9 Genomic Southern blot analysis

Genomic DNA from sugarcane variety N19 (extracted using the method of Honeycutt et al., 1992) was treated separately with each of the restriction endonucleases BamHI, EcoRI and Hind III. The three restriction reactions were incubated for 24 h at 37°C using 20 μg of DNA and 60 units of each enzyme in a 200 μl reaction volume. After ethanol precipitation, the samples were dried and resuspended in 20 μl TE buffer. Electrophoresis in a 0.8% (w/v) agarose gel was followed by ethidium bromide staining (0.01 mg ml⁻¹) to confirm the satisfactory restriction of the DNA. The restricted DNA was then transferred to a nylon membrane (MagnaGraph; MSI). Downward capillary blotting was done as described (3.3.6), and the DNA was linked to the nylon membrane by exposing it to UV light for 2.5 min at 120 mJ cm⁻². As a positive control for the labelling of the culm-preferential cDNA fragments, 0.5 ng of each fragment was also included in the gel.

Twenty five nanogram of each fragment was radioactively labelled using the Prime-It II random primer labelling kit (Stratagene) in the presence of [α-^32P]dCTP (Amersham; 3000 Ci mmol⁻¹). After 4 h prehybridisation in RapidHyb buffer (Amersham) at 55°C, the probe was denatured for 5 min in boiling water and added to the prehybridisation buffer. Hybridisation was done for 16 h at 55°C, after which the membrane was washed twice at 55°C for 20 min in a solution containing 2x SSC and 0.1% (w/v) SDS, once at 60°C for 15 min in a solution containing 2x SSC and 0.1% (w/v) SDS and once at 60°C for 20 min in a solution containing 0.2x SSC and 0.1% (w/v) SDS. The membrane was sealed in a plastic bag and exposed to X-ray
film for 48 h. After the film was developed, the radioactive probe was stripped from the membrane by washing in a solution containing 50% (v/v) formamide and 2x SSC at 65°C for 1 h. The membrane was rinsed in 0.1x SSC before it was prehybridised again and probed with the next fragment. The membrane was kept moist between hybridisations and was stripped three times before being discarded.

3.3.10 Northern blot analysis of isolated cDNA fragments
Twenty microgram of total RNA from leafroll, leaf, young culm (internode 2) and mature culm (internode 7) was separated in a 1.2% (w/v) agarose gel. Samples were prepared by adding 50% (v/v) formamide, 6% (v/v) formaldehyde and 1x MOPS buffer (200 mM 3-[N-Morpholino]propanesulphonic acid (MOPS); 50 mM NaOAc; 5 mM EDTA) and heating to 65°C for 10 minutes before loading. One sample was loaded separately together with the 0.24 - 9.5 kb RNA Ladder (GibcoBRL), which was treated the same as the samples. After separation in the gel, these two lanes were cut off and stained in ethidium bromide (0.1 mg ml⁻¹). The sizes of the fragments were verified using the RNA Ladder. The rest of the samples were transferred overnight to a positively charged nylon membrane (Boehringer Mannheim) by downward capillary blotting in 10x SSC. The RNA was linked to the nylon membrane by exposing it to UV light for 2.5 min at 120 mJ cm⁻².

Prehybridisation was done in a solution containing 250 mM phosphate buffer (pH 7.0), 7% SDS and 1 mM EDTA at 65°C for at least 4 h. Labelling of fragments was done as described in 3.3.6, and the specific activities ranged from 1.2 x 10⁹ to 2.3 x 10⁹ cpm µg⁻¹. Hybridisation followed in the same buffer for 20 h at 65°C. The membrane was washed twice at 65°C in a solution containing 0.2x SSC and 0.1% (w/v) SDS for 20 min after which it was sealed in a plastic bag and exposed to X-ray film. The film was developed after 24 h, but a signal could only be observed after 6 days. Radioactive probes were stripped from the nylon membrane by washing 3 times in a solution containing 0.1x SSC, 1% (w/v) SDS and 40 mM Tris-HCl (pH 7.5) at 80°C.
3.3.11 Characterisation of the culm-preferential fragment, SA11, by specific amplification

Sequence specific amplification primers were designed to amplify the SA11 fragment. The primers were both 20-mers with a GC-content of 50 -55%: SA11-F: 5'-ATC GCC GTC AAT GAG CAC TC-3' and SA11-R: 5'-AGA GGT TGT AGT GAC ATC GG-3'. Specific amplification reactions were done in 15 μl volumes using 1 μM of each primer. Different templates were used in these reactions and were as follows: cDNA from internode 7 tissue: 5 μl of purified reverse transcription mixture as described in 3.3.3, RNA from leafroll and internode 7 tissues: 1 μg of total RNA, genomic DNA from sugarcane varieties NCo310 and 84F3097: 100 ng, and the cloned SA11 fragment as positive control: 15 ng. The rest of the PCR mixture was the same as described in 3.3.3, except that it contained 2 units of Taq DNA Polymerase (Boehringer Mannheim) instead of the AmpliTaq Stoffel fragment. The cycling parameters were as follows: 1 cycle of 94°C for 2 min, 39 cycles of 94°C for 1 min, 55°C for 2 min and 72°C for 2 min. The last cycle was followed by a 7 min extension at 72°C. Amplification products were separated in 2% (w/v) agarose gels.

3.3.12 Characterisation of the culm-preferential fragment, SA11, by specific RT-PCR

The GeneAmp RNA PCR Kit (Perkin Elmer) was used for the specific RT-PCR of SA11. The reverse transcription reaction was done at 42°C for 15 min in a volume of 10 μl containing 5 mM MgCl₂, 1x PCR buffer II (supplied in the kit), 1 mM of each dNTP, 2.5 units of RNase inhibitor, 2.5 μM oligo(dT)₁₆ primer or 0.75 μM of the specific primer, 1 μg of total RNA and 5 units of M-MuLV Reverse Transcriptase, and was overlaid with 30 μl of mineral oil. The reaction was terminated by heating to 99°C for 5 minutes and then cooling to 5°C. The amplification reaction mixture was added directly to the cDNA after reverse transcription and contained 2 mM MgCl₂, 1x PCR buffer II, 2.5 units of AmpliTaq DNA Polymerase and 0.15 μM of each primer. This mixture was spun through the oil of the RT mixture and was subjected to 42 cycles of PCR using the following cycling parameters: 1 cycle of 94°C for 2 min, 40 cycles of 94°C for 30 s, 45°C for 45 s and 72°C for 3 min. The last cycle was followed by
an 8 min extension at 72°C. The amplification products were separated in 2% (w/v) agarose gels.

3.4 RESULTS

3.4.1 Optimisation of PCR template concentration

The RNA isolation technique that was used yielded intact RNA averaging between 40 μg g\(^{-1}\) fresh mass (fm) tissue from the mature culm to 150 μg g\(^{-1}\) fm tissue from the leaf. Initially, aliquots of 2.5 μl and 5.0 μl of the reverse transcription reaction mixture were used in amplification reactions, and the oligo(dT)
\(_{15}\) primer was used together with a random decamer primer.

Only smears could be seen after the amplification reactions. The possibility that the template concentration was too high was investigated by using a dilution series of the reverse transcription reaction mixture ranging from 1:5 to 1:1 000 (results not shown). When amplification of these dilutions also yielded smeared profiles, the interference of the reverse transcription primer was investigated. The oligo(dT)
\(_{15}\) primer was removed by the purification of the reverse transcription reaction mixture through an affinity column. Initially, undiluted purified cDNA was used in amplification reactions, but only a few fragments were formed (results not shown). The possibility of an excess of template was investigated. A dilution series was prepared from mature culm (internode 7) cDNA ranging from 1:50 to 1:5 000 to determine the optimal concentration for amplification, and the amplification was done using a single random decamer primer, OPF12. Four reactions were done with each template concentration.
to verify reproducibility. In all the amplification reactions, a volume of 5 µl of the template was used. The banding profiles are shown in Figure 3.3.

![Figure 3.3](image)

**Figure 3.3** Optimisation of the PCR template concentration for differential display. A dilution series of the reverse transcribed cDNA from internode 7 was prepared. Four identical samples were amplified from each concentration to test the reproducibility of the amplification profiles. Primer OPF12 was used in these reactions. Lanes 1 and 18: molecular weight marker 3, lanes 2 - 5: cDNA dilution 1:50, lanes 6 - 9: cDNA dilution 1:500, lanes 10 - 13: cDNA dilution 1:1000, lanes 14 - 17: cDNA dilution 1:5000.

The only reproducible banding pattern was produced when the 1:50 dilution (Figure 3.3 lanes 2 - 5) was used. At lower template concentrations (higher dilutions) the banding patterns were highly variable between supposedly identical samples, and therefore these template concentrations could not be used for differential display. A second dilution series was prepared for finer optimisation of template concentration, and this ranged from 1:50 to 1:500. These dilutions were amplified using primer OPF06. Four identical samples were used from each dilution to ensure reproducibility (Figure 3.4).

A reproducible banding pattern was produced when the 1:50 (Figure 3.4 lanes 2 - 5) and 1:100 (Figure 3.4 lanes 6 - 9) dilutions were used. With lower template concentrations, the banding patterns were once again very variable. The 1:100 dilution of the purified reverse transcription reaction mixture, in which 1 µg of total RNA was reverse transcribed, was used as template in the series of amplifications that followed. The concentration of the cDNA was not determined.
Figure 3.4 Optimisation of the PCR template concentration for differential display. A dilution series of the reverse transcribed cDNA from internode 7 was prepared. Four identical samples of each dilution were amplified to test the reproducibility of the amplification profiles. In these reactions primer OPF06 was used. Lanes 1 and 18: Molecular weight marker 3, lanes 2 - 5: cDNA dilution 1:50, lanes 6 - 9: cDNA dilution 1:100, lanes 10 - 13: cDNA dilution 1:250, lanes 14 - 17: cDNA dilution 1:500.

3.4.2 Differential display: the comparison of random amplified fragments from different sugarcane tissues

A series of 120 random decamer primers was used in amplification reactions and cDNA fragments from leafroll, leaf, young culm (internode 2) and mature culm (internode 7) were compared. A total of 1767 fragments were amplified, resulting in an average of 15 fragments per primer. The size of amplified fragments ranged from 100 to 2000 bp. Thirty five (2%) of these fragments were identified as possible culm-specific fragments. Results obtained with some of the primers, for example OPF18 in Figure 3.5 (A), showed the presence of many polymorphisms, even in the two internode 7 samples, while with others, for example OPB10 in Figure 3.5 (B), the banding pattern was identical in all tissues.

To minimise false positives due to the synthesis of spurious fragments, two identical internode 7 samples were included in each reaction as suggested by Liang et al. (1993). This was not only used to verify banding profile reproducibility, but it also facilitated the characterisation of the isolated fragments (discussed in 3.4.3).
Figure 3.5  Amplification profiles of four sugarcane tissues using different random primers. A.) Amplification profiles obtained using primer OPF18 where numerous polymorphisms were detected. Some polymorphic fragments are indicated with arrows. B.) Amplification profiles obtained using primer OPB10 where no polymorphisms were detected. Lanes 1: molecular weight marker 3, lanes 2: leafroll, lanes 3: leaf, lanes 4: young culm (internode 2), lanes 5 and 6: identical samples from mature culm (internode 7).

Since this method relies on the amplification of DNA, very low concentrations of contaminating DNA can skew the display of amplified expressed fragments. After each RNA isolation that was done, a control reaction was included in the first set of amplifications where the RNA was amplified without prior reverse transcription. This ensured that any genomic DNA contamination that might be present in the template would be detected. No amplification products were ever seen after these reactions (results not shown). Further confirmation that genomic DNA was not amplified during the differential display reactions was obtained by comparing the banding profiles to those obtained from random amplification of polymorphic DNA (RAPD) analysis of sugarcane genomic DNA using the same random primer. These profiles were very different (results not shown).

3.4.3 Characterisation of isolated fragments using RT-PCR Southern blots
To determine whether the specific fragments that were isolated were present in more than one isoform or whether they possibly belonged to multi-gene families, the fragments were characterised by hybridisation analysis. Twenty one of the possible
35 culm-specific fragments were characterised using a RT-PCR Southern blot technique. These fragments were chosen because they could be successfully reamplified or could be labelled directly without prior reamplification. Fragments that did not reamplify were not characterised further. After a seemingly unique fragment was identified from the culm tissue, it was removed from the gel and used as a probe. A range of different results was obtained. Some fragments that were not tissue-specific hybridised strongly to the same size fragments in all tissues, while others hybridised to fragments of different sizes in the same and other tissues indicating the possibility of the presence of isoforms or multi-gene families. In four cases the isolated fragment hybridised to the expected size fragment in the culm and a weaker signal could be seen in other tissues. These fragments were termed culm-preferential. None of the characterised fragments were culm-specific.

The four culm-preferential fragments were designated 2A08 (667 bp), SA11 (640 bp), 7C16 (589 bp) and 7F18 (~1100 bp). The first part of the name identifies the tissue from which the fragments were isolated, i.e. internode 2, internode 7, or both of the culm tissues(S), while the second part indicates the primer that was used to amplify the fragment, e.g. OPA08. The differential display gel as well as the hybridisation result after RT-PCR Southern blot analysis of each of the culm-preferential fragments are shown in Figures 3.6 to 3.9.

3.4.4 Sequence determination of the cloned fragments
The complete sequences of 2A08, SA11 and 7C16 and a partial sequence of 7F18 were determined. Due to limitations in the length of sequences that could be read from polyacrylamide gels, only a partial sequence could be obtained of 7F18. Sequences are shown in Figures 3.10 to 3.13.
Figure 3.6 Identification and characterisation of culm-preferential fragment 2A08. A.) Amplification profile of sugarcane tissues produced by primer OPA08. The 667 bp fragment (2A08) isolated from lane 4 is indicated with an arrow. B.) Southern blot of the amplified fragments in (A) probed with isolated fragment 2A08. The hybridisation of the fragment to the same size fragments in the culm lanes (4, 5 and 6) is indicated with an arrow. Lane 1: molecular weight marker 3, lanes 2: leafroll, lanes 3: leaf, lanes 4: young culm (internode 2), lanes 5 and 6: mature culm (internode 7).

Figure 3.7 Identification and characterisation of culm-preferential fragment SA11. A.) Amplification profile of sugarcane tissues produced by primer OPA11. The 640 bp fragment (SA11) visible in lanes 4, 5 and 6 is indicated with an arrow. B.) Southern blot of the amplified fragments in (A) probed with isolated fragment SA11. The preferential hybridisation of the fragment to the same size fragments in the culm lanes (4, 5 and 6) is indicated with an arrow. Lane 1: molecular weight marker 3, lanes 2: leafroll, lanes 3: leaf, lanes 4: young culm (internode 2), lanes 5 and 6: mature culm (internode 7).
Figure 3.8 Identification and characterisation of culm-preferential fragment 7C16. A.) Amplification profile of sugarcane tissues produced by primer OPC16. The 589 bp fragment (7C16) isolated from lane 6 is indicated with an arrow. B.) Southern blot of the amplified fragments in (A) probed with isolated fragment 7C16. The hybridisation of the fragment to the same size fragments in the culm lanes (4, 5 and 6) is indicated with an arrow. Lane 1: molecular weight marker 3, lanes 2: leaf roll, lanes 3: leaf, lanes 4: young culm (internode 2), lanes 5 and 6: mature culm (internode 7).

Figure 3.9 Identification and characterisation of culm-preferential fragment 7F18. A.) Amplification profile of sugarcane tissues produced by primer OPF18. The 1100 bp fragment (7F18) visible in lanes 5 and 6 is indicated with an arrow. B.) Southern blot of the amplified fragments in (A) probed with isolated fragment 7F18. The specific hybridisation of the fragment to the same size fragment in one of the two mature culm lanes (6) is indicated with an arrow. Lane 1: molecular weight marker 3, lanes 2: leaf roll, lanes 3: leaf, lanes 4: young culm (internode 2), lanes 5 and 6: mature culm (internode 7).
Figure 3.10 Complete nucleotide sequence of cloned fragment 2A08. Sequences determined using T7 and pUC18/M13 reverse primers were overlapped and a 667 bp sequence was obtained. Bases that could not be clearly read from the sequencing gel are indicated with an N.

Figure 3.11 Complete nucleotide sequence of fragment SA11. Sequences determined using T7 and pUC18/M13 reverse primers were overlapped and a 641 bp sequence was obtained. The specific primers, SA11-F on the 5' end and SA11-R on the 3' end, are underlined in the sequence.
Figure 3.12 Complete nucleotide sequence of fragment 7C16. Sequences determined using T7 and pUC18/M13 reverse primers were overlapped. A sequence of 589 bp was obtained.

5' CACACTCCAGCATTATCATGATTAAAGGTAAAGATCATTGTCATATCAGCTACTGCACAAAACTGCACGTTTCTCAGAGCTGAGCTGCTTCTCAGACATTCACAAAAGCACCAGGTAACCGTGCGACCGATCCTACACCTGAACTTTTGAAAGAGAGAGGGACCTGGAGTGTG 3' 

Figure 3.13 Partial nucleotide sequence of fragment 7F18. Sequences were determined using T7 (394 bp) and pUC18/M13 reverse primers (673 bp). A sequence of ~100 bp is missing (--.--).

The sequences were compared to both the non-redundant (NR) nucleotide sequence as well as the expressed sequence tag database (dbEST) by BLASTN, and to the deduced amino acid database by BLASTX (Altschul et al., 1990). No significant homologies were found, and only very short pieces of the fragments were aligned in the analyses by the basic local alignment search tool (BLAST) programmes, as can be seen in Table 3.1.
Table 3.1 Summary of results obtained from sequence similarity BLAST searches of the culm-preferential fragments. Only the highest percentage homology results obtained with an "expected" value of 5.0 are shown in the table. The latest results that were obtained on 8/1/98 are included in the table.

<table>
<thead>
<tr>
<th>Sequence</th>
<th>%</th>
<th>Length</th>
<th>Sequence</th>
<th>%</th>
<th>Length</th>
</tr>
</thead>
<tbody>
<tr>
<td>BLASTN NR</td>
<td>1. Human BAC clone RG072E11 from 7q21-7q22</td>
<td>73</td>
<td>49 bp</td>
<td>1. S. cerevisiae chromosome XIV reading frame ORF YNR075w</td>
<td>75</td>
</tr>
<tr>
<td>BLASTX</td>
<td>2. Methanococcus small extrachromosomal element</td>
<td>72</td>
<td>58 bp</td>
<td>2. P. falciparum mRNA for AARP1 protein</td>
<td>64</td>
</tr>
<tr>
<td>BLASTN dbEST</td>
<td>3. Yeast DNA cut14 protein</td>
<td>69</td>
<td>56 bp</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BLASTX</td>
<td>None</td>
<td></td>
<td></td>
<td>1. Orgyia pseudotsugata hypothetical protein</td>
<td>35</td>
</tr>
<tr>
<td>BLASTN dbEST</td>
<td>None</td>
<td></td>
<td></td>
<td>2. Nicotiana tabacum DNA-binding protein</td>
<td>34</td>
</tr>
</tbody>
</table>

Sequence: The sequences in the databases to which the highest percentage homology were recorded. %: The percentage similarity recorded for the specific fragment that was compared in the search. Length: The number of base pairs (bp) or amino acids (aa) that were compared in the BLAST search. NR: Non-redundant databases of Genbank and EMBL sequences. EST: Non-redundant database of Genbank EST division.

3.4.5 Genomic Southern blot analysis of culm-preferential fragments

Southern blot analysis of all 4 fragments showed positive signals. The number of bands observed ranged from 2 to 11, and the results are summarised in Table 3.2. With each of the 4 fragments, the positive control gave a very strong signal which confirmed that the probe labelling was satisfactory.
Table 3.2 Characterisation of culm-preferential fragments by genomic Southern blot analysis.

<table>
<thead>
<tr>
<th></th>
<th>2A08</th>
<th>SA11</th>
<th>7C16</th>
<th>7F18</th>
</tr>
</thead>
<tbody>
<tr>
<td>BamHI</td>
<td>6</td>
<td>6</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>EcoRI</td>
<td>4</td>
<td>4</td>
<td>2</td>
<td>11</td>
</tr>
<tr>
<td>Hind III</td>
<td>10</td>
<td>11</td>
<td>6</td>
<td>7</td>
</tr>
</tbody>
</table>

3.4.6 Northern blot analysis of culm-preferential fragments

No clear signals were obtained from the northern blot analysis. Even after an exposure time of 6 days, only the following faint signals were observed: 2A08: no signal whatsoever was visible on the film; SA11: faint bands at 650 bp in the leafroll and internode 7 lanes were visible in smears; 7C16: 2 faint bands at 1400 and 900 bp, respectively, were visible in the leaf, internode 2 and internode 7 lanes; 7F18: a very faint band at approximately 1 500 bp was seen in smears in the leafroll and internode 7 lanes. No strong culm-specific signals were obtained (results not shown).

3.4.7 Characterisation of the culm-preferential fragment, SA11

Based on the result of the northern blot, and the fact that it was the first fragment from which the complete sequence was obtained, the SA11 fragment was characterised further. Sequence-specific 20-mer primers were designed for the amplification of this fragment. The sequence and location of these primers are indicated in the sequence of SA11 (Figure 3.11) The forward primer, SA11-F, contained eight of the ten bases of the OPA11 random primer on the 5' end. Because of sequence ambiguity on the 3' end of the SA11 sequence, the reverse primer, SA11-R, did not contain any part of the random primer sequence. PCR using the 20-mer primers was done using cDNA from internode 7 and a genomic DNA template. The results are shown in Figure 3.14. The amplification products from both templates were of identical size (Figure 3.14 lanes 2 and 8).
Figure 3.14 Specific amplification of fragment SA11 from cDNA and genomic DNA. A concentration series of internode 7 cDNA was prepared after the RT mixture was purified. Lane 1: molecular weight marker 3, lane 2: cDNA undiluted, lane 3: cDNA 1:10 dilution, lane 4: cDNA 1:50 dilution, lane 5: cDNA 1:100 dilution, lane 6: cDNA 1:250 dilution, lane 7: cDNA 1:500 dilution, lane 8: 120 ng genomic DNA (variety N19).

The presence of genomic DNA contamination in the RNA samples was investigated by an amplification reaction where RNA from leafroll and internode 7 tissues, without prior reverse transcription, was used as template. Genomic DNA from two sugarcane varieties, NCo310 and 84F3097, was also amplified and the cloned SA11 fragment was included as positive control for the reaction. The two SA11-specific primers were used for the amplification of all these templates. Results are shown in Figure 3.15.

Figure 3.15 Determining the origin of fragment SA11. RNA from leafroll (lane 2) and internode 7 (lane 3) was amplified without prior reverse transcription. Genomic DNA from variety NCo310 (lane 4) and 84F3097 (lane 5) was amplified using the same primers. The cloned SA11 fragment was used as a positive amplification control (lane 6). Molecular weight marker 3 (lane 1) was used to determine the size of the amplified products.

No amplification products were formed from the RNA templates (Figure 3.15 lanes 2 and 3), but a 640 bp fragment was amplified from both the genomic DNA samples (Figure 3.15 lanes 4 and 5). In variety NCo310 (lane 4) a high concentration of the fragment was amplified, but in variety 84F3097 (lane 5), only a very faint product could be seen in the gel. In the positive control reaction (lane 6), the expected 640 bp fragment was amplified.

The dependence of the synthesis of SA11 on the reverse transcription reaction was investigated by a series of RT-PCR reactions.
Firstly, reverse transcription of RNA from leafroll and internode 7 tissues was primed with either the forward specific primer, SA11-F, the reverse specific primer, SA11-R, or the oligo(dT)$_{16}$ primer. Amplification of the reverse transcribed cDNA was done using the SA11-F and SA11-R primers, and the results are shown in Figure 3.16(A). In all six reactions, the 640 bp SA11 fragment was synthesised. The only difference that was observed was that the fragments were amplified at different concentrations. Secondly, the RT-PCR was done using either SA11-R or SA11-F primers to prime reverse transcription, but no reverse transcriptase enzyme was added to the reaction. Amplification was then done using SA11-F and SA11-R primers and the results that were obtained are shown in Figure 3.16(B). No amplification products were formed. In the third RT-PCR, no primers were used in the reverse transcription reaction, but after amplification with SA11-F and SA11-R, the 640 bp SA11 fragment had been synthesised as can be seen in Figure 3.16(C).

**Figure 3.16** Dependence of the formation of fragment SA11 on reverse transcription. A.) Reverse transcription of RNA from leafroll (lanes 2 - 4) and internode 7 (lanes 5 - 7) was primed with 3 different primers: SA11-F (lanes 2 and 5), SA11-R (lanes 3 and 6) and oligo(dT)$_{16}$ (lanes 4 and 7). Amplification of the reverse transcribed cDNA was done with SA11-F and SA11-R in all cases. B.) Reverse transcription of RNA from leafroll using primers SA11-F (lane 2) and SA11-R (lane 3) without the addition of the reverse transcriptase enzyme. Subsequent amplification was done using SA11-F and SA11-R. C.) Reverse transcription of RNA from leafroll using reverse transcriptase without the addition of a primer. Amplification was done using SA11-F and SA11-R (lane 2). In each case, the sizes of amplified fragments were determined with molecular weight marker 3 (lane 1).
3.5 DISCUSSION

The differential display methods of Liang and Pardee (1992) and Welsh et al. (1992) were combined and simplified to generate discrete cDNA bands from different sugarcane tissues with ease. In each step of the procedure the conditions had to be optimised for use with sugarcane.

Instead of the sets of anchored dT primers that were used in the original method, a simple oligo(dT)$_{15}$ was used for reverse transcription to produce cDNA in a single reaction. Random hexanucleotides have sometimes been used for reverse transcription to ensure that the amplified products were not all situated on the 3' end of the mRNA (Sokolov and Prockop, 1994). However, an oligo(dT)$_{15}$ primer targets mRNA and was therefore used in this study to reduce the possibility of obtaining ribosomal and transfer RNA. The influence of this oligo(dT)$_{15}$ primer in the amplification reaction following reverse transcription did, however, have to be determined.

Where the oligo(dT)$_{15}$ as well as a random decamer primer were used for the amplification after reverse transcription, smears were observed in the gel, as have also been observed by other groups (Liang et al., 1993; Hadman et al., 1995). This can be due to the oligo(dT)$_{15}$ primer, without anchoring 3' bases, annealing to any part of the poly-A tail of a given mRNA. Subsequent reverse transcription will thus result in a series of amplified cDNA products of different sizes from the same mRNA. Supporting evidence for this was obtained from the experiment where the oligo(dT)$_{15}$ primer was removed on completion of the reverse transcription reaction, and amplification was done using a single random decamer primer. Discrete amplification products were observed in the gel.

The importance of template concentration to obtain reproducible amplification profiles when using the differential display technique was investigated. The purified template cDNA was initially used without prior dilution, and only a few fragments could be seen in the gels. This may have been due to an excess of template in the amplification
reactions. Results obtained after several dilutions of the template cDNA had been used in PCR confirmed this speculation, and a 1:100 dilution of the cDNA gave the best reproducible banding pattern. The observation reported by Liang and co-workers (Liang et al., 1993) that some fragments were expressed at different concentrations or were differentially amplified, was also made in this study. The conclusions that can be made from this part of the investigation are that cDNA template concentration for RT-PCR is very important for the generation of reproducible amplification profiles and furthermore, the differential display method is not quantitative under the PCR conditions that are used.

Results obtained from RT-PCR using different random primers proved to be variable for the numbers of amplification products, the concentration of fragments amplified, the number of polymorphisms, and the clarity of fragments in the agarose gel after staining. An average of 15 clear fragments per primer were detectable in an agarose gel after RT-PCR, which is comparable to the 10 to 20 amplified fragments that were detected by Welsh et al. (1992), and the 15 to 24 fragments that were amplified by Sokolov and Prockop (1994). Two percent of the 1767 fragments that were scored appeared to be culm-specific. This is higher than the 1% differentially expressed fragments that were amplified by Liang et al. (1993). This may be due to the fact that they compared different developmental stages of the same tissue, and more variation can be expected when different tissues are compared.

A different approach to the original method (Liang and Pardee, 1992) was followed for the analysis of the differential display profiles after RT-PCR. Instead of denaturing polyacrylamide gels to separate the amplification products after PCR, agarose gels and ethidium bromide staining was used, thereby eliminating the need for radioactivity. Although modified, nonradioactive differential display techniques have been described since this study was started, and silver staining of the acrylamide gels can be done (Lohmann et al., 1995; Doss, 1996), agarose gelelectrophoresis combined with ethidium bromide staining has the advantages of being quicker and simpler to perform. Although the resolution of agarose gels is lower than that of
polyacrylamide gels, the number of primers that were used for the amplification of cDNA fragments was increased to overcome this drawback. The use of agarose gels also facilitated the characterisation of isolated fragments since the complete amplification profile could be transferred to a nylon membrane for Southern blot analysis.

The RT-PCR Southern blot technique was developed to allow the immediate characterisation of isolated fragments that appeared to be unique to mature culm tissue. Apart from cases where a fragment that appeared to be culm-specific was actually present in the other tissues, but was not amplified from that particular cDNA sample because of incomplete cDNA synthesis or priming mismatches (Bauer et al., 1993), this hybridisation technique eliminated obvious false positives. From the different results that were obtained from the Southern blot analyses, it was clear that not all the amplified fragments could be visualised in an agarose gel. In some cases, fragments of the same size as the isolated fragment were seen in the other tissues after hybridisation. These fragments were, however, not visible in the gel. In other cases the probe hybridised to fragments of different sizes in the same and in the other tissues. This could mean that the primer found multiple priming sites on the same mRNA, which can be explained by the fact that mismatches in the priming site were tolerated at the low stringency PCR conditions that were used (annealing temperature of 35°C and MgCl₂ concentration of 3.8 mM) (Bauer et al., 1993). Another possible explanation for this is that those mRNAs could be present in multiple isoforms that may be tissue-specific, or that they may be members of a multi-gene family.

In contrast to other groups that reported some heterogeneity in the purified PCR products obtained by differential display (Bauer et al., 1993; Callard et al., 1994), none of the four culm-preferential fragments that were cloned and sequenced in the present study contained contaminating sequences. Although the results of RT-PCR Southern blot analyses indicated that four of the isolated fragments were culm-preferential, northern blot analyses of fragments 7C16 and 7F18 revealed that they
were expressed in other tissues as well. Fragment SA11 showed a faint signal in leafroll and a slightly stronger signal in internode 7, which may indicate that the fragment could be culm-preferential. The northern blot results also indicated that these fragments were not abundant since the signals were very low, even after using 20 μg of total RNA per sample lane, and probe specific activities of $10^9$ cpm μg$^{-1}$ as suggested by Zebrowski et al. (1994). No signal was observed on the blot probed with fragment 2A08. Similar results have been reported by Liang et al. (1993) where 6 of 15 isolated fragments failed to give a signal with northern blotting. Five of these fragments were confirmed to be tissue-specific and the rest were false positives.

Sequence analysis of the fragments showed that no significant homology to differentially expressed sequences in the international database could be found. Only three plant genes showed a degree of homology to these sequences. The database search results may indicate that the isolated fragments are novel sequences that have not yet been characterised. Genomic Southern blot analysis showed that these fragments are indeed part of the sugarcane genome. Judging from the number of bands that were observed after hybridisation, these fragments could either be present as multiple copies in the genome, or located in repetitive DNA regions.

Specific characterisation of the SA11 fragment led to the conclusion that the formation of the 640 bp product was dependent on the reverse transcription reaction, but independent of the primers used in the reaction. The fragment was also not a result of genomic DNA contamination in the RNA samples. This phenomenon could not be explained.

Results of northern blot and sequence analyses indicated that the isolated fragments were not tissue-specific and could thus not be used in this study to obtain a culm-specific promoter sequence.
3.6 CONCLUSIONS

The original differential display method was simplified to facilitate the identification and isolation of differentially expressed gene sequences in four different sugarcane tissues. Although the results that were obtained compared favourably to other groups who have successfully used the same technique, the fragments that were isolated proved to be false positives. The modified differential display technique could not be used successfully in this study to obtain differentially expressed gene sequences from sugarcane mature culm tissue.

During the course of this study, several modifications of the differential display technique were reported in the literature. The following suggestions to minimise false positives were summarised and may be useful in future optimisation of this technique for use in sugarcane:

- The use of long composite primers and higher annealing temperatures during PCR may increase the sensitivity of the technique (Zhao et al., 1995).
- A combination of mRNA isolation and DNase I treatment of RNA prior to reverse transcription can reduce the level of contaminating DNA to a minimum (Doss, 1996). Although this was not a problem in the present study, it may be included as part of the method when other tissues are analysed.
- Reverse transcription primed with a simple oligo(dT)$_n$ primer and subsequent amplification reactions using anchored dT primers together with random decamers, can make the technique more specific while still maintaining its ease of operation (Doss, 1996).
CHAPTER 4

EXPRESSION OF REPORTER GENES UNDER CONTROL OF THE rolC PROMOTER IN SUGARCANE

4.1 ABSTRACT

The expression patterns of GUS and GFP reporter genes driven by the Agrobacterium rhizogenes rolC promoter were investigated for the first time in transgenic sugarcane culm tissue. The constructs were introduced into sugarcane callus via particle bombardment and stable transformants were selected on the basis of resistance to geneticin, conferred by the cotransferred plasmid. Seven GUS transformants and four GFP transformants were analysed. Results of DNA and RNA analyses confirmed the stable integration of the foreign DNA into the genome and mRNA transcription. Histochemical analysis revealed that reporter gene expression was localised in the vascular bundles and specifically in the phloem cells. The expression pattern in the culm was directly comparable to the sucrose gradient in mature sugarcane plants. Expression was higher in the mature parts of the plant and increased from the peripheral region towards to core of the culm.

These results demonstrate that, as in the cases of tobacco, rice and hybrid aspen, the rolC promoter drives phloem-specific expression in sugarcane. Furthermore, it can be concluded that sucrose may activate the phloem-specific expression of the rolC promoter in sugarcane plants.

4.2 INTRODUCTION

The rolA, B and C genes that are located on the TL-DNA regions of the Ri plasmid of Agrobacterium rhizogenes (Chilton et al., 1982; Cardarelli et al., 1987) are involved in the pathogenesis of the hairy-root syndrome upon infection in most dicotyledonous plants. Several investigators have shown that these genes establish phenotypic alterations characteristic for each gene when expressed separately in transgenic tobacco plants (Oono et al., 1987; Schmülling et al., 1988; Sinkar et al., 1988).
The intergenic region separating the rolB and rolC genes was found to represent a bidirectional promoter that regulates transcription in a similar fashion for both genes in the aerial organs of the plant, but in different ways in the roots. In the leaves expression is mainly limited to the vascular system. In the roots rolB is active mainly in the meristematic regions, while rolC is active in the phloem (Schmülling et al., 1988). Expression of the rolC promoter has been found to be phloem-specific in transgenic tobacco plants (Schmülling et al., 1988), rice (Matsuki et al., 1989) and hybrid aspen (Nilsson et al., 1996).

In an investigation to find the signals related to cell-specific and developmentally regulated expression of the rolC gene, the effects of various external stresses and molecules on the rolC promoter were analysed (Yokoyama et al., 1994). The results indicated that the rolC promoter is activated by sucrose in phloem cells of transgenic tobacco seedlings bearing the rolC promoter-uidA chimeric fusion gene (Yokoyama et al., 1994).

The rolC promoter has already been used in studies where phloem-specific expression of genes were required. In a study involving the investigation of the role of inorganic phosphate in phloem metabolism, the ppa gene from E.coli was expressed using the rolC promoter (Lerchl et al., 1995). The rolC promoter was also used together with the phloem-specific Sh promoter of maize sucrose synthase-1 to engineer potato leafroll virus (PLRV)-resistant plants (Graham et al., 1997).

For the present study it was postulated that the rolC promoter will allow phloem-specific expression in sugarcane. Since the vascular bundles in the sugarcane culm are concentrated mainly on the periphery, the use of such a promoter will ensure preferential expression, with significant enrichment of the transgene product in the outer peripheral area of the culm. This could be of great importance in transgenic plants where phloem-feeders such as aphids and leafhoppers that carry viruses, and borer insects are targeted.
Phloem-specific expression is also advantageous for engineering viral resistance (Graham et al., 1997). A number of plant viruses, such as luteoviruses, reoviruses and most geminiviruses, replicate exclusively in phloem-associated tissue. For such viruses, phloem-specific expression of virus resistance genes is desirable since constitutive expression might unnecessarily increase the risks of transcapsidation or viral recombination in non-vascular tissues (Green and Alison, 1994). Furthermore, phloem-specific gene expression may impose a decreased metabolic load on the plant in comparison to constitutive expression.

To evaluate the activity of the rolC promoter in transgenic sugarcane plants, two reporter genes were used. The first of these was the bacterial uidA gene coding for β-glucuronidase (GUS) (Jefferson et al., 1987). This is the most popular reporter gene in plant histochemical research since the blue staining is easy to localise, and gives visually striking results.

The green fluorescent protein (GFP) of the jellyfish Aequorea victoria (Chalfie et al., 1994) is the second reporter gene that was used. The use of GFP allows direct imaging of the fluorescent protein in living cells without the need for prolonged and lethal histochemical staining procedures. The wild-type GFP is mis-spliced when expressed in Arabidopsis, resulting in little or no fluorescence in transgenic plants (Haseloff and Amos, 1995). However, a modified version of the gene was recently constructed which rectified the splicing problem and therefore resulted in improved fluorescence (Haseloff et al., 1997). Since the performance of the wild-type gene has not been determined in sugarcane cells, the modified gene (mGFP4) was used as reporter gene in the present study.
4.3 MATERIALS AND METHODS

4.3.1 Construction of expression vectors

Standard techniques were used to construct the expression vectors (Sambrook et al., 1989; Ausubel et al., 1997). The rolC promoter was derived from the plasmid pBINrolC (provided by Dr. J. Schell, Max Planck Institute, Köln, Germany) and was subcloned into pBluescript (SK+) to form pRolC100.

pGUS800 was constructed as follows: an 1120 bp EcoR I-BamHI fragment containing the rolC promoter was isolated from pRolC100. The promoter, first exon, first intron and a part of the second exon of rice actin was removed from the plasmid pAct1-D (Zhang et al., 1991) by an EcoR I-BamHI digestion and the rolC promoter was inserted in its place upstream of the uidA gene coding for β-glucuronidase (GUS) (Jefferson et al., 1987). Restriction characterisation of the vector was done. Because of the presence of a Hind III site in the rolC promoter, this enzyme could be used to confirm the inclusion of the promoter in the vector. The rolC promoter-containing fragment was excised by an EcoR I-BamHI double digestion, and the correct size of the construct was verified by a BamHI digestion. Although β-glucuronidase is encoded by the uidA gene, this gene was denoted as GUS for the present study.

pGFP800 was constructed as follows: a green fluorescent protein gene (mGFP4) that was modified for expression in plants was obtained from J. Haseloff (Division of Cell Biology, MRC, Cambridge) (Haseloff et al., 1997) in the plasmid pBIN35S-mGFP4. The mGFP4 gene together with the nopaline synthase terminator sequence (NOS) were subcloned into pBluescript (SK+). A 1240 bp Xba I-Xho I fragment containing the mGFP4 gene and the NOS terminator was cloned into the Xba I-Xho I site of the plasmid pRolC100 downstream of the rolC promoter. Characterisation of the vector by restriction analysis was done. The size of the vector was confirmed by a BamHI digestion, and the inserted mGFP4-NOS terminator sequence was excised by a Xba I-Xho I digestion to confirm both the presence and size of the insert.
4.3.2 Sugarcane tissue culture

Embryogenic (Type III; Taylor et al., 1992) sugarcane callus was produced and maintained as described from mature, field grown plants of variety NCo310 (Bower and Birch, 1992; Snyman et al., 1996).

4.3.3 Sugarcane transformation

Transformation of embryogenic callus by particle bombardment was done using a particle inflow gun (PIG) and DNA-coated tungsten particles. Precipitation of plasmid DNA onto tungsten particles (M17; Bio-Rad Laboratories, CA) was done as described by Snyman et al. (1996). The conditions used for microprojectile bombardment using a PIG (Finer et al., 1992) had previously been optimised for sugarcane transformation (Bower and Birch, 1992; Snyman et al., 1996). The callus was placed onto osmoticum medium (MS basal medium (Sigma)(Murashige and Skoog, 1962) supplemented with 3% (w/v) sucrose, 0.1% (w/v) casein hydrolysate, 0.5% (w/v) agar gel, 3.64% (w/v) D-mannitol and 3.64% (w/v) D-sorbitol) for at least 4 h before and 1 h after bombardment (Vain et al., 1993a). Cotransformation was done using the pEmuGN plasmid (Last et al., 1991) containing the neomycin resistance gene (2.4.2.1), and either pGUS800 or pGFP800. The ratio of the two plasmids were 1:1 and the total amount of DNA was kept constant at 5 μg. Transient gene expression was monitored by using 2 plasmids:

A. pDP687 (Ludwig et al., 1990), which contains the maize anthocyanin transcriptional activation factors. Anthocyanin is produced after successful transformation, and can be visualised as red foci (Figure 4.1 A).

B. pAHC27 (Christensen et al., 1992), which contains the bacterial uidA gene coding for β-glucuronidase (GUS) under control of the maize polyubiquitin (Ubi-1) promoter. Incubation in the substrate-containing GUS assay solution (100 mM sodium phosphate buffer (pH 7.0); 5 mM potassium ferrocyanide (K₄Fe(CN)₆·3H₂O); 5 mM potassium ferricyanide (K₂Fe(CN)₆); 2 mM X-Gluc (5-bromo-4-chloro-3-indolyl-β-glucuronic acid); 0.3% (v/v) Triton X-100) results in the production of blue foci after incubation at 37°C (Jefferson et al., 1987) (Figure 4.1 B).
Stable transformants were selected on medium containing 45 mg l\(^{-1}\) geneticin (G418). Selection was performed for 2 to 3 weeks in the dark on callus maintenance medium (MS basal medium supplemented with 3% (w/v) sucrose, 0.1% (w/v) casein hydrolysate, 0.5% (w/v) agar and 3 mg l\(^{-1}\) 2,4-dichloro-phenoxyacetic acid (2,4-D)) and was continued for several weeks in the light on regeneration medium (MS basal medium supplemented with 3% (w/v) sucrose, 0.1% (w/v) casein hydrolysate and 0.5% (w/v) agar) until the plantlets that regenerated started rooting. The plants were then grown on regeneration medium without antibiotic until they could be hardened off.

4.3.4 Characterisation of transformed plants: DNA analyses
DNA was isolated from 9 month old transformed plants using the modified method of Dellaport et al. (1983). Specific primers for the amplification of the reporter genes were used to verify their presence in the genome of transformed sugarcane plants. Two 22-mer primers were used for the amplification of the GUS reporter gene:
GUS-F: 5'-ATG TTA CGT CCT GTA GAA ACC C-3' and GUS-R: 5'-TCA TTG TTT GCC TCC CTG CTG C-3'. These primers should produce an amplification product of 1 812 bp in length. The amplification of mGFP4 was primed with GFP-F: 5'-AGA TGG TGA TGT TAA TGG GC-3' and GFP-R: 5'-GAT TGT GTG GAC AGG TAA
TGG-3'. These primers should produce an amplification product of 562 bp in length. The PCR mixture was the same as described for the specific amplification of a DNA fragment (3.3.11), and the cycling parameters were as follows: 1 cycle of 94°C for 2 min; 10 cycles of 94°C for 30 s, 50°C for 1 min, 72°C for 1.5 min; 25 cycles of 94°C for 30 s, 45°C for 45 s, 72°C for 1 min. The last cycle was followed by a 5 min extension at 72°C. Amplification products were separated in 1.5% (w/v) agarose gels.

The DNA of transformed plants was analysed by Southern blot hybridisation. Fifteen micrograms of genomic DNA from each transformed plant was treated with the restriction endonuclease, EcoRV. In addition, DNA from the GUS transformants was treated with the restriction endonuclease Alu I, and the GFP transformants with SacI. These enzymes were chosen because of the presence of their recognition sites inside the regions that were targeted with the probes. The rest of the method was as described in 3.3.9, except for the use of either the GUS or the GFP coding sequence as the probe.

4.3.5 Characterisation of transformed plants: RNA analysis
RNA was isolated from the leaves of 9 month old transformed plants using the RNeasy Plant RNA isolation kit (QIAGEN). The presence of transgene mRNA was verified by RT-PCR using the Titan RT-PCR kit (Boehringer Mannheim) and the sequence-specific primers described in 4.3.4. The method was followed as recommended by the manufacturer and the annealing temperature during the amplification steps was 52°C. The amplification products were separated in a 1.5% (w/v) agarose gel.

4.3.6 Characterisation of transformed plants: Histochemical analysis
To assay GUS activity in the mature culm internodal and nodal tissues, a hand microtome was used to cut 0.5 - 0.8 mm transverse sections. The sections were incubated for 48 h at 37°C in GUS assay solution (4.3.3) which was introduced into the cells by a brief (~1 min) vacuum infiltration. After staining the sections were incubated in chlorallactophenol (CLP, 2:1:1 chloral hydrate: lactic acid: phenol) for
10 min to 1 h to clear the pigments from the tissue (Beeckman and Engler, 1994). The cellular location and intensity of GUS staining were assessed using a Zeiss Axioskope microscope and the photographs were taken using AGFA HDC 100 film.

Fluorescence microscopy was done to detect GFP expression in the culm. Sections of 0.5 mm of the internodal tissue were cut using a hand microtome and were immediately mounted on microscope slides in water. A Zeiss Axioptot microscope was used with the following filter set (487909): excitation filter BP450 - 490 nm, chromatic beam splitter FT510, barrier filter LP520. The light source was an HBO 50 W mercury vapour short arc lamp. Photography was done using AGFA HDC100 film.

4.4 RESULTS

4.4.1 Construction of the expression vectors, pGUS800 and pGFP800

Characterisation of the two expression vectors by restriction analysis was done using several different combinations of restriction enzymes. The expected and obtained sizes of restriction fragments of both constructs are summarised in Table 4.1. The results of the restriction analyses are shown in Figure 4.2 (A) and (B). Maps of the constructs are shown in Figure 4.3.

<table>
<thead>
<tr>
<th>Enzyme</th>
<th>Expected size (bp)</th>
<th>Actual size (bp) in gel</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>pGUS800</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BamH I</td>
<td>6710</td>
<td>6700</td>
</tr>
<tr>
<td>Hind III</td>
<td>6311, 405</td>
<td>6300, 400</td>
</tr>
<tr>
<td>EcoR I–BamH I</td>
<td>5590, 1120</td>
<td>5600, 1100</td>
</tr>
<tr>
<td><strong>pGFP800</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BamH I</td>
<td>5321</td>
<td>5300</td>
</tr>
<tr>
<td>Hind III</td>
<td>3356, 1965</td>
<td>3300, 2000</td>
</tr>
<tr>
<td>Xba I–Xho I</td>
<td>4081, 1240</td>
<td>4000, 1250</td>
</tr>
</tbody>
</table>

1 The sizes of the fragments as determined by computer analysis of the sequences.
2 The sizes of the fragments as determined by comparison to molecular weight markers.
Figure 4.2 Restriction analysis of the reporter gene expression vectors, pGUS800 and pGFP800. The expected and obtained results of these analyses are summarised in Table 4.1. A.) pGUS800 was treated with BamH I (lane 2), Hind III (lane 3) and EcoR I-BamH I (lane 4). Sizes of the resulting fragments were determined using molecular weight markers 3 (lane 1) and 5 (lane 5). B.) pGFP800 was treated with BamH I (lane 2), Hind III (lane 3) and Xba I-Xho I (lane 4). Sizes of the fragments were determined by comparing them to molecular weight marker 3 (lane 1).

Figure 4.3 Schematic representation of the two reporter gene expression vectors, pGUS800 and pGFP800. A.) pGUS800. The rolC promoter-containing EcoR I-BamH I fragment was removed from pBINrolC and inserted upstream of the GUS reporter gene after the rice actin promoter was removed from pAct1-D. The resulting chimeric construct was 6710 bp in length. B.) pGFP800. An Xba I-Xho I fragment containing mGFP4 and the NOS terminator sequence was removed from the subclone pBluescript-mGFP4(SK+). This fragments was inserted downstream from the rolC promoter in pBINrolC to replace the OCS terminator sequence. The resulting chimeric construct was 5321 bp in length.
4.4.2 Sugarcane transformation
A total of 11 plants survived the selection on geneticin-containing medium. These included 7 plants from a single bombardment with the pGUS800 construct and 4 plants from 2 separate bombardments with the pGFP800 construct.

4.4.3 Characterisation of transformed plants: DNA analyses
Amplification of the inserted GUS and GFP reporter genes from DNA isolated from the transformed plants showed that all 7 of the GUS transformants and 3 of the 4 GFP transformants contained the foreign DNA. All the GFP transformants that showed positive results were obtained from the same bombardment. The same amplification reactions using DNA from the control untransformed sugarcane did not produce the expected size fragments (Figure 4.4).

![Figure 4.4](image)

*Figure 4.4* PCR amplification of reporter genes from transformed plants. Reporter gene coding regions were amplified from genomic DNA isolated from transformed plants using sequence-specific primers. A.) Amplification of the GUS coding region from the untransformed NCo310 control (lane 2), transformants GUS 4.2.1 (lane 3), 4.2.2 (lane 4), 4.2.3 (lane 5), 4.2.4 (lane 6), 4.3E (lane 7) 4.3.1 (lane 8) and 4.3.2 (lane 9). B.) Amplification of the GFP coding region from the untransformed NCo310 control (lane 2), transformants GFP 1.1E (lane 3), 1.1 (lane 4), 1.2 (lane 5) and 1.3 (lane 6). The sizes of amplified fragments were verified using markers 3 (lane 1) and 5 (lane 7).

Southern blot analysis confirmed the PCR results. Positive signals were obtained from all GUS transformants and were as follows: samples treated with *EcoR* V had two bands of sizes 3400 and 1400 bp, while two bands of sizes 1150 and 700 bp were observed when samples were treated with *Alu* I. No hybridisation signal was
seen in the untransformed control sample (Figure 4.5 A).

Positive signals were obtained from the same 3 GFP transformants that produced positive results with the PCR, namely GFP1.1E, 1.2 and 1.3, but not GFP1.1. A single band of 2000 bp was seen in the positive samples when the DNA was treated with EcoR V, and three bands of 3000, 1200 and 750 bp, respectively, were seen when the DNA was treated with Sac I. No bands were seen in the untransformed control sample (Figure 4.5 B).

![Figure 4.5](image)

**Figure 4.5** Southern blot analysis of transformed plants to confirm the presence of the reporter genes. 15 μg of genomic DNA from transformed plants and an untransformed control plant were treated with endonucleases, separated in 0.8% agarose gels and transferred to nylon membranes. A.) i) DNA treated with EcoR V and probed with a radioactively labelled GUS fragment. Untransformed NCo310 control (lane 1), GUS 4.2.1 (lane 2). ii) DNA treated with Alu I and probed with a radioactively labelled GUS fragment. Untransformed NCo310 control (lane 3), GUS 4.2.1 (lane 4). The other transformants showed exactly the same hybridisation pattern as observed for GUS 4.2.1. B.) i) DNA treated with EcoR V and probed with a radioactively labelled GFP fragment. Untransformed NCo310 control (lane 1), GFP 1.1E (lane 2). ii) DNA treated with Sac I and probed with a radioactively labelled GFP fragment. Untransformed NCo310 control (lane 3), GFP 1.1E (lane 4). The other transformants showed exactly the same hybridisation pattern as observed for GFP 1.1E.

### 4.4.4 Characterisation of transformed plants: RNA analysis

RT-PCR of the RNA isolated from the GUS transformed plants produced the expected 1812 bp fragment, which was absent from the untransformed control plant,
when the GUS-specific primers were used (Figure 4.6 A). When the GFP-specific primers were used, the expected 526 bp GFP fragment was formed in 3 of the 4 GFP transformants, but not in the untransformed control plant (Figure 4.6 B).

Figure 4.6 RT-PCR of transformed plant RNA to confirm the presence of the reporter gene transcripts. RT-PCR was done using sequence-specific primers and amplification products were separated in 1.5% (w/v) agarose gels. A.) RT-PCR of the GUS transcript from RNA isolated from the untransformed control plant and the GUS transformants. GUS 4.2.1 (lane 2), 4.2.2 (lane 3), 4.2.3 (lane 4), 4.2.4 (lane 5), 4.3E (lane 6), 4.3.1 (lane 7), 4.3.2 (lane 8) and control (lane 9). B.) RT-PCR of the GFP transcript from RNA isolated from the untransformed control plants and the GFP transformants. GFP 1.1 (lane 2), 1.1E (lane 3), 1.2 (lane 4), 1.3 (lane 5) and control (lane 6). The sizes of amplified fragments were verified using marker 3 (lane 1) and marker 5 (lane 7).

4.4.5 Characterisation of transformed plants: Histochemical analysis
GUS expression could be visualised in all GUS transformants after staining with the assay solution for 48 h. In the thicker (0.8 mm) sections, staining was clearer. Blue coloration was limited to the vascular bundles and specifically to the phloem (Figure 4.7 B). Due to the thickness of the sections it was difficult to focus the microscope. In transverse sections of the nodal region of young culm, horizontal vascular bundles also stained intensely (Figure 4.7 D). No specific staining was observed in the untransformed control tissue as shown in Figure 4.7 (A,C).
Figure 4.7 Histochemical analysis of GUS expression in sugarcane cells. Culm sections (0.5 - 0.8 mm) of transformed plant GUS 4.2.1 and an untransformed control plant were stained in assay solution for 48 h and cleared in CLP for 1 h. All sections were visualised with brightfield illumination. A.) A single vascular bundle in the culm of untransformed NCo310 internodal region (400 x). B.) A single vascular bundle in the culm of GUS 4.2.1 internodal region (400 x). The position of the phloem is indicated with an arrow. C.) Section of mature node of untransformed NCo310 (16 x). D.) Section of mature node of GUS 4.2.1 (16 x). A horizontal vascular bundle is indicated with an arrow.

Because plant material contains various autofluorescent and light-scattering components which can obscure the green fluorescence of GFP, four different filter sets were used for fluorescence microscopy to determine the optimal set for the visualisation of GFP in sugarcane. GFP expression in the vascular bundles was visualised using a standard fluorescein isothiocyanate (FITC) filter set with maximum excitation at 485 nm (460 - 510 nm) and 540 nm (515 - 565 nm)(Figure 4.8 A), a standard FITC filter set with barrier filter at 420 nm (Figure 4.8 B), a filter with excitation at 450 - 490 nm, and barrier filter at 520 nm (Figure 4.8 C), and filter set 487909 with excitation filter BP450 - 490 nm, chromatic beam splitter FT510 and barrier filter LP520 (Figure 4.8 D). Different microscopes were used for these analyses. Filter sets (A) and (B): Olympus BH2 (PM-20), (C): Zeiss MC 63 and (D):
Zeiss Axiophot. The light source of the Zeiss MC 63 microscope illuminates the sample from below. With the other two microscopes that were used, the samples were illuminated from above. Optimal visualisation of GFP in sugarcane cells was obtained using the Zeiss Axiophot microscope and filter set 487909.

Figure 4.8 Visualisation of GFP expression in phloem cells using different filter sets and microscopes. A.) Standard FITC filter set with maximum excitation at 485 nm (460 - 510 nm) and 540 nm (515 - 565 nm). B.) Standard FITC filter set with barrier filter at 420 nm. C.) Excitation filter at 450 - 490 nm, chromatic beam splitter at 510 nm and barrier filter at 520 nm. D.) Filter set 487909 with excitation filter BP450 - 490 nm, chromatic beam splitter FT510 and barrier filter LP520. In A, B and D the samples were illuminated from above and in C illumination was from below. X: xylem; P: phloem.

In 3 of the 4 GFP transformants GFP expression was observed in the phloem cells of vascular bundles of the mature culm as shown in Figure 4.9 (D, F). A lower level of fluorescence was observed in younger tissues and in vascular bundles situated on the periphery of the culm sections. The control plant sections did not show any fluorescence in the phloem, as shown in Figure 4.9 (C, E).
4.5 DISCUSSION

It is evident from the results that the rolC promoter from Agrobacterium rhizogenes was successfully engineered into sugarcane through microprojectile bombardment. This was done with two different types of genetic constructs where either GUS or GFP was used as a reporter gene. The only other monocotyledonous plant in which rolC has been evaluated before is rice (Matsuki et al., 1989).

DNA analyses confirmed that the foreign DNA was incorporated into the sugarcane genome. The expected size products were amplified from the transformants indicating that no rearrangement of the genes had taken place in the plant cells. Southern blot results revealed that all the GUS transformants and all the GFP transformants resulted from single transformation events and that, in each case, different plants were regenerated from a single piece of transformed callus. The fact that all the plants transformed with GUS and GFP constructs respectively contained the rolC promoter and reporter gene inserted into the same locus in the sugarcane genome, makes it impossible to determine whether the insertion of the DNA into different regions of the genome can lead to variation in reporter gene expression levels. For the determination of tissue-specificity of the rolC promoter in sugarcane, however, different transformants were not essential, and consequently the above-mentioned plants were used in all subsequent analyses.

Fragments of the expected sizes for the reporter gene mRNAs were obtained from RT-PCR analysis which confirmed that the intact messengers were transcribed in the sugarcane cells. The RT-PCR technique that was used to determine the presence of the transcripts is not a quantitative method and northern blots were not performed. No conclusion can therefore be made about the level of reporter gene mRNA in the plants. Expression of GUS and GFP was detected in the sugarcane transformants by histochemical analysis, allowing the tissue-specificity of the promoter to be determined.
GUS staining patterns in the culm tissues revealed a tissue-specific expression of the β-glucuronidase in the phloem. This result is consistent with expression studies in other transgenic plants using the rolC promoter, including tobacco (Scmulling et al., 1989), rice (Matsuki et al., 1989) and hybrid aspen (Nilsson et al., 1996). Although "phloem" can include companion cells, sieve tubes, phloem parenchyma and fibre cells, which are very different cell types (Graham et al., 1997), the specific cell type where rolC is active in the sugarcane phloem tissues was not determined. It does, however, warrant further investigation, and the analysis of reporter gene expression in the leaf tissue where the cell types in the vascular bundles are better defined (Graham et al., 1997) can be done. Strong staining was also seen in the horizontal vascular bundles in the nodal tissue (Fahn, 1990). The different cell types in these bundles were not distinguished and it is therefore difficult to explain the strong signal in the nodes. Clearing the tissues in CLP after incubation in the GUS assay solution (Beeckman and Engler, 1994) increased the intensity of the blue color in sugarcane cells where the protein is being expressed.

GFP expression was also restricted to the phloem cells in the sugarcane culm tissue. A very specific filter set and microscope had to be used for the visualisation of GFP, because plant tissues contain various autofluorescent and light-scattering components which can obscure the green fluorescence, especially when expression levels are low. Clear results could only be obtained with one of the four filter sets (Figure 4.8 D) that were used. The same filter set was used to obtain different results with two different microscopes (Figure 4.8 C and D). It was evident that the best results were obtained when the samples were illuminated from above. Laser-scanning confocal microscopy is often used for the visualisation of GFP in plant tissue (Haseloff and Amos, 1995; Kain et al., 1995), but was not available for the present study.

It has been reported that plants expressing high levels of GFP are difficult to regenerate (Haseloff et al., 1997). The protein may be a source of fluorescence-related free radicals and it has been advised that the protein should be targeted to
a more localised compartment within the cell (Haseloff et al., 1997). In jellyfish photocytes, where naturally high levels of GFP are well tolerated, the protein is found sequestered in cytoplasmic granules. In contrast, the mature protein is found throughout the cytoplasm and nucleoplasm of transformed Arabidopsis cells (Haseloff et al., 1997). GFP expression was limited to the vascular bundles in the current investigation, and these same problems were not encountered. The results discussed here were obtained by using the modified version of the GFP gene (mGFP4), and the possibility that different results will be obtained when using the wild-type gene was not investigated.

The organ-related activation of reporter gene expression (roots>stems>leaves) that has been reported for transgenic tobacco (Sugaya et al., 1989) was not investigated in sugarcane tissues although higher GUS expression was observed in the mature culm tissue than in the younger parts of the culm. GFP expression in the mature culm was stronger in the vascular bundles situated in the core and decreased toward the peripheral regions. This may, at least in part, be attributed to the fact that peripheral vascular bundles in sugarcane contain particularly small or no phloem cells (Jacobsen et al., 1992).

The pattern of reporter gene expression in the sugarcane culm can be directly compared to the sucrose gradient in the plant. Sucrose levels increase from young to mature tissue (from top to bottom) and from the peripheral region to the core in the internodes (Whittaker and Botha, 1997). These results suggest that sucrose influences rolC activity in sugarcane. The rolC promoter has previously been shown to be activated by sucrose in the phloem cells of transgenic tobacco (Yokoyama et al., 1994). Results obtained in the present study were not compared to previous reports, since the sucrose levels were not quantified. Such analyses can be included in future investigations.

Activity of the rolC promoter was determined in the planted crop only and not in subsequent ratoon crops. This indicates that the promoter is stable at least under
glass house conditions. In potatoes the expression of GUS under control of the rolC promoter has been shown to be stable when the potatoes were stored at 10°C for extended periods (Graham et al., 1997). Because unstable transgene expression poses a major difficulty for biotechnology, the stability of the expression driven by the rolC promoter in sugarcane ratoon crops and under field conditions still has to be investigated.

4.6 CONCLUSIONS
Both GUS and GFP are expressed in sugarcane plants under the control of the rolC promoter from Agrobacterium rhizogenes, confirming that the rolC promoter is active in sugarcane. Expression of these proteins is restricted to the phloem, and expression levels are influenced by the sucrose content in the mature sugarcane plant.
CHAPTER 5

ISOLATION OF THE PYROPHOSPHATE-DEPENDENT PHOSPHOFRUCTOKINASE β-SUBUNIT (PFP-β) PROMOTER

5.1 ABSTRACT

Activity of pyrophosphate-dependent phosphofructokinase (PFP) in sugarcane increases with culm maturity and is negatively correlated with sucrose levels in the culm. The enzyme is also apparently preferentially expressed in the peripheral areas of the mature culm. These properties suggest that the PFP promoter may be used to drive transgene expression in the mature sugarcane culm. Enrichment of transgene expression products in this area of the plant will be of great benefit for the manipulation of sucrose metabolism.

Sequence information of the PFP-β gene of castor bean and potato was used to design degenerate primers for the PCR screening of sugarcane cDNA libraries. Two fragments were isolated and their sequences showed high homology to the castor bean and potato genes. A mature culm cDNA library was screened using one of the amplified fragments as a probe and three positive clones were identified. These clones all contained the same length incomplete cDNA fragment and sequence analysis revealed that the clones were identical. The incomplete cDNA fragment was used to screen a sugarcane genomic library and four positive clones were isolated. Three of these contained the 5' end and the other one the 3' end of the PFP-β gene. A Deletion Factory™ system was used to analyse the sequence of one of the clones that contained the 5' end of the gene.

A part of the coding region containing the first five exons as well as the 5' flanking region of the gene was identified. Sequence-specific primers were designed for the amplification of two fragments in the promoter area. These fragments were of different lengths, one containing 1 072 bp of the promoter region and the other containing the same fragment plus the first intron of the gene. Expression vectors
containing these promoter fragments together with the GUS reporter gene were constructed. PFP promoter activity was determined by the quantification of transient reporter gene expression in sugarcane callus. Transient activity was compared to that of promoters that have previously been tested in sugarcane.

Results indicated that the region of the PFP-β promoter that was included in the constructs is capable of driving foreign gene expression in sugarcane callus. Although expression levels were low in comparison to the other promoters that were tested, further analysis is needed to ensure that all sequences that are crucial for optimal promoter activity are included.

5.2 INTRODUCTION

Pyrophosphate dependent phosphofructokinase (PFP) catalyses the reversible conversion of fructose 6-phosphate (Fru-6-P) and pyrophosphate (PPI) to fructose 1,6-bisphosphate (Fru-1,6-P2) and inorganic phosphate (Pi). The activity of the enzyme is high in plant cells, especially those that have a high biosynthetic activity (Dennis and Greyson, 1987).

PFP is, in most plants, composed of two immunologically unrelated subunits, α and β (Kruger and Dennis, 1987). The α-subunit of 65 - 68 kD may be involved in the regulation of the enzyme by binding the allosteric activator fructose 2,6-bisphosphate (Cheng and Tao, 1990), while the smaller β-subunit of 60 - 64 kD is the catalytic subunit. It contains putative binding sites for Fru-6-P and PPI. The PFP enzyme from the intestinal anaerobe Giardia lamblia has been shown to have only one subunit which is similar to the β- rather than the α-subunit of plant PFP (Rozario et al., 1995). Both subunits are expressed in most tissues (Carlisle et al., 1990), but differential expression of the genes has been reported (Botha and Botha, 1991; Botha and Botha, 1993a; Botha and Botha, 1993b; Blakeley et al., 1992). This differential expression of the subunits, both of which are required to produce an enzyme with significant activity, is difficult to explain (Dennis and Blakeley, 1995).
Although the kinetic properties of PFP have been studied in detail, the physiological role of the enzyme is still being investigated. Possible roles for the enzyme in both the regulation of gluconeogenesis and biosynthetic activity in plant cells have been suggested (Dennis and Greyson, 1987; Huber and Akazawa, 1986; Nakamura et al., 1992; Botha and Botha, 1993a; Botha and Botha, 1993b).

In sugarcane, sucrose accumulation is suggested to be primarily regulated within the translocation system and at the level of the sink (Moore, 1995). There is likely to be substantial competition for incoming sucrose for storage and biosynthetic activity, including structural fiber in sugarcane sink tissue (Whittaker and Botha, 1996). PFP activity was shown to increase from internode 3, reaching optimal activity at internode 6 and 7, after which the enzyme declines to internode 10. An inverse correlation between peak PFP activity/expression levels and the prevailing sucrose content of internodal tissue, and a positive relationship between PFP activity and the insoluble matter content was observed (Whittaker and Botha, 1996). Sucrose content in the sugarcane culm was shown to increase from internodes 3 to 9 (Whittaker and Botha, 1997).

The distribution of sucrose, fiber and PFP activity within an internode was investigated, and PFP activity correlated positively with the increase in fiber deposition from the core to the periphery. There was an inverse trend in sucrose accumulation, which correlated negatively with PFP activity (Whittaker and Botha, 1996).

It is hypothesised that PFP activity in sugarcane sink tissues is likely to be associated with metabolic processes involving sucrose degradation and utilisation, rather than increased storage (Whittaker and Botha, 1996). PFP and sucrose synthase have also been suggested to play a role in controlling sink strength (Black et al., 1995).

Molecular technology, through the manipulation of key enzymes, can be used to increase sucrose yield as well as facilitate the study of sucrose metabolism in
sugarcane. In future studies where enzymes will be down regulated or over expressed, the use of specific promoter elements that are active in the mature culm where the changes are required will be of great importance. The properties of PFP in sugarcane have made it an ideal candidate to use in these studies. Since the β-subunit has been shown to be the catalytic subunit of the enzyme, it was hypothesised that the promoter of PFP-β may confer differential expression in the sugarcane culm.

5.3 MATERIALS AND METHODS

5.3.1 Degenerate and specific primer design
Degenerate primers were designed after the PFP-β gene sequences from castor bean (accession number Z32850) and potato (accession number J05629) were compared and an area of high homology identified. Initially, two primers were designed for the amplification of sugarcane PFP from cDNA libraries (5.3.2), namely PFP-B1, a forward primer, and PFP-B2, a reverse primer. Four primers were then designed for the sequencing and amplification of the 5' flanking region of the PFP-β gene (5.3.12 and 5.3.13). The details of all the primers that were designed for this study are summarised in Table 5.1.

Table 5.1 Oligonucleotide primers designed for the amplification of PFP-β cDNA and 5' flanking region.

<table>
<thead>
<tr>
<th>Primer</th>
<th>Sequence (5' to 3')</th>
<th>Length (bp)</th>
<th>Tm (°C)</th>
<th>Degenerate/Specific</th>
</tr>
</thead>
<tbody>
<tr>
<td>PFP-B1</td>
<td>ATI-GAT-TTC-ATI-CCI-GAG-GT</td>
<td>20</td>
<td>51.2</td>
<td>D</td>
</tr>
<tr>
<td>PFP-B2</td>
<td>TCA-TCI-ACA-ACA-TCA-TGI-GCC</td>
<td>21</td>
<td>55.9</td>
<td>D</td>
</tr>
<tr>
<td>PFP-PROM1</td>
<td>GGA-AGA-ACG-TCG-TCT-GAT-TGC</td>
<td>21</td>
<td>59.8</td>
<td>S</td>
</tr>
<tr>
<td>PROM-F</td>
<td>TCA-CTG-TGG-ATC-CCT-AGT-CAA-AGG-GAG</td>
<td>27</td>
<td>66.5</td>
<td>S</td>
</tr>
<tr>
<td>PROM-R</td>
<td>CGC-CGC-CGG-ATC-CTG-CCT-TGC-CCT-TGG</td>
<td>27</td>
<td>75.6</td>
<td>S</td>
</tr>
<tr>
<td>PFP-E2</td>
<td>CAA-ACA-AGT-TAG-GAT-GCA-GCT-TGG-C</td>
<td>25</td>
<td>63</td>
<td>S</td>
</tr>
</tbody>
</table>

BamHI sites in sequences of primers PROM-F, PROM-R and PFP-E2 are underlined.

5.3.2 Amplification of PFP fragments from cDNA libraries
Two cDNA libraries from sugarcane (leafroll and mature culm) were screened by PCR using each of the degenerate primers, PFP-B1 and PFP-B2, in combination with
one of the library vector-specific primers, T7 and pUC18/M13 reverse primer (R). The
PCR reaction mixtures contained 1.5 µl of cDNA library template (~3 x 10⁹ pfu ml⁻¹),
1x Taq Polymerase buffer (10 mM Tris-HCl; 1.5 mM MgCl₂; 50 mM KCl; 1 mg ml⁻¹
gelatin; pH 8.3, 20°C), 0.5 µM of each primer, 0.2 mM of each dNTP and 1 unit of
Taq DNA Polymerase (Boehringer Mannheim) in a volume of 20 µl. The samples
were overlaid with 30 µl mineral oil and subjected to 42 cycles of PCR using the
following cycling parameters: 1 cycle of 94°C for 1 min; 10 cycles of 94°C for 45 s,
50°C for 30 s and 72°C for 2 min; 30 cycles of 94°C for 30 s, 45°C for 30 s and 72°C
for 1 min. The last cycle was followed by a 2 min extension at 72°C. PCR
amplification products were separated in 1.5% (w/v) agarose gels. The fragments
were extracted from the gel using the QIAquick Gel Extraction Kit (QIAGEN).

5.3.3 Cloning of cDNA fragments
The fragments were cloned using the pCR-Script (SK+) cloning kit (Promega). The
manufacturer's instructions were followed, except for the use of E.coli HB101 cells
instead of Epicurian coli supercompetent cells for transformation. Clones were
identified using blue/white selection on X-Gal-containing LB medium (3.3.7), and
insert sizes were determined by PCR amplification of the inserted fragments using
the vector-specific T7 promoter and pUC18/M13 reverse primers.

5.3.4 Sequence determination of cDNA fragments
Sequence determination of amplified cDNA fragments was done using the Dye
Terminator Cycle Sequencing Ready Reaction with AmpliTaq DNA Polymerase FS
(Perkin Elmer Applied Biosystems), and sequences were analysed using a Perkin
Elmer ABI Prism 310 Genetic analyser. An average of 500 nucleotides were
sequenced per reaction.

5.3.5 Screening cDNA and genomic DNA libraries
Apart from the host cell strains and the fragments used as probes, the same method
was used to screen the cDNA and genomic libraries. The amplified mature culm
(internode 7) cDNA library in Lambda Zap II (Stratagene) had a titre of 3x10⁹ pfu ml⁻¹
and the amplified genomic library in Lambda Fix II (Stratagene), $1.7 \times 10^9$ pfu ml$^{-1}$. The libraries were screened using MagnaGraph nylon membranes (MSI). In the first round, 250 000 plaque forming units (pfu) were screened using 10 plates.

Host cells of the cDNA library were XL1-Blue MRF$^+$ and of the genomic library, XL1-Blue MRA (P2). The cells were grown to an OD$_{600}$ of 1.0 in Luria broth (1% (w/v) NaCl; 1% (w/v) tryptone; 0.5% (w/v) yeast extract) supplemented with 0.2% (w/v) maltose and 10 mM MgSO$_4$. Cells were centrifuged and suspended in 10 mM MgSO$_4$. Before use, the cells were diluted to an OD$_{600}$ of 0.5 with 10 mM MgSO$_4$. For the first round, 200 µl of diluted cells were added to the calculated volume of the library and incubated at 37°C for 15 min. Top agarose (1% (w/v) NZ-amine (casein hydrolysate); 0.5% (w/v) NaCl; 0.5% (w/v) bacto yeast extract; 0.2% (w/v) MgSO$_4.7H_2O$; 0.7% (w/v) agarose) was melted, dispensed into 4 ml aliquots and left to cool to 48°C. The top agarose was added to the cell-phage suspension and poured onto prewarmed, dry NZCYM plates (1% (w/v) NZ-amine (casein hydrolysate); 0.5% (w/v) NaCl; 0.5% (w/v) bacto yeast extract; 0.2% (w/v) MgSO$_4.7H_2O$; 1.5% (w/v) agar). After an incubation of 16 h at 37°C plaques were visible on the agar surface. The plates were stored at 4°C for at least 2 h before the phages were transferred to nylon membranes. Dry nylon membranes were placed on top of the agar and were left for 20 min. The phage DNA was denatured by soaking the membrane in denaturing solution (1.5 M NaCl; 0.5 N NaOH) for 2 min, neutralised by soaking in neutralisation solution (1.5 M NaCl; 0.5 M Tris, pH 8.0) for 5 min and rinsed in 2x SSC; 0.2 M Tris, pH 7.5 for 30 s. The membrane was then blotted dry on filter paper and the DNA was linked to the nylon membrane by exposing it to UV light for 2.5 min at 120 mJ cm$^{-1}$.

Prehybridisation was done in RapidHyb buffer (Amersham) for 4 h at 50°C. Random primer labelling of probes was done using the Prime It II kit (Stratagene) and hybridisation was done for 16 h at 50°C. The membranes were washed in 2x SSC; 0.1% (w/v) SDS for 20 min at 42°C and again with 0.5x SSC; 0.1% (w/v) SDS for 20 min at 50°C. The membranes were exposed to X-ray film for 24 h. Based on the
hybridisation results, positive plaques were punched out of the agar plates with the back of a pasteur pipette. These agar plugs were placed into 500 µl SM-buffer (0.58% (w/v) NaCl; 0.2% (w/v) MgSO₄·7H₂O; 50 mM Tris (pH 7.5); 0.01% (v/v) gelatin) with 20 µl chloroform. Second and third round screening was carried out in exactly the same manner, except that fewer pfu's were plated on each plate to facilitate single plaque isolation.

In the case of positive cDNA library clones, a phagemid rescue protocol (Stratagene) was followed to isolate the pBluescript (SK+) plasmid containing the positive insert from the phage by use of a helper phage.

5.3.6 Southern blot analysis
DNA from the following plants were used for a Southern blot analysis of PFP-β:
1.) Saccharum officinarum (Black Cheribon), 2.) Saccharum spontaneum (Coimbatore), 3.) N21 (commercial sugarcane variety), 4.) NCo310 (commercial sugarcane variety), 5.) Solanum tuberosum (potato), 6.) Ricinus communis (castor bean) and 7.) Zea mays (maize). Twenty micrograms of genomic DNA from each plant was treated with the restriction endonuclease EcoR I. The rest of the procedure was done as described (3.3.9), with the incomplete PFP cDNA fragment isolated from sugarcane clone C3 (5.4.2) being used as a probe.

5.3.7 Northern blot analysis
The size of the sugarcane PFP-β mRNA was determined using northern blot analysis. Twenty micrograms of total RNA from leafroll tissue (variety NCo310) was separated in a 1.2% (w/v) agarose gel. The rest of the procedure was done as described (3.3.10). The incomplete PFP-β cDNA fragment in clone C3 (5.4.2) was used as the probe.

5.3.8 Isolation of Lambda-DNA
Lambda DNA was isolated from positive genomic library clones using the QIAGEN Lambda mini kit (QIAGEN). Plate lysates were prepared according to the method
described in Sambrook et al. (1989), and these lysates were used directly for the isolation of Lambda DNA. The DNA was treated with 0.2 mg ml\(^{-1}\) RnaseA (Boehringer Mannheim) for 1 h at 37\(^{\circ}\)C, after which a chloroform:isoamylalcohol (24:1) extraction was done. The purified Lambda DNA was quantified spectrophotometrically and by visual comparison to a known concentration of Lambda DNA (Boehringer Mannheim) in an ethidium bromide-stained (0.01 mg ml\(^{-1}\)) 0.8% (w/v) agarose gel.

5.3.9 Characterisation of positive genomic library clones
The insert sizes of the positive genomic library clones were determined by treating 1 \(\mu\)g of the Lambda DNA from each clone with the restriction enzyme \(N\ot\)l I for 2 h at 37\(^{\circ}\)C. The \(N\ot\)l I restriction sites are located on both sides of the inserted plant genomic DNA in Lambda Fix II (Stratagene). After restriction the DNA fragments were separated in an ethidium bromide-containing (0.01 mg ml\(^{-1}\)) 1% (w/v) agarose gel. The DNA was then transferred to a nylon membrane as described in 3.3.6.

Two separate probes were prepared from the incomplete PFP-\(\beta\) cDNA fragment (5.4.2) for the characterisation of the positive genomic clones. One probe, 5'EIB340, was 340 bp in length and was prepared by an \(EcoR\ I-Bgl\ II\) double digestion of the cDNA. This fragment was located on the 5' end of the PFP-\(\beta\) cDNA. The second fragment, 3'EVEI650, was 650 bp in length and was located on the 3' end of the cDNA fragment. The fragment was prepared by treating the cDNA with \(EcoR\ V\) and \(EcoR\ I\). These probes were prepared to determine whether the 5' end or the 3' end of the PFP-\(\beta\) gene was contained in the genomic clones.

Twenty five nanograms of each fragment was radioactively labelled using the Prime-It II random primer labelling kit (Stratagene) in the presence of \([\alpha-^{32}\P]dCTP\) (Amersham; 3000 Ci mmol\(^{-1}\)). After a 4 h prehybridisation in RapidHyb buffer (Amersham) at 48\(^{\circ}\)C, the probe was denatured for 5 min in boiling water and added to the prehybridisation buffer. Hybridisation was done for 16 h at 48\(^{\circ}\)C after which the membrane was washed once at 45\(^{\circ}\)C for 20 min in a solution containing 2x SSC and 0.1% (w/v) SDS. The membrane was sealed in a plastic bag and exposed to X-ray
film for 16 h. After the film was developed, the radioactive probe was stripped from
the membrane by pouring boiling 0.1% (w/v) SDS onto it and leaving it to cool to
room temperature. The membrane was rinsed in 0.1x SSC before it was
prehybridised again and probed with the second fragment.

5.3.10 Construction of a Deletion Factory™

The analysis of the positive genomic clones was done using the Deletion Factory
System Version 2.0 (GibcoBRL). This system is based on the generation of nested
deletions across large DNA inserts by intramolecular transposition using the
engineered transposon, ϕC31 (Tn 1000). After transposition, one transposon end always
abuts a deletion endpoint, serving as the sequencing primer binding site. Deletion
derivatives are recovered by plating on microbiological media, and selecting for loss
of a contra-selectable marker. This selection, in conjunction with screening for the
presence of a downstream selectable marker, results in nested deletions that extend
for varying distances from one transposon end, through the contra-selectable marker,
and into, but not beyond, the cloned insert DNA. Deletion endpoints can be mapped
by determining the size of the subclone using agarose gel electrophoresis.

Four steps are involved in the Deletion Factory System, namely cloning of the insert
DNA into the cosmid pDELTA 2, transforming competent cells, screening for
deletions and isolating plasmid DNA for sequencing.

– Cloning of the genomic DNA insert into pDELTA 2

The Lambda DNA from the positive genomic library clone G2, was treated with the
restriction enzyme Not I to remove the insert from the phage arms. Two insert
fragments were formed by this digestion and they were both included in a ligation
reaction with pDELTA 2, which was prepared by a Not I digestion and an alkaline
phosphatase treatment using standard protocols (Ausubel et al., 1997). The phage
arms had only one cohesive end each and could therefore not be cloned during this
reaction. The ligation reaction mixture contained 1x Ligase buffer (66 mM Tris-HCl;
5 mM MgCl₂; 1 mM DTT; 1 mM ATP; pH 7.5, 20°C), 5 units of T4 DNA Ligase
Boehringer Mannheim, 2.5 µg Lambda DNA from clone G2 and 250 ng of prepared pDELTA 2 cosmid, and was incubated at 22°C for 5 h. Transformation of DH10B cells (GibcoBRL) was done following the protocol supplied with the Deletion Factory System, and positive clones were selected on Luria plates (3.3.7) containing: 0.2 mg ml⁻¹ ampicillin, 0.05 mg ml⁻¹ kanamycin and 0.05 mg ml⁻¹ X-Gal.

Plasmid DNA was prepared from selected clones using an alkaline lysis method (modified method of Ish-Horowicz and Burke, 1981). Because two clonable fragments were formed after Not I digestion of clone G2 Lambda DNA, the pDELTA 2 clones containing the 5' end region of PFP-β had to be identified. The DNA was treated with restriction enzymes EcoR I and Hind III, and after separation of the fragments in a 0.8% (w/v) agarose gel, they were transferred to a nylon membrane. This Southern blot was probed with 5'EIB340 (5.3.9).

The insert size of the positively identified clone, pΔ2-PFPC, was determined by removing it from pDELTA 2 by Not I digestion.

**Generating nested deletions**

pΔ2-PFPC plasmid DNA was used to transform E.coli DF1 cells (GibcoBRL) where deletions were generated by the action of γδ transposase (tnpA), which specifically recognises the transposon end sequence. DF1 provides tnpA on a compatible chloramphenicol resistance plasmid. Transformants were grown in Luria broth (1% (w/v) NaCl; 1% (w/v) tryptone; 0.5% (w/v) yeast extract) containing 0.2 mg ml⁻¹ ampicillin, 0.05 mg ml⁻¹ kanamycin and 0.015 mg ml⁻¹ chloramphenicol. Clockwise deletion transformants were subsequently grown in Luria broth containing 0.2 mg ml⁻¹ ampicillin and 0.015 mg ml⁻¹ chloramphenicol, and counterclockwise deletion transformants in Luria broth containing 0.05 mg ml⁻¹ kanamycin and 0.015 mg ml⁻¹ chloramphenicol. Plasmid DNA was isolated from these transformants and used to transform DH10B cells. The cells were plated out on two different sets of Luria plates, the one set containing 0.2 mg ml⁻¹ ampicillin and 5% (w/v) sucrose, and the other containing 0.05 mg ml⁻¹ kanamycin and 0.1 mg ml⁻¹ streptomycin. Plasmid DNA was
purified from these deletion subclones and the insert sizes determined by comparing the mobility of undigested and digested DNA in a 1% (w/v) agarose gel.

5.3.11 Sequence determination of cloned inserts
Inserts of deletion subclones were sequenced as described in 5.3.4. The Contig Manager function of DNASIS for Windows Version 2.0 (Hitachi Software) was used to consecutively arrange the sequence of each clone with those of its neighbors by locating regions at the ends of clones that overlapped with those of neighboring clones. Open reading frames were identified and a partial gene sequence of PFP-β, including the 5' untranslated region, was obtained. The sequences were compared to the non-redundant (NR) nucleotide sequence database by BLASTN, and to the deduced amino acid database by BLASTX (Altschul et al., 1990). A restriction map of the sequence was obtained and was used in the construction of the expression vectors.

Comparative sequence analysis of the obtained sequences and that of castor bean and potato PFP-β was done using the multiple alignment (Higgins-Sharp) function of DNASIS for Windows Version 2.0.

5.3.12 Isolation of the 5' flanking region of sugarcane PFP-β
Analysis of the obtained sequence upstream of the translation initiation codon, ATG, was used to design three primers, namely PROM-F, PROM-R and PFP-E2 (Table 5.1), for the specific amplification of two fragments in this region. PROM-F was used in combination with PROM-R to amplify a 1 088 bp fragment containing the 5' flanking region excluding the ATG codon, and with PFP-E2 to amplify a 1 380 bp fragment containing this 1 088 bp region as well as the first intron, the first exon and 17 bp of the second exon of PFP-β. The primers contained BamHI restriction sites for the cloning of these fragments.

The PCR mixtures were as described in 5.3.2, except for the use of 100 ng cosmid DNA from subclone K7 as template. Cycling parameters were as follows: 1 cycle of
94°C for 3 min, 55°C for 1 min and 72°C for 2 min; 30 cycles of 94°C for 1 min, 55°C for 2 min and 72°C for 3 min. The last cycle was followed by a 7 min extension at 72°C.

5.3.13 Construction of expression vectors with the GUS reporter gene

A promoterless vector, pGUS000, was used in the construction of expression vectors. For the determination of PFP-β promoter activity, transient reporter gene expression in sugarcane callus had to be quantified. Two expression vectors were constructed for this analysis:

a.) pGUS1072 was constructed by the cloning of the amplified, BamH I treated 1072 bp fragment (5.3.12) into the BamH I cloning site of pGUS000.

b.) pGUS1355 was constructed by cloning the BamH I treated 1355 bp fragment containing the promoter area, first exon and first intron of PFP-β (5.3.12) into the BamH I cloning site of pGUS000.

Standard cloning techniques were used in the construction of these vectors (Sambrook et al., 1989), and characterisation of the final constructs was done by restriction analysis.

5.3.14 Transient GUS expression in sugarcane callus

The following seven constructs were used for the transient GUS expression analysis of different promoters in sugarcane callus:

a.) pGUS520: Maize Ubi-1 promoter + GUS, b.) pAHC27: Maize Ubi-1 promoter + intron + GUS (Christensen et al., 1992), c.) pBl221: CaMV 35S promoter + GUS (Jefferson et al., 1987), d.) pMON19468: CaMV 35S promoter + intron + GUS, e.) pGUS 1072: PFP-β promoter + GUS, f.) pGUS1355: PFP-β promoter + intron + GUS and g.) pGUS000: Promoterless vector + GUS.

Particle bombardment of sugarcane callus was done as described in 4.3.3. Three bombardments were performed with each construct. Histochemical GUS assays were performed 48 h after bombardment as described in 4.3.6. Blue foci in the callus were counted after staining for 36 h.
5.4 RESULTS

5.4.1 Amplification of PFP-β from cDNA libraries

Amplification products were formed with 3 of the 4 primer combinations that were used in each of the cDNA libraries, as shown in Figure 5.1. No products were amplified using the primer combination B2/R.

![Figure 5.1](image_url)

Figure 5.1 Amplification of PFP-β cDNA from leafroll and mature culm cDNA libraries. Amplification was done using each of the degenerate primers, PFP-B1 (B1) and PFP-B2 (B2), in combination with one of the library vector-specific primers, T7 and pUC18/M13 reverse (R) primer. Lanes 3 to 6: amplification products from the leafroll cDNA library, lanes 7 to 10: amplification products from the mature culm library, lanes 3 and 7: B1/T7, lanes 4 and 8: B2/T7, lanes 5 and 9: B1/R, lanes 6 and 10: B2/R. The sizes of amplified fragments were determined by comparing them to marker 3 (lanes 1 and 12) and marker 5 (lanes 2 and 11). Amplified fragments that were isolated for further analysis, B1/R from the leafroll library and B2/T7 from the mature culm library, are indicated with arrows.

The sizes of the amplified fragments varied from 250 bp to 1 100 bp and the result confirmed that the PFP-β cDNA was present in these libraries in both orientations. The two largest fragments that could be reamplified, namely the 1 100 bp B1/R fragment from the leafroll library (Figure 5.1 lane 5) and the 800 bp B2/T7 fragment from the mature culm library (Figure 5.1 lane 8), were isolated for further characterisation.

The sequences of these fragments were compared to the international database. The only homologies that were reported were to the PFP-β genes of castor bean (75%), rice (84%) and potato (73%).
5.4.2 Screening the sugarcane mature culm cDNA library

Because PFP activity was shown to reach a peak in internode 7 of the sugarcane culm (Whittaker and Botha, 1996), the library from this tissue was screened for the PFP-β cDNA fragment. The amplified fragment, B2/T7 (5.4.1), which contained the most 5' region of the cDNA, was used as the probe.

A number of 250 000 pfu's of the library were screened. Three rounds of screening were completed and 3 positive clones, namely C3, C4 and C6, were finally isolated after the third round. After phagemid rescue, sequence analysis revealed that the 3 final clones all contained an incomplete PFP-β cDNA fragment of 1 800 bp. No initiation codon was present in the fragments and the reported size of PFP-β cDNA in castor bean is closer to 2000 bp (Carlisle et al., 1990). The inserts of clones C3 and C4 were cloned in the same orientation in pBluescript, while that of C6 was inserted in the opposite orientation. Clone C6 contained a second insert of 1100 bp. Sequence analysis was done and this insert was partially homologous to mitochondrial malate dehydrogenase. No further analysis of this fragment was done.

5.4.3 Southern and northern blot analysis

Different hybridisation patterns were observed for different plants after the Southern blot as shown in Figure 5.2. The sugarcane cDNA probe hybridised to the DNA of all the plants that were analysed.

All four sugarcane varieties (lanes 1 to 4) as well as Zea mays (lane 7) showed similar results. Two bands of 4 160 and 2 280 bp respectively were seen in these sample lanes. In the Solanum lane (lane 5) 2 bands of sizes 3 100 and 2 800 bp were seen, while bands of 2 500 and 1 500 bp respectively were present in the Ricinus lane (lane 6). The signals were of approximately the same strength.
Figure 5.2 Southern blot analysis of PFP-β in sugarcane and other plants. 20 µg of genomic DNA from different plants was digested with EcoR I and separated in a 0.8% (w/v) agarose gel. Lane 1: *Saccharum officinarum* (Black Cheribon), lane 2: *Saccharum spontaneum* (Coimbatore), lane 3: N21 (commercial sugarcane variety), lane 4: NCo310 (commercial sugarcane variety), lane 5: *Solanum tuberosum* (potato), lane 6: *Ricinus communis* (castor bean), lane 7: *Zea mays* (maize). The incomplete sugarcane PFP-β cDNA fragment (5.4.2) was used as probe.

The sugarcane PFP-β cDNA fragment hybridised strongly to an mRNA of approximately 2 300 bp that can be seen in the result of the northern blot in Figure 5.3.

Figure 5.3 Northern blot analysis of PFP-β in sugarcane leafroll. Twenty micrograms of total RNA was separated and the incomplete PFP-β cDNA fragment (5.4.2) was used as a probe. Lane 1: RNA ladder, lane 2: leafroll RNA. The sizes of the fragments are given in kb.

5.4.4 Screening of a sugarcane genomic library and characterisation of positive clones
A total number of $2 \times 10^8$ pfu's of the genomic library was screened in 3 sets of screening experiments consisting of 3 rounds each. The incomplete cDNA fragment isolated from the cDNA library clone C3 (5.4.2) was used as a probe. Four final positive clones, namely G2, G3, G4 and G9, were isolated.
Insert sizes of the positive clones were determined by Not I digestion and two types of clones were identified. Two fragments were formed from inserts in clones G2, G3 and G4, while Not I digestion removed a single fragment from the phage arms in clone G9 as shown in Figure 5.4 A.

Because both types of clones were isolated from the genomic library using the same probe, the possibility that the one type of clone contained the 5' end and the other type the 3' end of the PFP-β gene sequence, was investigated. A Southern blot of the 4 clones digested with the restriction enzyme Not I, was probed consecutively with 2 probes (5.3.9) located on different ends of the PFP-β cDNA fragment. The results are shown in Figure 5.4 B and C.

**Figure 5.4** Characterisation of positive genomic library clones by restriction and hybridisation analyses. A.) Lambda DNA from genomic library clones G2 (lane 2), G3 (lane 3), G4 (lane 4) and G9 (lane 5) was treated with Not I. Sizes of the insert fragments and phage arms were determined by comparing to marker 3 (lane 1) and 2 (lane 6). To clarify, a diagram of the fragments are shown next to the gel photograph. B.) Southern blot analysis of insert fragments obtained in (A) probed with 5'EIB340. C.) Southern blot analysis of insert fragments obtained in (A) probed with 3'EVEI650.
Hybridisation of 5'EIB340 to one of the two insert fragments of G2, G3 and G4, but not to G9 was observed. The other probe, 3'EVEI650, only hybridised to the insert fragment of G9. Clone G2, which contained the 5' end of the gene, was chosen for further analysis and the isolation of the promoter sequence.

5.4.5 Construction of a Deletion Factory™

The insert fragments of G2 were cloned into the cosmid pDELTA2 and the resulting cosmids containing parts of the coding sequence of PFP-β were identified by Southern blot analysis. The 5'EIB340 fragment (5.3.9) was used as a probe. The positively identified cosmid, pΔ2-PFPC, was 15 kb in length with an insert size of approximately 7 kb.

Deletion subclones were obtained after DF1 cells were transformed with pΔ2-PFPC. Clockwise deletion transformants (A) were selected on medium containing ampicillin and sucrose and a total of 9 clones were isolated. A total number of 32 counterclockwise deletion transformants (K) were selected on medium containing kanamycin and streptomycin.

The sequences of all the clones were analysed and overlapping sequences were identified by DNASIS analysis. The overlaps of the A-clone sequences were very large possibly because the clockwise deletion reactions were not efficient. The K-clones, on the other hand, showed sufficient overlapping and large fragments could be analysed.

Initially the K-clones were clustered together in 4 separate groups (CONA to COND) that did not overlap with each other. After the translation initiation codon was identified, the 5' upstream sequence was analysed and a specific primer, PFP-PROM1 (Table 5.1) was designed for the sequencing of an intermediate region (PROM) which linked 2 of the groups (CONC and COND) together. The overlapping fragment clusters as well as the sizes of the consensus sequences are shown in Figure 5.5.
Sequence analysis also revealed that CONB, CONC and COND contained parts of the coding sequence and the 5' flanking region of PFP-β as shown in Figure 5.5.

![Figure 5.5](image)

**Figure 5.5** Deletion Factory: Counterclockwise deletion subclones (K). A.) The 32 K-clones are clustered together in four groups, CONA to COND, according to deletion size. The specific primer, PFP-PROM1, was used to sequence an intermediate region (PROM, shown by an open arrow) which links CONC and COND together. B.) The exons that are included in the sequence are illustrated by dark boxes and introns by solid lines. The 5' flanking region is shown as a dashed line.

5.4.6 Comparative sequence analysis of the sugarcane, castor bean and potato PFP-β coding sequences

The first 5 exons of the PFP-β gene was identified (Appendix I) and the sequence was compared to that of castor bean and potato. Multiple base substitutions and in-frame insertions or deletions were observed. Most of the mutations that had an effect on the amino acid sequence occurs at the 5' end of the gene. Intron lengths and sequences are not conserved between these plants. For the region that was analysed, a 72.7% homology between sugarcane and castor bean and a 71.9% homology between sugarcane and potato was observed on nucleotide level. Comparison of the same
region in castor bean and potato showed a 75.4% homology. On amino acid level the observed homology in the same region was as follows: 77.9% homology between sugarcane and castor bean, 74.9% homology between sugarcane and potato and 78.7% homology between castor bean and potato.

5.4.7 Sequence analysis of the sugarcane PFP-\(\beta\) promoter

The complete sequence of a 1 747 bp fragment upstream of the ATG codon is shown in Figure 5.6.

![Sequence of 1 747 bp of the 5' flanking region of sugarcane PFP-\(\beta\).](image)

Useful restriction sites and known promoter motifs with their positions in the sequence are shown in Table 5.2.
Table 5.2  Restriction sites (A) and promoter motifs (B) in the promoter region of PFP-β.

A.

<table>
<thead>
<tr>
<th>Restriction endonuclease</th>
<th>Sequence</th>
<th>Cutting position</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acc I</td>
<td>GT/MKAC</td>
<td>1495</td>
</tr>
<tr>
<td>Dra I</td>
<td>TTT/AAA</td>
<td>614 787 1030 1223</td>
</tr>
<tr>
<td>EcoR I</td>
<td>G/AATTC</td>
<td>598 1014</td>
</tr>
<tr>
<td>Fok I</td>
<td>GGATGNNNNNNNNN</td>
<td>332</td>
</tr>
<tr>
<td>Fok I</td>
<td>/NNNNNNNNNNNNNNCATCC</td>
<td>1304</td>
</tr>
<tr>
<td>Hinf II</td>
<td>GTY/RAC</td>
<td>1496</td>
</tr>
<tr>
<td>Nco I</td>
<td>C/CATGG</td>
<td>653 1079</td>
</tr>
<tr>
<td>Nde I</td>
<td>CA/TATG</td>
<td>503 1616</td>
</tr>
<tr>
<td>Nhe I</td>
<td>G/CTAGC</td>
<td>36 40</td>
</tr>
<tr>
<td>Nru I</td>
<td>TCG/CGA</td>
<td>2026</td>
</tr>
<tr>
<td>Rsa I</td>
<td>GT/AC</td>
<td>643</td>
</tr>
<tr>
<td>Sal I</td>
<td>G/TCGAC</td>
<td>1494</td>
</tr>
<tr>
<td>Sma I</td>
<td>CCC/GGG</td>
<td>1909</td>
</tr>
<tr>
<td>Spe I</td>
<td>A/CTAGT</td>
<td>607 1023</td>
</tr>
<tr>
<td>Sph I</td>
<td>GCATG/C</td>
<td>1929</td>
</tr>
<tr>
<td>Taq I</td>
<td>T/CGA</td>
<td>1495 1876</td>
</tr>
</tbody>
</table>

B.

<table>
<thead>
<tr>
<th>Motif</th>
<th>Pattern</th>
<th>Position</th>
</tr>
</thead>
<tbody>
<tr>
<td>HSP-70.5</td>
<td>G-A-T-T-G-G</td>
<td>556 - 561</td>
</tr>
<tr>
<td>INF.1</td>
<td>A-A-G-T-G-A</td>
<td>1153 - 1158</td>
</tr>
<tr>
<td>MSP-CS</td>
<td>A-G-A-T-C-T</td>
<td>418 - 423</td>
</tr>
<tr>
<td>napB</td>
<td>C-A-T-G-C-A</td>
<td>1622 - 1627</td>
</tr>
<tr>
<td>NF-E1.5</td>
<td>T-A-T-C-T-T</td>
<td>1255 - 1260</td>
</tr>
<tr>
<td>OCT-B2-SV40</td>
<td>C-T-T-G-C-A-T</td>
<td>2290 - 2296</td>
</tr>
<tr>
<td>SEF4</td>
<td>R-T-T-T-T-R</td>
<td>1006 - 1012, 1309 - 1315</td>
</tr>
<tr>
<td>UAS2UP1</td>
<td>T-G-A-T-G-G-T</td>
<td>555 - 562</td>
</tr>
</tbody>
</table>
5.4.8 Transient GUS expression analysis in sugarcane

Although sequence analysis of the 5' flanking region of PFP-β revealed the presence of promoter motifs (Table 5.2), the ability of this sequence to act as a promoter and drive foreign gene expression had to be determined. Due to the fact that the 5' flanking sequence was extremely AT-rich, an area had to be identified that would allow the design of a primer with a GC-content of approximately 50%. Such an area was located approximately 1 090 bp upstream from the ATG codon. The primers, PROM-F and PROM-R, were thus designed for the amplification of this portion of the 5' flanking region. Since the inclusion of intron sequences in promoter constructs has resulted in higher foreign gene expression levels in monocotyledonous plants (Mascarenhas et al., 1990; Maas et al., 1991), the first intron of PFP-β was included in one of the promoter constructs. The two expression vectors containing the promoter (pGUS1072) and the promoter plus the first intron (pGUS1355) of sugarcane PFP-β, respectively, are shown in Figure 5.7.

Figure 5.7 Schematic representation of the expression vectors, pGUS1072 and pGUS1355. A.) pGUS1072. The amplified and BamHI treated 1 072 bp fragment containing the putative PFP-β promoter (PFP PROM) was cloned into the BamHI cloning site of the promoterless vector, pGUS000. B.) pGUS1355. The 1 355 bp fragment containing the putative promoter as well as the first exon and first intron of PFP-β (PFP P + I), was treated with BamHI and cloned into pGUS000.
Blue foci were observed with all constructs tested, with the exception of the promoterless control, pGUS000. A variety of intensities of blue coloration ranging from extremely dark to pale blue was observed. Furthermore, blue foci produced from transient expression driven by both PFP promoter constructs were mostly located in the non-embryogenic callus (Taylor et al., 1992). With all the other constructs tested, blue foci were predominantly seen in the white embryogenic callus (Taylor et al., 1992). The results obtained in the transient expression assay were normalised against the levels of expression with pAHC27 (expressed as a percentage), and are shown in Figure 5.8. pAHC27 was used as reference since *Ubi-1* has previously been shown to be the most efficient promoter for driving GUS expression in sugarcane (Gallo-Meagher and Irvine, 1993).

**Figure 5.8** Transient GUS expression in sugarcane callus using different promoter constructs. Blue foci were counted after transformation of sugarcane callus. The result of pAHC27 was set to an arbitrary value of 100%, and the other results were calculated in comparison to this value. The constructs are described in 5.3.14.
Blue foci were observed with all constructs tested, with the exception of the promoterless control, pGUS000. A variety of intensities of blue coloration ranging from extremely dark to pale blue was observed. Furthermore, blue foci produced from transient expression driven by both PFP promoter constructs were mostly located in the non-embryogenic callus (Taylor et al., 1992). With all the other constructs tested, blue foci were predominantly seen in the white embryogenic callus (Taylor et al., 1992). The results obtained in the transient expression assay were normalised against the levels of expression with pAHC27 (expressed as a percentage), and are shown in Figure 5.8. pAHC27 was used as reference since Ubi-1 has previously been shown to be the most efficient promoter for driving GUS expression in sugarcane (Gallo-Meagher and Irvine, 1993).

Figure 5.8 Transient GUS expression in sugarcane callus using different promoter constructs. Blue foci were counted after transformation of sugarcane callus. The result of pAHC27 was set to an arbitrary value of 100%, and the other results were calculated in comparison to this value. The constructs are described in 5.3.14.
5.5 DISCUSSION

Two tissue-specific cDNA libraries from sugarcane were screened for PFP-β by PCR using degenerate primers. Amplification products were formed with most primer combinations indicating that the PFP-β cDNA is present in both the leafroll and the mature culm library in both orientations. The sizes of the amplified fragments also indicate that cDNAs of different sizes are present in the libraries, probably due to incomplete cDNA synthesis in some cases. Because PFP activity reaches a peak in internode 7 in the mature sugarcane plant (Whittaker and Botha, 1996), the mature culm cDNA library was screened for the full-length cDNA fragment. However, all the clones that were isolated from this library contained a 1 800 bp incomplete cDNA fragment. The absence of an initiation codon in the open reading frame and the fact that the reported size of PFP-β cDNA is close to 2 000 bp (Carlisle et al., 1990), revealed that the fragments were not complete. Although the full-length cDNA could not be obtained from this library, the isolated cDNA fragment was successfully used as probe in Southern and northern hybridisation analyses and for the screening of the genomic library.

Southern blot analysis of sugarcane genomic DNA revealed that PFP-β may be encoded by a single gene. Similar results have been reported for castor bean (Blakeley et al., 1992). This result is, however, surprising for sugarcane because of the complex polyploid nature of the plant. The monocotyledonous plants that were analysed, namely sugarcane and maize, showed the same hybridisation pattern when probed with PFP-β cDNA. Potato and castor bean showed different hybridisation patterns, and in all plants two bands were observed which may be an indication of the low copy number of the gene in these plants. Identical hybridisation patterns were observed for all sugarcane varieties that were analysed.

The presence of a single mRNA of approximately 2 300 bp was revealed by northern blot analysis. This result is consistent with northern blots from castor bean (Blakeley et al., 1992) and potato (Carlisle et al., 1990). The signal was strong and indicates a high expression level of the gene in leafroll. Expression levels in different
sugarcane tissues still have to be determined to confirm the results that were obtained in metabolic studies of PFP (Whittaker and Botha, 1996).

Two different types of clones were isolated from the genomic library and these contain the 5' and the 3' end of the PFP-β gene respectively. The present study is focused on the isolation of the promoter area and therefore the clones containing the 5' end of the gene were analysed first. The complete sequence of the gene will, however, be determined in the future since the sequence of sugarcane PFP-β is not yet available and will be of value for further studies.

Sequences of deletion subclones were analysed and the first five exons and the 5' untranslatable region of PFP-β were identified. The gaps that exist in the regions between CONB and COND, and between CONA and CONC (Figure 5.4) can be filled by analysing more deletion subclones. For the purpose of the present study, however, the available sequences were adequate. When the sequence of the coding region that was contained in these clones was compared to the same region in castor bean and potato genes, a homology of approximately 72% was observed.

A portion of 1 747 bp of the 5' flanking region of PFP-β has so far been sequenced and the presence of promoter motifs has been indicated. In the absence of data on the binding of nuclear factors in relation to PFP expression, the significance of these motifs cannot be determined before deletion analyses of the 5' flanking region have been carried out. For the present study, however, the question as to whether the 5' flanking region of the isolated PFP-β gene was a functional promoter or not had to be answered. The possibility existed that the 5' flanking region that was analysed would not drive high levels of reporter gene expression. This was taken into account by including the first intron of PFP-β with the putative promoter region in pGUS1355 to possibly enhance reporter gene expression. Introns of some genes have been postulated to stimulate gene expression, for example the addition of an intron to the 5' untranslated region of the maize Sh1 gene increased reporter gene expression approximately 100-fold (Maas et al., 1991).
Results obtained from transient GUS expression analysis indicated that the 5' flanking region of PFP-β that was analysed was capable of driving foreign gene expression in sugarcane callus. The expression levels obtained were very low in comparison to those obtained with the Ubi (pAHC27) and 35S (pMON19468) promoters, which included introns in the constructs. However, the difference in expression levels was not as pronounced when compared to the constructs containing the Ubi and 35S promoters without the introns, pGUS520 and pBI221, respectively. Expression levels obtained with pAHC27 and pBI221 in callus can be compared to a previous report where the same constructs were used for transient GUS expression in expanding sugarcane leaves (Gallo-Meagher and Irvine, 1993).

Expression levels obtained with the construct containing the first exon and intron (pGUS1355) were slightly lower than with the construct containing only the promoter (pGUS1072). This result may be explained by the possible presence of cis-active regulatory elements in the first exon of PFP-β that could influence gene expression. However, in a study where the effects of promoter, intron and enhancer elements on transient expression in sugarcane protoplasts were investigated, the potential influence of precise sequence context around regulatory elements was emphasised. It was reported that the inclusion of the Adh1 intron in constructs containing either the nos or the 35S promoter caused a 30 - 70% reduction in GUS expression in sugarcane (Rathus et al., 1993), which was in contrast to the trend from cereals (Mascarenhas et al., 1990). Since the source, size and spatial position of the intron, as well as the combination of the intron enhancer with a surrounding exon region have been shown to strongly influence gene expression (Maas et al., 1991), further analyses with constructs containing the first PFP-β intron should be carried out. It would appear from the results that the expression of the PFP-β promoter was predominant in non-embryogenic callus (Taylor et al., 1992). The implications of this observation can be determined by testing expression in other sugarcane tissues.

Now that it has been established that the 5' flanking region of PFP-β is a functional promoter, further characterisation can be done to determine the sequences required
for optimal activity and possible tissue-specificity. The stability of gene expression driven by this promoter must also be determined in transgenic sugarcane.

5.6 CONCLUSIONS

The gene of PFP-β was isolated from sugarcane, and the sequences of the first five exons and 1 747 bp of 5' flanking region were determined. Transient expression analysis of a reporter gene in sugarcane callus indicated that the 5' flanking region of the isolated gene is a functional promoter that can be used to drive foreign gene expression.
CHAPTER 6
CONCLUSIONS

Two different approaches were followed in the present study to isolate tissue-specific promoter sequences from sugarcane. Due to the limited knowledge of the sugarcane genome, and the fact that very few sugarcane genes have been characterised, a non-directed approach was initially followed to obtain differentially expressed genes.

The original RNA differential display method (Liang and Pardee, 1992) was modified to facilitate the rapid screening of cDNA populations with a large number of primers. Although all four putative culm-preferential fragments that were isolated from the sugarcane culm were shown to be false positives, several important suggestions regarding the differential display technique can be made from the results obtained and from reports in the literature.

Smears resulting from non-specific amplification reactions using an oligo(dT) primer can be eliminated by removing the oligo(dT) primer after reverse transcription, and using random primers for PCR. Alternatively, reverse transcription primed with a simple oligo(dT) primer, and subsequent PCR using an anchored dT primer together with random decamers, can increase the specificity of the amplification reactions. The cDNA template concentration for RT-PCR was shown to be very important for the generation of reproducible amplification profiles. Two identical samples from each tissue should be amplified to verify the reproducibility of the banding pattern. It should also be noted that differential display is not a quantitative method under the low stringency PCR conditions that are used. Longer primers and higher annealing temperatures could increase the sensitivity of the technique.

In future, other techniques could be evaluated for the isolation of differentially expressed genes from sugarcane. Genes that are differentially expressed during fruit ripening in strawberries were identified recently by using a procedure that combines a differential screening technique with a Southern blot screening of a subtractive
library by means of PCR (MAST-PCR-SBDS). The cDNA library was generated by using a modified magnet-assisted subtraction technique (MAST) (Medina-Escobar et al., 1997). The authors describe the technique as being a sensitive, easy, low-cost and reliable means of obtaining differentially expressed genes. In contrast to the small fraction (10 to 20%) of fragments identified by the differential display technique that turn out to be true differentially expressed genes, 31 of the 45 clones (69%) that were characterised by MAST-PCR-SBDS were positive (Medina-Escobar et al., 1997). The results described in the report, and the fact that it was developed in a plant system, make this technique attractive for the identification of differentially expressed genes in sugarcane.

The *Agrobacterium rhizogenes* rolC promoter was shown to drive phloem-specific reporter gene expression in sugarcane culm tissue. The specific cell type in the phloem where expression occurs, and the stability of gene expression under control of the rolC promoter in ratoon crops and under field conditions still have to be investigated. The reporter gene expression pattern observed in this study could be directly compared to the sucrose gradient, which suggested that sucrose influences rolC promoter activity in sugarcane. The advantage of a promoter that is activated by high sucrose levels is that it will only drive expression in plant tissues where the sucrose levels are high enough for its activation. The properties of the rolC promoter that have been observed in the present study have shown that it could be a very useful promoter in future pest and disease resistance programmes, as well as in metabolic studies in sugarcane.

The GUS and mGFP4 reporter genes were used with success in mature sugarcane tissue. Although a staining period of 48 h was required for the visualisation of GUS expression driven by the rolC promoter, clear results were nevertheless obtained. The increased staining time may be a reflection of the promoter strength. Clearing of pigmented tissue with CLP was shown to increase the coloration in tissues where GUS was expressed. GFP expression could only be visualised when an appropriate filter set and microscope were used for fluorescence microscopy because of the
presence of autofluorescent and light scattering components in the plant tissues. Several other modified versions of GFP are available and should be evaluated for use as reporter genes in sugarcane.

A directed approach was followed to isolate the gene and promoter of PFP-β from sugarcane. Screening of cDNA libraries using PCR revealed that PFP is present in the leafroll and in internode 7 of the sugarcane culm. Since no other sugarcane cDNA libraries were available, the presence of the transcript in other tissues was not investigated. Northern hybridisation analysis or RT-PCR of different sugarcane tissues can be done to determine whether PFP is expressed in all tissues, and the levels of expression can be compared to results obtained from metabolic studies. Southern blot analysis revealed that PFP is a low copy number gene in sugarcane, castor bean, potato and maize. PFP-β mRNA in sugarcane was shown to be 2.3 kb in length, which is also the reported length of the transcripts in castor bean and potato.

Two types of clones containing different portions of the PFP-β gene were isolated from the sugarcane genomic library. The complete sequence of the PFP-β gene from sugarcane has not been available, but can now be determined from the genomic library clones. Apart from the first five exons of the gene, 1747 bp of the 5' flanking area of the PFP-β gene was sequenced, and a number of promoter motifs were identified. Preliminary activity analysis revealed that the 5' flanking region of the isolated PFP-β gene is a functional promoter. Further analyses can now be done to characterise the promoter, and identify cis-acting sequences that are essential for optimal promoter activity. The tissue-specificity and stability of the promoter will be investigated by stable expression studies in different sugarcane tissues. Since the enzyme has been postulated to play a key role in sucrose breakdown in sugarcane, the effect of sucrose on promoter activity can also be investigated.

The two most significant contributions of the present study are the following:

- A functional promoter has been isolated from sugarcane.
The *Agrobacterium rolC* promoter can drive phloem-specific expression in the sugarcane culm. These elements may be used in future investigations into the genetic manipulation of sugarcane.
REFERENCES


ap Rees, T., Green, J.H. and Wilson, P.M. 1985 Pyrophosphate:fructose 6-phosphate 1-phosphotransferase and glycolysis in non-photosynthetic tissues of higher plants. Chemical Journal, 227: 229 - 304.


Delauney, A.J. and Verma, D.P.S. 1990 A soybean gene encoding Δ¹-pyrroline-5-carboxylate reductase was isolated by functional complementation in *E.coli* and is found to be osmoregulated. *Molecular and General Genetics*, 221: 299 - 305.


Guerrero, F.D. and Mullet, J.E. 1988 Reduction of turgor induces rapid changes in leaf translatable RNA. *Plant Physiology*, 88: 401 - 408.


Guivarc’h, A., Spena, A., Noin, M., Besnard, C. and Chriqui, D. 1996 The pleiotropic effects induced by the rolC gene in transgenic plants are caused by expression restricted to protophloem and companion cells. *Transgene Research*, 5: 3 - 11.


CURRICULUM VITAE

Sarita Groenewald (née Greeff) was born on 19 November 1968 in Johannesburg.

1986 - Matriculated at Hoërskool Riebeeck, Randfontein.

1989 - Obtained a Bachelor of Science degree with distinction, at Potchefstroom University for Christian Higher Education (PU for CHE), Potchefstroom, with majors in Biochemistry and Microbiology.

1990 - Obtained a Bachelor of Science (Honours) degree in Biochemistry, with distinction, at PU for CHE.

1993 - Obtained a Master of Science degree in Biochemistry with specialisation in Biotechnology, with distinction, at PU for CHE. Title of thesis: Cloning and expression of the human genes for GM-CSF and G-CSF.

1994 - Enrolled for a PhD in Plant Molecular Biology at the University of Natal, Pietermaritzburg. Experimental work was carried out in the Biotechnology Department at the South African Sugar Association Experiment Station, Mount Edgecombe.

1994 - Became the first recipient of the FRD Aaron Klug PhD scholarship for molecular biology.

APPENDIX I

Nucleotide sequence of the first five exons of the PFP-β gene from sugarcane.
The exons are highlighted and the introns are given in lower case. In introns 3 and 4 positions where sequences are still lacking are indicated with ///.