ISOLATION AND EVALUATION OF THE SUGARCANE UDP-GLUCOSE DEHYDROGENASE GENE AND PROMOTER

By

Jennie van der Merwe

Dissertation presented for the

Degree of Doctor of Philosophy (Plant Biotechnology)

Stellenbosch University

Supervisor: Prof. Frederik C. Botha

Co-Supervisor: Dr. Sarita Groenewald

December 2006

Declaration

I the undersigned, hereby declare that the work contained in this dissertation is my own original work and that I have not previously in its entirety or in part submitted it at any university for a degree.

December 2006

ACKNOWLEDGEMENTS

- My sincere thanks go to my supervisor, Prof. Frikkie Botha, for his guidance, encouragement and patience through the duration of this study. His enthusiasm for science was contagious.

- I would like to thank my co-supervisor, Dr. Sarita Groenewald for her support, scientific guidance and useful suggestions throughout this study.

- I thank the South African Sugar Association, the National Research Foundation, the Department of Trade and Industry, the Harry Crossley Foundation and the University of Stellenbosch for the financial support that made this study possible.

- Thank you to all the staff and students at the IPB for the moral support, technical assistance and friendship.

ABSTRACT

The young internodes of sugarcane are ideal targets for altering metabolism, through genetic manipulation, to potentially control known fungal diseases such as Smut or to increase sucrose yields in these regions that are currently being discarded. At present, no regulatory sequences that specifically drive transgene expression in young developing sugarcane tissues are available. The objective of this study was therefore to isolate and evaluate such a sequence. The promoter targeted for isolation in this study regulates the expression of UDP-glucose dehydrogenase (EC 1.1.1.22), an enzyme which catalyses the oxidation of UDP-glucose to UDP-glucuronic acid, a precursor for structural polysaccharides which are incorporated into the developing cell wall. A strong correlation between the expression of UDP-glucose dehydrogenase and a demand for structural polysaccharides in developing tissues could therefore be expected.

The first part of this study addressed the general practicality of promoter isolation from sugarcane, a complex polyploid. A gene encoding UDP-glucose dehydrogenase was isolated from a sugarcane genomic library. The gene contains an open reading frame (ORF) of 1443 bp, encoding 480 amino acids and one large intron (973 bp), located in the 5'-UTR. The derived amino acid sequence showed 88 – 98% identity with UDP-glucose dehydrogenase from other plant species, and contained highly conserved amino acid motifs required for cofactor binding and catalytic activity. Southern blot analysis indicates a low copy number for UDP-glucose dehydrogenase in sugarcane. The possible expression of multiple gene copies or alleles of this gene was investigated through comparison of sequences amplified from cDNA prepared from different tissues. Although five Single Nucleotide Polymorphisms (SNP) and one small-scale insertion/deletion (INDEL) were identified in the aligned sequences, hundred percent identity of the derived amino acid sequences suggested the expression of different alleles of the same gene rather than expression of multiple copies. The finding that multiple alleles are expressed to provide the required level of a specific enzyme, rather than the increased expression of one dominant allele, is encouraging for sugarcane gene and promoter isolation.

In the second part of the study the suitability of UDP-glucose dehydrogenase as a target for the isolation of a developmentally regulated promoter was investigated. The contribution of UDP-

glucose dehydrogenase to pentan synthesis, as well as the expression pattern and subcellular localisation of the enzyme in mature sugarcane plants was studied at the tissue and cellular level. Radiolabelling with positionally labelled glucose was used to investigate the relative contributions of glycolysis, the oxidative pentose phosphate pathway and pentan synthesis to glucose catabolism. Significantly (P=0.05) more radiolabel was released as CO^2 from [6-^{14}C]-glucose than [1-^{14}C]-glucose in younger internodes 3, 4 and 5, demonstrating a significant contribution of UDP-glucose dehydrogenase to glucose oxidation in the younger internodes. In addition, there was significantly (P=0.05) more radiolabel in the cell wall (fiber) component when the tissue was labelled with [1-^{14}C]-glucose rather than [6-^{14}C]-glucose. This also demonstrates a selective decarboxylation of glucose in position 6 prior to incorporation into the cell wall and is consistent with a major role for UDP-glucose dehydrogenase in cell wall synthesis in the younger internodes.

Expression analysis showed high levels of expression of both the UDP-glucose dehydrogenase transcript and protein in the leafroll, roots and young internodes. *In situ* hybridisation showed that the UDP-glucose dehydrogenase transcript is present in virtually all cell types in the sugarcane internode, while immunolocalisation showed that the abundance of the protein declined in all cell types as maturity increased. Results obtained confirmed that this enzyme plays an important role in the provision of hemicellulose precursors in most developing tissues of the sugarcane plant, indicating that UDP-glucose dehydrogenase was indeed a suitable target for promoter isolation.

Lastly, the promoter region and first intron, located in the 5'-untranslated region (UTR) of this gene, were isolated and subsequently fused to the GUS reporter gene for transient expression analysis and plant transformation. Transient expression analysis showed that the presence of the intron was essential for strong GUS expression. Analysis of stably transformed transgenic sugarcane plants, evaluated in a green house trial, showed that the isolated promoter is able to drive GUS expression in a tissue specific manner under these conditions.

OPSOMMING

Die jong internodes van suikerriet is ideale teikens vir genetiese manipulering om sodoende bekende siektes soos suikerriet brand, te beheer, of om die suikerinhoud in hierdie weefsels wat tans weggegooi word, te verhoog. Daar is geen regulerende elemente of promotors, wat transgeenuitdrukking in jong ontwikkelende suikerrietweefsel kan aandryf, huidiglik beskikbaar nie. Die doel van hierdie studie was dus om so ‘n volgorde te isoleer en te evalueer. Die promotor wat in hierdie studie geteiken is, reguleer die uitdrukking van UDP-glukose dehidrogenase (EC 1.1.1.22), ‘n ensiem wat die oksidering van UDP-glukose na UDP-glukoroonsuur kataliseer. UDP-glukoroonsuur is ‘n voorloper vir strukturele polisakkariede wat in die ontwikkelende selwand geïnkorporeer word. ‘n Sterk korrelasie tussen die uitdrukking van UDP-glukose dehidrogenase en ‘n behoefte aan strukturele polisakkariede in ontwikkelende weefsels kan dus verwag word.

Die eerste gedeelte van hierdie studie ondersoek die praktiese aspekte verbonde aan promotorisolering uit suikerriet, ‘n komplekse poliploïed. ‘n Geen wat kodeer vir UDP-glukose dehidrogenase is uit ‘n suikerriet genomiese biblioteek geïsoleer. Hierdie geen bevat ‘n oop leesraam van 1443 bp, wat vir 480 aminosure kodeer, en een groot intron (973 bp) wat in die 5’-ongetransleerde gebied geleë is. Die afgeleide aminosuurvolgorde is 88 – 98% soortgelyk aan UDP-glukose dehidrogenases van ander plantspesies en bevat hoogs gekonserveerde motiewe wat vir kofaktorbinding en katalitiese aktiwiteit vereis word. ‘n Southern-hibridiseringsanalise het ‘n lae kopiegetal vir UDP-glukose dehidrogenase in suikerriet aangedui. Die moontlike uitdrukking van veelvoudige kopieë of allele van hierdie geen, is ondersoek deur volgordes wat geamplifiseer is uit kDNS afkomstig van verskillende weefsels, te vergelyk. Alhoewel daar vyf enkel-nukleotied polimorfismes en een kleinskaalse invoeging/delesie geïdentifiseer is, was die afgeleide aminosuurvolgorde van die geamplifiseerde fragmente identies. Die uitdrukking van verskillende allele en nie verskillende kopiëe van die geen, is dus hiermee bevestig. Die bevinding dat vereisde ensiemvlakke eerder dmv die uitdrukking van veelvoudige allele as deur die verhoogde uitdrukking van ‘n enkele alleel bereik word, is bemoedigend vir die isolering van promotors en geenvolgordes uit suikerriet.

In die tweede gedeelte van die studie word die geskiktheid van UDP-glukose dehidrogenase as teiken vir die isolering van ‘n jong weefselspesifieke promotor, ondersoek. Die bydrae van UDP-glukose dehidrogenase tot pentaansintese sowel as die uitdrukkingspatroon en subsellulêre lokalisering van die ensiem in volwasse suikerrietplante, is bepaal. Glukose wat radioaktief gemerk is in verskillende posisies, is gebruik om die relatiewe bydraes van glikolise, die oksidatiewe pentosefosfaatweg en pentaansintese tot die katabolisme van glukose te bepaal. In die jonger internodes, 3, 4 en 5, is beduidend (P=0.05) meer radioaktiwiteit vanaf [6-^{14}C]-glukose as [1-^{14}C]-glukose in CO^2 vrygestel, wat op ‘n aansienlike bydrae van UDP-glukose dehidrogenase tot die oksidering van glukose in jonger internodes dui. Daar was ook beduidend (P=0.05) meer radioaktiwiteit in die selwandkomponent (vesel) waar die weefsel eerder met [1-^{14}C]-glukose as [6-^{14}C]-glukose gemerk is. Daar vind dus ‘n selektiewe dekarboksilering van glukose in posisie 6 voor inkorporering in die selwand plaas, wat op ‘n belangrike rol vir UDP-glukose dehidrogenase in selwandsintese in jong internodes dui.

Hoë uitdrukkingsvlakke van beide die UDP-glukose dehidrogenase geentranskrip en -proteïen is dmv uitdrukkingsanalises in die blaarrol, wortels en jong internodes bevestig. *In situ*-hibridisering het gewys dat die UDP-glukose dehidrogenase transkrip in feitlik elke seltipe in die suikerriet internode teenwoordig is. Immunolokalisering het verder aangedui dat die hoeveelheid proteïen met toenemende volwassenheid in alle seltipes afneem. Hierdie resultate bevestig dat die UDP-glukose dehidrogenase ensiem ‘n belangrike rol in die verskaffing van voorlopers vir hemisellulose in meeste ontwikkelende weefsels van die suikerrietplant speel. Dit beteken dat die promotor van hierdie geen ‘n geskikte teiken vir die doel van hierdie studie was.

Laastens is die promotor en eerste intron van hierdie geen geïsoleer. Promotoraktiwiteit is op beide tydelike sowel as stabiele vlakke geëvalueer deur van die GUS-verklikkergeen gebruik te maak. Tydelike uitdrukkingsanalises het gewys dat sterk GUS-uitdrukking van die teenwoordigheid van die intron afhanklik was. Analise van stabiel-getransformeerde transgeniese suikerrietplante het verder aangedui dat die geïsoleerde promotor onder glashuiskondisies, GUS-uitdrukking op ‘n weefselspesifieke manier kon reguleer.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

The South African sugar industry is one of the world's leading producers of high quality sugar. Current commercial sugarcane varieties are obtained through a multi-stage selection scheme over a period of approximately 10 years to identify a few elite clones in a very large group of seedlings. Some elite clones have to be abandoned because of a single fault such as disease susceptibility. Genetic transformation can correct single faults in elite cultivars, possibly by the insertion of a single gene to complement, rather than replace traditional breeding methods. It can also provide a better understanding of the role of specific sugarcane genes in complex processes such as sugar accumulation, and can introduce valuable novel genes for new properties in sugarcane.

Although the transformation of sugarcane is well established (Birch and Franks, 1991), a major obstacle limiting progress in this area is the availability of promoters. An absolute prerequisite for the use of genetic engineering for sugarcane varietal improvement is the stable and predictable expression of introduced genes. Very simply, gene expression is controlled by promoter sequences, generally located immediately upstream of the coding region, which determine the strength, developmental timing and tissue specificity of expression of the adjacent coding region (a detailed discussion of the regulation of gene expression is presented in Chapter 2). The shortage of such promoter sequences, as well as patent considerations (Birch, 1997), has made it necessary to isolate novel promoters that could be used for sugarcane transformation.

Several promoters that direct near-constitutive expression in monocotyledonous plants have been isolated. These include promoters isolated from plants, such as the maize polyubiquitin (*ubi-1*) promoter (Christensen and Quail, 1996) and the rice actin (*Act1*) promoter (McElroy *et al*., 1990), and viral promoters such as the cauliflower mosaic virus (CaMV) 35S promoter (Benfey *et al*., 1990; Terada and Shimamoto, 1990), sugarcane bacilliform badnavirus promoter (Tzafrir *et al*., 1998), and promoters isolated from the banana streak badnavirus (Schenk *et al*., 2001). Though constitutive expression of a transgene may sometimes be required, targeting the

expression to a specific tissue where the action of the transgene is required will greatly decrease the metabolic load resulting from transformation.

When sugarcane is harvested, the top internodes are traditionally discarded due to low juice purity. Increasing the sucrose content in these internodes, thereby providing additional tissue from which sucrose can be extracted, could result in an increased sucrose yield per cane. Increasing sucrose yields is one of the main goals of sugarcane breeders. The top of the cane is also the point of infection for Smut, the most important fungal disease of sugarcane in South Africa. Genetic manipulation has the potential to alter metabolism in these tissues to increase the sucrose yield, or to control Smut, possibly through the insertion of a single gene. The main aim of this study was therefore to provide a promoter, which could be used to regulate the expression of a transgene exclusively in developing sugarcane tissues.

One possible approach to obtain promoters which direct specific levels and distribution of expression is to identify endogenous genes already expressed in the desired pattern in the organism targeted for transformation, in this instance, sugarcane. The corresponding promoter can then be isolated from the genome of the target organism. Following this approach, the promoter of a gene encoding uridine 5-diphosphate-glucose dehydrogenase (UDP-glucose dehydrogenase) was selected as a potential target for promoter isolation in this study, based on what is known about the function of the enzyme that it encodes.

UDP-glucose dehydrogenase (EC 1.1.1.22) catalyses the oxidation of UDP-glucose to UDP-glucuronic acid (Nelsestuen and Kirkwood, 1971), a precursor for sugar nucleotides which are incorporated into pectin and hemicelluloses. Both pectin and hemicellulose are key components of cell walls, providing a matrix that strengthens the cell wall structure (Gibeaut, 2000). As UDP-glucose dehydrogenase is required for growth and development, the promoter of this gene could possibly be used to drive transgene expression in young developing tissues.

Promoter isolation is technically difficult in most species. In sugarcane this process is further complicated by a highly polyploid genome. Modern sugarcane cultivars (*Saccharum* spp. Hybrids) appear to have a basic chromosome number of 10 and 2n chromosome numbers of

between 100 and 130 (Butterfield *et al*., 2001). This implies that for each single copy of a gene, up to ten alleles can be present. It is not currently known, however, whether all these alleles are expressed. It is possible, even likely, that some of the gene copies have accumulated sequence changes inhibiting their expression. As a result the sugarcane genome may contain many sequences that represent silent copies of a specific gene, adjacent to non-functional promoters. The first part of this study addressed this potential problem by investigating the possible expression of multiple alleles of the gene targeted for promoter isolation, and thereby the general practicality of promoter isolation from sugarcane and other polyploids. Chapter 3 describes the isolation of a gene encoding UDP-glucose dehydrogenase from a sugarcane genomic library, and provides evidence for the simultaneous expression of distinct alleles of this gene in a single sugarcane plant. The finding that multiple alleles are expressed to provide the required level of a specific enzyme, rather than the increased expression of one dominant allele, is encouraging for sugarcane gene and promoter isolation.

The next part of the study investigated the suitability of UDP-glucose dehydrogenase as a target for the isolation of a developmentally regulated promoter. The distribution of the target gene in different tissue types, and different cell-types within a specific tissue will determine the usefulness of the promoter for transgene expression. A strong correlation between the expression of UDP-glucose dehydrogenase and a demand for structural polysaccharides in tissues that are actively synthesising cell walls, has been reported for several plant species (Tenhaken and Thulke, 1996; Seitz *et al*., 2000; Johansson *et al*., 2002). UDP-glucose dehydrogenase has previously been purified from rapidly expanding culm tissues of sugarcane (Turner and Botha, 2002). Although the kinetic properties of the sugarcane enzyme were studied, no information is currently available about the distribution of the enzyme in sugarcane. It was previously shown that significant levels of UDP-glucose are present in the sugarcane culm (Whittaker and Botha 1997). However, to date, most carbon partitioning research in sugarcane has focussed on the accumulation of sucrose and partitioning within the sugar pool, and little attention has been paid to the allocation of carbon to structural components such as the cell wall. In Chapter 4 the role of UDP-glucose dehydrogenase in pentan synthesis in younger and more mature internodes was investigated. In addition, the distribution of the enzyme in different cell types present in the sugarcane internode was examined by *in situ* hybridisation, while immunolocalisation in

internodal sections from different developmental stages was used to investigate the abundance of the protein as tissue maturity increases. Results obtained in this part of the study indicated that UDP-glucose dehydrogenase was indeed a suitable target for promoter isolation.

Having found that promoter isolation from a complex polyploid such as sugarcane is viable, and that the proposed target gene was expressed in the desired pattern in sugarcane, the next part of the study involved the isolation of the promoter sequence adjacent to the isolated UDP-glucose dehydrogenase coding region. Chapter 5 describes the characterisation of this promoter. The isolated sequence was evaluated for its ability to drive transgene expression in a transient system and stably transformed sugarcane. It was demonstrated that an active promoter, able to drive highly tissue specific expression in transgenic sugarcane, was isolated. Also, the sequence of the promoter was investigated through computer analysis for possible clues relating to the regulation of the expression of UDP-glucose dehydrogenase. As this study presents the first demonstrated isolation of a developmentally regulated promoter from sugarcane, valuable knowledge about the regulation of gene expression in sugarcane can be gained.

REFERENCES

Benfey, P. N., Ren, L., and Chua, N.-H. 1990, Combinatorial and synergistic properties of CaMV 35S enhancer subdomains, *EMBO Journal*, **9** (6): 1685-1696.

Birch, R. G. 1997, Plant transformation: problems and strategies for practical application, *Annual Review of Plant Physiology and Plant Molecular Biology*, **48**: 297-326.

Birch, R. G. and Franks, T. 1991, Development and optimisation of microprojectile systems for plant genetic transformation, *Australian Journal of Plant Physiology*, **18**: 453-469.

Butterfield, M., D'Hont, A., and Berding, N. 2001 The sugarcane genome: A synthesis of current understanding, and lessons for breeding and biotechnology. *Proc Soc Afr Sugarcane Technol Ass*, **75**: 1-5.

Christensen, A. H. and Quail, P. H. 1996, Ubiquitin promoter-based vectors for high-level expression of selectable and/or screenable marker genes in monocotyledonous plants, *Transgenic Research*, **5**: 213-218.

Gibeaut, D. M. 2000, Nucleotide sugars and glycosyltransferases for synthesis of cell wall matrix polysaccharides, *Plant Physiology and Biochemistry*, **38**: 69-80.

Johansson, H., Sterky, F., Amini, B., Lundeberg, J., and Kleczkowski, L. A. 2002, Molecular cloning and characterization of a cDNA encoding poplar UDP-glucose dehydrogenase, a key gene of hemicellulose/pectin formation, *Biochimica et Biophysica Acta-Gene Structure and Expression*, **1576** (1-2): 53-58.

McElroy, D., Zhang, W., Cao, J., and Wu, R. 1990, Isolation of an efficient actin promoter for use in rice transformation, *Plant Cell*, **2**: 163-171.

Nelsestuen, G. L. and Kirkwood, S. 1971, The mechanism of action of uridine diphosphoglucose dehydrogenase, *Journal of Biological Chemistry*, **246** (12): 3828-3834.

Schenk, P. M., Remans, T., Sagi, L., Elliott, A. R., Dietzgen, R. G., Swennen, R., Ebert, P. R., Grof, C. P., and Manners, J. M. 2001, Promoters for pregenomic RNA of banana streak badnavirus are active for transgene expression in monocot and dicot plants, *Plant Molecular Biology* **47** (3): 399-412.

Seitz, B., Klos, C., Wurm, M., and Tenhaken, R. 2000, Matrix polysaccharide precursors in Arabidopsis cell walls are synthesized by alternate pathways with organ-specific expression patterns, *Plant Journal*, **21** (6): 537-546.

Tenhaken, R. and Thulke, O. 1996, Cloning of an enzyme that synthesizes a key nucleotide-sugar precursor of hemicellulose biosynthesis from soybean: UDP-glucose dehydrogenase, *Plant Physiol*, **112** (3): 1127-1134.

Terada, R. and Shimamoto, K. 1990, Expression of CaMV 35S-GUS gene in transgenic rice plants, *Molecular and General Genetics*, **220**: 389-392.

Turner, W. and Botha, F. C. 2002, Purification and kinetic properties of UDP-glucose dehydrogenase from sugarcane, *Archives of Biochemistry and Biophysics*, **407** (2): 209-216.

Tzafrir, I., Torbert, K. A., Lockhart, B. E., Somers, D. A., and Olszewski, N. E. 1998, The sugarcane bacilliform badnavirus promoter is active in both monocots and dicots, *Plant Molecular Biology*, **38**: 347-356.

Whittaker, A. and Botha, F. C. 1997, Carbon partitioningduring sucrose accumulation in sugarcane internodal tissue, *Plant Physiology*, **115**: 1651-1659.

CHAPTER 2

LITERATURE REVIEW

2.1 INTRODUCTION

The successful transformation of any organism is dependent on reliable methodology for the introduction of a foreign gene, and the predictable and stable expression of such a gene. This chapter will provide a short overview of plant transformation, with a specific focus on progress made towards the successful transformation of sugarcane. Transformation strategies, as well as the availability of regulatory sequences will be briefly discussed.

In addition to regulatory sequences, transcribed sequences, including the 5'- and 3'-untranslated region, introns and the coding region, also appear to play an important role in the regulation of gene expression, particularly in monocotyledonous plant species. The occurrence and importance of these sequences, and specifically the role of introns in the regulation of gene expression will be discussed in this chapter.

The main aim of this study was to identify, isolate and characterise a promoter that could be used to regulate the expression of a foreign gene in developing sugarcane tissues. The approach followed was firstly to identify an endogenous gene expressed in the desired pattern, and then to isolate the corresponding promoter from the sugarcane genome. The promoter of a gene encoding UDP-glucose dehydrogenase was selected for isolation, based on what is known about the function of the enzyme it encodes. UDP-glucose dehydrogenase catalyses the oxidation of UDP-glucose to UDP-glucuronic acid, a precursor for structural polysaccharides which are incorporated into the developing cell wall. This enzyme was previously purified from rapidly expanding culm tissues of sugarcane (Turner and Botha, 2002). The promoter of this gene provides a possible candidate for a regulatory sequence specific for young developing tissues. Some background information about this enzyme is therefore also included in this chapter.

2.2 PLANT TRANSFORMATION

Plant transformation is an essential research tool in plant biology (Stitt and Sonnewald, 1995) and a practical tool for cultivar improvement (Birch, 1997; Newell, 2000; Mendoza, 2002). The requirements for the successful transformation of any organism include dependable methodology for the introduction of a foreign gene, and the predictable and stable expression of such a gene.

2.2.1 Direct transformation

Many methods have been devised to introduce DNA into plant cells. There are two types of gene transfer systems: direct gene transfer in which naked DNA is introduced into cells via any physical and/or chemical treatment, and indirect gene transfer in which another organism is used as a vector to effect the transfer and/or integration. Several direct gene transfer methods have been developed to transform plant species. The most popular of these seems to be microprojectile bombardment, which involves high velocity acceleration of microprojectiles carrying foreign DNA, penetration through the cell wall and membrane by the microprojectile, and the delivery of the associated DNA into plant cells. The method of microprojectile bombardment has demonstrated its broad utility and appears to be effective for all plant species tested so far (for reviews see Vasil, 1994; Casas *et al.*, 1995; Birch, 1997; Maenpaa *et al*., 1999; Taylor and Fauquet, 2002; Lorence and Verpoort, 2004). Other direct gene transfer methods include electroporation, infiltration, and microinjection (reviewed by Newell, 2000; Rakoczy-Trojanowska, 2002).

An advantage of direct gene transfer is that any piece of DNA may be transferred without using specialised vectors. Direct transformation is also very useful for transient expression analysis. When stable transformation is not the objective, a transgene can be transcribed in the nucleus and translated in the cytosol, independent of integration of the transgene into the host nuclear genome. Such transient expression can be used, for example, for promoter analysis (Rathus *et al*., 1993; Wei *et al*., 1999; Atienzar *et al*., 2000; Basu *et al*., 2003; Ono *et al*., 2004). A drawback of direct genetic transformation is that transformed cells will often contain multiple insertions of the transgene of interest, as well as fragmented copies of the transgene and vector (Pawlowski and Somers, 1996; Makarevitch *et al*., 2003). Multiple insertions could lead to co-suppression (Matzke and Matzke, 1995; Wu and Morris, 1999).

2.2.2 Indirect transformation

Indirect gene transformation relies almost exclusively on the use of the soil bacterium, *Agrobacterium tumefaciens*. This bacterium has the natural ability to transfer a particular DNA segment (transferred DNA or T-DNA) into the nucleus of an infected cell, where it is then stably integrated into the host genome and transcribed (Binns and Tomashow, 1988). Initial studies of the T-DNA transfer process to plant cells demonstrated that foreign DNA placed between the T-DNA borders could be transferred into plant cells, regardless of the origin of the DNA. This allowed for the first vector and bacterial strain systems for plant transformation to be developed (for review, see Hooykaas and Shilperoort, 1992).

Monocotyledonous plants, and particularly graminaceous crop species, were initially considered to be outside the *Agrobacterium* host range, since these plants are not natural hosts for this bacterium (Binns and Tomashow, 1988; DeCleene, 1985; Potrykus, 1990). Only a few years later *Agrobacterium*-mediated transformation of maize (Gould *et al*., 1991; Ishida *et al*., 1996) and rice (Chan *et al*., 1993; Hiei *et al*., 1994) proved that this was not true. Since then, many other monocotyledonous species have successfully been transformed following *Agrobacterium*-mediated gene transfer (Komari *et al*., 1998). The use of *Agrobacterium tumefaciens* in plant transformation and the molecular mechanisms involved have been reviewed extensively (De la Riva *et al*., 1998; Gelvin, 2000; Tzfira and Citovsky, 2000; Zupan *et al*., 2000; Tzfira and Citovsky, 2002; Gelvin, 2003; Tzfira *et al*., 2004).

Agrobacterium-mediated transformation has some advantages over direct transformation methods. Relatively large segments of DNA can be transferred with little rearrangement, and integration of low numbers of gene copies occurs in the host genome (Ishida *et al*., 1996). A disadvantage of *Agrobacterium*-mediated transformation is that the remaining infecting bacteria must be removed after transformation (for review see Lorence and Verpoort, 2004).

2.2.3 Regulation of transgene expression

To efficiently introduce a foreign gene to a plant, or to manipulate a metabolic process, the gene must be expressed in a suitable and predictable manner (Birch, 1997). Successful transformation is therefore, to a large extent, dependent on the availability of different promoters to achieve specific or induced expression. Practical application of most potentially useful new genes will require not only sustained expression without silencing over many vegetative generations, but also tailored levels and developmental or inducible patterns of expression, appropriate to the desired effect of the transgene product (Laporte *et al*., 2001).

Appropriate genetic constructs containing a promoter, the transgene and a terminating signal (An and Kim, 1993), are required to facilitate the integration and expression of foreign DNA in plants. Promoters are regions within a genome, located upstream of a gene transcription start site. Promoter elements determine the transcription initiation point, transcription specificity and rate. Previous studies suggest that promoters are constructed as a linear array of promoter elements, or *cis*-acting elements, each recruiting different transcription factors. Depending on the distance from the transcription initiation site, these elements form part of the 'proximal' or 'distal' promoter. Both proximal and distal promoters contribute to the process of cell-, tissue-, developmental stage-, and organ-specific regulation of transcription (for a review on promoter structure, see Guilfoyle, 1997; Lefebvre and Gellatly, 1997).

Traditionally, promoter elements were identified by fusing the putative promoter region to a reporter gene, such as GUS, and then making a deletion series of the promoter driving expression of the reporter gene (An and Kim, 1993). After transformation and determination of the expression level and pattern of the reporter gene, promoter regions required for regulation of transcription are identified. A good example of such an analysis is the dissection of the CaMV 35S promoter (Benfey *et al*., 1990). Use of deletion analysis has identified a whole array of plant promoter elements. Several public databases containing a collection of these *cis*-acting elements have been established, e.g. PlantProm DB (http://mendel.cs.rhul.ac.uk/mendel.php), PLACE (www.dna.affrc.go.jp/PLACE), PlantCARE (http://intra.psb.ugent.be:8080/PlantCARE), and TRANSFAC (http://www.gene-regulation.com/pub/databases.html). A new bioinformatics-based approach, which makes use of such databases in conjunction with motif-detection software

and an increasing number of large scale expression-profiling techniques, is fast replacing traditional deletion analysis. New promoter elements are being identified based on the hypothesis that the transcription of genes with a similar expression profile will be regulated by the same transcription factors. Computational, or *in silico*, analysis of the promoter regions of such genes is then used to identify over-represented elements. These and other related methods have recently been reviewed (Hehl and Wingender, 2001; Aarts and Fiers, 2003; Rombauts *et al*., 2003; Venter and Botha, 2004).

2.3 TRANSFORMATION OF SUGARCANE

During the twentieth century, highly productive sugarcane varieties with enhanced resistance to disease and insect pests were successfully developed in conventional breeding programs, but few modern crop varieties retain the same degree of resistance as exhibited by their wild relatives. Important traits, such as resistance to insect pests and herbicides, appear to be absent from the sugarcane parental germplasm. Also, many elite varieties produced by traditional breeding methods have to be abandoned due to a single "fault", such as susceptibility to a specific disease. The use of plant transformation methods to introduce new genes, and thereby new traits, into the sugarcane genome may have an important impact on sugarcane yields.

2.3.1 Transformation methodology

Methodology for the stable transformation of sugarcane is well established. In 1992, Rathus and Birch (1992) produced stably transformed sugarcane callus by electroporation of protoplasts, but no plants could be regenerated. A few months later, Bower and Birch (1992) reported the production of the first transgenic sugarcane plants by particle bombardment of embryogenic calli. Around the same time Arencibia and coworkers (1992) recovered transgenic sugarcane plants following electroporation of meristematic tissue. This group later also developed a method for sugarcane transformation by electroporation of intact cells (Arencibia *et al*., 1995). Some years later the first successful *Agrobacterium*-mediated transfer of DNA to sugarcane meristems was demonstrated (Enríques-Obregón *et al*., 1997; Arencibia *et al*., 1998). Since then, successful

Agrobacterium-mediated transformation of sugarcane callus (Elliott *et al*., 1998; Santosa *et al*., 2004) and axillary buds (Manickavasagam *et al*., 2004) has also been achieved.

Most sugarcane cultivars tested to date have yielded regenerable calli (Ingelbrecht *et al*., 1999), making the introduction of specific desirable traits directly into elite sugarcane varieties a realistic goal. The feasibility of using transformation technology to introduce specific genetic improvements into sugarcane is further demonstrated by the fact that all of the above-mentioned transformation methods have subsequently been used to introduce specific traits into sugarcane, specifically herbicide-resistance (Gallo-Meagher and Irvine, 1996; Enríques-Obregón *et al*., 1998; Snyman *et al*., 1998; Falco *et al*., 2000; Manickavasagam *et al*., 2004) and insect-resistance (Arencibia *et al*., 1997; Arencibia *et al*., 1999; Setamou *et al*., 2002; Tomov and Bernal, 2003).

2.3.2 Availability of regulatory sequences

Although the transformation of sugarcane is well established, a major obstacle limiting the use of genetic transformation for the varietal improvement of sugarcane is the availability of regulatory sequences, or promoters, to drive stable transgene expression. Most studies to date, including those mentioned above, have made use of three promoters to regulate the constitutive expression of the gene of interest and/or the selectable marker gene. These are the maize ubiquitin promoter (*Ubi-1*) (e.g. Gallo-Meagher and Irvine, 1996; Enríques-Obregón *et al*., 1998; Falco *et al*., 2000), the 35S promoter from the cauliflower mosaic virus (CaMV) (modified for use in sugarcane, see 2.4.1.4) e.g.(Arencibia *et al*., 1997; Elliott *et al*., 1998; Enríques-Obregón *et al*., 1998) and the artificial Emu promoter (e.g. Bower and Birch, 1992). This artificial promoter is made up of a truncated maize *adh*1 promoter with additional enhancer elements including six anaerobic response elements from the *adh*1 gene of maize and four ocs-elements from the *ocs* gene of *Agrobacterium tumefaciens* (Last *et al*., 1991).

Three other promoters have been shown to drive constitutive (or near-constitutive) expression of reporter genes in green house-grown transgenic sugarcane. These are the rice ubiquitin promoter (RUBQ2) (Liu *et al*., 2003) and two promoters (Cv and My promoters) derived from Australian banana streak badnavirus (Schenk *et al*., 2001), though expression from the My promoter was

relatively weak. Only one case of tissue-specific transgene expression in sugarcane has been reported to date. Preliminary analysis of a stem-specific promoter isolated from sugarcane demonstrated reporter gene expression in the top half of the stems of transformed plants (Hansom *et al*., 1999). These promoters, however, have not been widely utilised in subsequent studies.

Other promoters that could possibly be used for sugarcane transformation include the pPLEX series, derived from the subterranean clover stunt virus (SCSV) genome, modified for use in monocotyledonous plants (Schünmann *et al*., 2003b). Although these promoters have not been evaluated for their usefulness in sugarcane, their activity in other monocotyledonous species suggests that they may be active in a wide range of species. Another promoter that may be useful is the sugarcane badnavirus promoter (ScBV) (Tzafrir *et al*., 1998). This promoter was only evaluated in transformed rice, but as it was derived from a virus that infects both rice and sugarcane, it is very likely that this promoter would also be active in transgenic sugarcane.

Although the lack of regulatory sequences present a major obstacle hindering sugarcane transformation, the above-mentioned studies do prove that promoters derived from different sources, i.e. viral, artificial, closely related species and the sugarcane genome, can be used successfully to drive transgene expression in transgenic sugarcane and that these sources can be further exploited to obtain a wider range of regulatory sequences for sugarcane transformation.

2.3.3 Promoter silencing in sugarcane

Levels of transgene expression in transgenic plants are often unpredictable and many transgenic plants become silenced. Many promoters tested to date have been silenced in transgenic sugarcane, even though some of these were able to drive strong transient expression and expression in transformed callus. These include promoters isolated from different sources. The sugarcane polyubiquitin promoters, *Ubi4* and *Ubi9* (Wei *et al*., 1999), drove high-level GUS expression in sugarcane callus, but were silenced in regenerated plants (Wei *et al*., 2003). Interestingly, the ubi-9 promoter was active in transgenic rice. Two other promoters isolated from sugarcane were also silenced in transgenic sugarcane plants. A promoter from a peroxidase gene was not functional in callus or plants (Hansom *et al*., 1999). The promoter from a metallothionein gene (*Rsg*) was progressively silenced in transgenic callus, and silent in resulting

plants (Hansom *et al*., 1999). Silencing in transgenic sugarcane has also been reported for promoters isolated from other monocotyledonous plants. The rice actin promoter, previously shown to drive strong constitutive expression in transgenic rice (McElroy *et al*., 1990; Wang *et al*., 1992), was silenced when introduced into sugarcane (Hansom *et al*., 1999). The root-specific rice *RCg2* promoter, active in transgenic rice (Xu *et al*., 1995), was also silenced in sugarcane (Hansom *et al*., 1999).

These findings show that the use of homologous or heterologous promoters does not necessarily provide protection against transgene silencing in sugarcane. Also, promoters are not generally active in all monocotyledonous species. One study found that silencing in sugarcane is promoter-dependent and copy number independent (Hansom *et al*., 1999). Even when high copy numbers (>10 integration sites) of genes driven by the Maize *Ubi1* promoter integrated into the genome of sugarcane, there was no evidence of gene silencing, while other promoters, *viz*. the rice actin promoter and artificial Osa promoter, were silenced regardless of the copy number. The authors concluded that the problem of gene silencing in sugarcane might derive from the type of promoter used to drive the gene rather than number of integration sites (Hansom *et al*., 1999). In contrast with these findings, another group found that gene expression under the control of Maize *Ubi1* promoter was greatly reduced after regeneration of transformed sugarcane (Wang *et al*., 2002). An investigation of the mechanisms of gene silencing in sugarcane by both of these groups, however, found that silencing was due to post-transcriptional effects. This means that the introduced gene is still active, but the RNAs transcribed from the transgene are targeted for degradation by a currently unclear process. Therefore no protein product is produced from the transgene. This was an unexpected finding, as the apparent promoter dependence and often-observed developmental onset of silencing in sugarcane appears more consistent with transcriptional silencing.

Post-transcriptional silencing is a natural regulatory mechanism in plants that can specifically recognize foreign RNA and target it for degradation (Vance and Vaucheret, 2001). This process is also the underlying molecular mechanism in many cases of engineered virus resistance in plants (Baulcombe, 1999), e.g. in sugarcane (Ingelbrecht *et al*., 1999). Detailed reviews have been published about both transcriptional and post-transcriptional gene-silencing (Iyer *et al*.,

2000; Sijen and Kooter, 2000; Vaucheret and Fagard, 2001; Matzke *et al*., 2002). However, further investigation of the mechanisms of transgene silencing is required before predictions can be made about the silencing of specific promoters. To date, no specific promoter sequences or features have been identified that could explain why some promoters are silenced and others not. Until such time promoters for sugarcane transformation will have to be evaluated on a 'trial and error' basis.

2.4 THE ROLE OF TRANSCRIBED SEQUENCES IN THE REGULATION OF GENE EXPRESSION

Traditionally, research into the regulation of gene expression in plants has focused on the role of promoters (Guilfoyle, 1997). More recently, however, the importance of sequences located downstream of the transcription initiation site has been recognised. These sequences have been found to contribute to both the level and the location of expression of the gene with which they are associated. Research to date has mainly considered the role of introns, although regulation of expression by other transcribed sequences has also been demonstrated. Enhancement of gene expression by sequences located within the 5'-untranslated leader has been demonstrated for several plant genes, including the maize *Shrunken-1* gene (Clancy *et al*., 1994), the spinach *PetE*, *PsaF* and *PetH* genes for thylakoid proteins (Bolle *et al*., 1994), the rice actin (*Act1*) (Zhang *et al*., 1991) and sucrose phosphate synthase (*sps1*) (Martnez-Trujillo *et al*., 2003) genes. Regulation of tissue specificity by transcribed sequences has also been reported. The pea ferrodoxin (*Fed-1*) gene, for example, requires the 5'-leader sequence (Dickey *et al*., 1998) and exon sequences (Elliott *et al*., 1989) for light responsiveness. Another pea gene that requires both the 5'-leader sequence and the coding region for light regulation is the plastocyanin (*PetE*) gene (Helliwell *et al*., 1997). A light responsive element has also been found in the coding region of the tobacco *psaDb* gene (Yamamoto *et al*., 1997). Examples of tissue specificity mediated by sequences located in the 3'-untranslated region include the *Flaveria bidentis Me1* gene which contains an enhancer-like element in its 3'-untranslated region that confers high-level expression in leaves (Marshall *et al*., 1997), and nodule parenchyma-specific expression of the *Sesbania rostrata* early nodulin (*SrEnod2*) gene (Chen *et al*., 1998).

2.4.1 The role of introns

Introns have been found to play a role in the regulation of gene expression in a broad range of organisms, including nematodes (Okkema *et al*., 1993), insects (Schultz *et al*., 1991; Meredith and Storti, 1993), birds (Sorkin *et al*., 1993), fungi (Xu and Gong, 2003) and mammals (Luo *et al*., 1998; Chan *et al*., 1999; Chen *et al*., 2000). Introns can affect gene expression in different ways.

2.4.1.1 Intron-mediated enhancement

Stimulation of gene expression by introns in plants was first demonstrated by Callis and co-workers (1987) who showed that the maize *Adh1* first intron increased the expression of several genes, a phenomenon later termed intron-mediated enhancement (Mascarenhas *et al*., 1990). Subsequently, many introns that mediate enhanced gene expression in plants have been identified. Other introns known to enhance gene expression in monocotyledonous species include those from the maize *Bz1*, (Callis *et al*., 1987), *AdhI* (Mascarenhas *et al*., 1990), *ShI* (Vasil *et al*., 1989), *UbiI* (Christensen *et al*., 1992), *Hsp82* and *GapAI* (Donath *et al*., 1995) genes, and the rice *SalT* (Rethmeier *et al*., 1997), *Wx* (Li *et al*., 1995), *tpi* (Xu *et al*., 1994), and *Ostub 16* (Morello *et al*., 2002) genes. Similarly, introns contained in genes of dicotyledonous species that elevate expression include those from the petunia *rbcS* (Dean *et al*., 1989) and *PhADF1* (Mun *et al*., 2002), the potato *Sus3* (Fu *et al*., 1995a) and *Sus4* (Fu *et al*., 1995b), and the *Arabidopsis UBQ3, UBQ10* (Norris *et al*., 1993), *PAT1* (Rose and Beliakoff, 2000), and *At eEF-1β* (Gidekel *et al*., 1996) genes.

2.4.1.2 Intron-mediated tissue specificity

In addition to enhancement of expression, cases in which introns were required for tissue-specific expression of plant genes have also been reported. For example, an intron sequence is required for plastid and light-dependant expression of the *PsaD* gene of the spinach plant (Bolle *et al*., 1996). Expression of the *AGAMOUS (AG)* floral homeotic gene in *Arabidopsis* flowers requires an enhancer sequence located within an intron (Busch *et al*., 1999). Tissue preferential expression in actively dividing tissues of the rice *OsTubA1* gene is mediated by the first intron

(Jeon *et al*., 2000). An intron also contains the sequence responsible for endosperm-specific expression of the barley *SbeIIb* gene during seed development (Ahlandsberg *et al*., 2002).

Another example of introns affecting the pattern of plant gene expression is found in higher-plant sucrose synthase genes. With the exception of one gene in *Arabidopsis* (Martin *et al*., 1993), higher-plant sucrose synthase genes cloned to date all contain a very large intron conserved in position in the 5'-untranslated region (UTR), located between a non-coding first exon and a coding second exon. These include sucrose synthase genes from potato (Fu *et al*., 1995a), maize (Shaw *et al*., 1994; Vasil *et al*., 1989), *Arabidopsis* (Chopra *et al*., 1992), and citrus (Komatsu *et al*., 2002). Removal of this intron from two classes of sucrose synthase genes from potato, *Sus3* (Fu *et al*., 1995a) and *Sus4* (Fu *et al*., 1995a), results in changes in the pattern of expression, though alterations in tissue-specific expression observed on removal of the intron are dependant on the presence of promoter and 3'-UTR sequences. This intron therefore confers positive and negative tissue-specific regulated expression. The 5'-UTR intron of the maize sucrose synthase gene, *ShI*, is also extremely important for *ShI* (Vasil *et al*., 1989) expression in maize and can confer a dramatic enhancement of gene expression to heterologous genes (Maas *et al*., 1991; Clancy *et al*., 1994).

2.4.1.3 Conservation of introns within the 5'-untranslated region

The occurrence of an intron within the 5'-UTR, which separates a first non-coding exon from a second coding exon, was until recently, believed to be very rare (Hawkins, 1988; Vasil *et al*., 1989). The conservation of large introns in the 5'-UTR of plant genes is, however, not a phenomenon restricted to sucrose synthase genes. A large intron is present in the 5'-leader sequence of two soybean actin genes (Pearson and Meagher, 1990). This led to the suggestion that an intron in this position could be a common feature in plant actin genes based on the conservation of a potential intron acceptor site in the 5'-UTR of other soybean actin genes, as well as sequences from maize, *Arabidopsis*, rice and petunia. Introns in this position have subsequently been found in actin genes isolated from rice (McElroy *et al*., 1990), *Arabidopsis* (An *et al*., 1996; Huang *et al*., 1997) and tobacco (Thangavelu *et al*., 1993). Higher plant polyubiquitin genes (from *Arabidopsis* (Norris *et al*., 1993), maize (Christensen *et al*., 1992), sunflower (Binet *et al*., 1991), tomato (Hoffman *et al*., 1991), potato (Garbarino *et al*., 1995),

tobacco (Plesse *et al*., 2001), sugarcane (Wei *et al*., 1999) and rice (Wang *et al*., 2000)) also possess a conserved intron in their 5'-UTR. Many of these introns have been shown to contribute to the regulation of the level and pattern of the expression of the associated genes (Vasil *et al*., 1989; McElroy *et al*., 1990; Christensen *et al*., 1992; Norris *et al*., 1993; Fu *et al*., 1995a; Fu *et al*., 1995b; Wei *et al*., 1999). Other introns that occur in this position and enhance expression include the rice *SalT* (Rethmeier *et al*., 1997), *Wx* (Li *et al*., 1995) and *Ostub16* (Morello *et al*., 2002) first introns. The significance of the intron position in the 5'-UTR is not known, but a functional requirement for the presence of such an intron may be correlated with the conservation of the 5'-non-coding exon.

2.4.1.4 Intron-mediated enhancement in monocotyledonous vs dicotyledonous plants

Intron-mediated enhancement occurs in both monocotyledonous and dicotyledonous plants, but it is generally greater in monocotyledonous plants where intron-mediated enhancement of up to a 100-fold is not uncommon (e.g. Callis *et al*., 1987; Vasil *et al*., 1989; Maas *et al*., 1991). In dicotyledonous plants it commonly ranges from 2- to 10-fold (e.g. Dean *et al*., 1989; Tanaka *et al*., 1990; León *et al*., 1991; Norris *et al*., 1993; Rose and Beliakoff, 2000). Also, introns which significantly enhance expression in monocotyledonous species, have little or no effect when tested in dicotyledonous plants (Li *et al*., 1995; Maas *et al*., 1991; Tanaka *et al*., 1990). Although many promoters are active in both dicotyledonous and monocotyledonous plant species, modification of the promoters is often required to achieve high levels of expression in monocotyledonous species (Schünmann *et al*., 2003). The most widely used strategy is the addition of an intron, usually derived from a monocotyledonous plant gene, between the promoter and the transgene.

Intron-mediated enhancement has been used extensively for virus-derived promoters, such as the cauliflower mosaic virus (CaMV) 35S promoter. In its native form, the promoter is only poorly active in monocotyledonous plants (McElroy *et al*., 1991; Rathus *et al*., 1993; Vasil *et al*., 1989), but the addition of an intron taken from maize *Adh1* (Callis *et al*., 1987; Mascarenhas *et al*., 1990; Cornejo *et al*., 1993), *Sh1* (Vasil *et al*., 1989; Maas *et al*., 1991; Clancy *et al*., 1994), *Bz1* (Callis *et al*., 1987), *Ubi1* (Vain *et al*., 1996), rice *SalT* (Rethmeier *et al*., 1997; Rethmeier *et al*. 1998), *Wx* (Li *et al*., 1995), and *Act1* (Vain *et al*., 1996) dramatically improved expression in rice,

maize and various grasses. Interestingly, introns derived from dicotyledonous genes also enhanced expression of the CaMV 35S promoter in monocotyledonous, but not in dicotyledonous plants, e.g. the french bean phaseolin intron (Mitsuhara *et al*., 1996), the castorbean catalase *cat-1* intron (Tanaka *et al*., 1990), and the petunia chalcone synthase intron *chsA* (Vain et al., 1996).

A suite of promoters isolated from the subterranean clover stunt virus (SCSV) shown to be active in a number of dicotyledonous species (Schünmann *et al*., 2003a), were also modified for use in monocotyledonous plant transformation. While the original viral vectors exhibited low levels of activity in transgenic rice, insertion of the maize *Ubi1*, *Adh1* or rice *Act1* introns increased the level of expression by 10- to 50-fold (Schünmann *et al*., 2003b). Addition of the maize *Ubi1* intron significantly increased GUS expression directed by promoters derived from the banana bunchy top virus (BBTV) in transgenic banana plants (Dugdale *et al*., 2000). Addition of an intron from maize *Adh1*, rice *Act1* and sugarcane *rbcS* genes also significantly enhanced promoter activity of BBTV promoters in embryogenic banana cells (Dugdale *et al*., 2001). The sugarcane bacilliform badnavirus (ScBV) promoter coupled to the maize *Adh1* intron was able to drive near-constitutive expression in transgenic rice (Tzafrir *et al*., 1998), although the promoter was not evaluated in the absence of the intron. Addition of a monocotyledonous plant-derived intron, rice *Act1* intron, also enhanced expression of a dicotyledonous promoter, potato *pin2* promoter, in transgenic rice (Xu *et al*., 1993).

Quantitative differences in intron-mediated enhancement between monocotyledonous and dicotyledonous plants indicate differences in underlying mechanisms. Supporting this, differences have also been observed between monocotyledonous and dicotyledonous plants in the processing of heterologous introns. Introns of dicotyledonous plant origin, and even mammalian introns, appear to be accurately and efficiently processed when expressed in monocotyledonous plant cells (Tanaka *et al*., 1990; Goodall and Filipowicz, 1991). In contrast, several introns of monocotyledonous plant origin are inefficiently processed or not processed at all when introduced into dicotyledonous plant cells (Keith and Chua, 1986; Goodall and Filipowicz, 1991; Mitsuhara *et al*., 1996). Differences in intron composition between monocotyledonous and dicotyledonous plants have also been observed (Goodall and Filipowicz, 1991). These findings

may indicate that different mechanisms exist for intron-mediated enhancement in monocotyledonous and dicotyledonous plants.

2.3.1.5 Features and mechanisms of intron-mediated enhancement

The mechanisms of intron-mediated enhancement are not yet fully understood. However, some common features have emerged from cases in which these mechanisms have been explored:

Introns must be contained within transcribed sequences and in the proper orientation to elevate gene expression, unlike transcriptional enhancers which are usually position and orientation independent (Callis *et al*., 1987; Vasil *et al*., 1989; Mascarenhas et al., 1990; McElroy *et al*., 1990; Maas *et al*., 1991; Clancy *et al*., 1994; Li *et al*., 1995; Bourdon *et al*., 2001; Rose, 2002). The ability of introns to enhance expression declines as distance from the promoter increases (Rose 2004). The length (Sinibaldi and Mettler, 1992) and composition (Luehrsen and Walbot, 1991; Maas *et al*., 1991; Clancy *et al*., 1994; Carle-Urioste *et al*., 1994; Donath *et al*., 1995) of flanking sequences, as well as the coding region of the expressed gene (Rethmeier *et al*., 1997; Rethmeier *et al*., 1998; Sinibaldi and Mettler, 1992) influence the degree of stimulation. Intron-mediated enhancement is generally greater for weaker promoters (Callis *et al*., 1987; Mascarenhas *et al*., 1990; Luehrsen and Walbot, 1994; Bourdon *et al*., 2001). The same introns can evoke different levels of expression in the context of different promoters (Callis *et al*., 1987; Vasil *et al*., 1989), and different introns may evoke different levels of expression in the context of the same promoter (Vasil *et al*., 1989; Mascarenhas *et al*., 1990). Enhancement also depends on the tissue type and physiological conditions (Tanaka *et al*., 1990; Sinibaldi and Mettler, 1992; Gallie and Young, 1994; Fu *et al*., 1995a; Fu *et al*., 1995b; Plesse *et al*., 2001). Large, overlapping internal deletions can be made without affecting the ability of the intron to enhance expression (Clancy *et al*., 1994; Luehrsen and Walbot, 1994; Rose and Beliakoff, 2000; Clancy and Hannah, 2002), indicating that specific sequences required for enhancement must be present in multiple copies, making them redundant.

Very few specific intron sequences required for enhancement have been identified to date. A study of the maize *GapA1* gene showed that an octameric sequence motif contained within the first intron, which appeared to bind a maize nuclear factor, partially restored intron-dependent gene expression in the absence of the intron (Donath *et al*., 1995). The authors note that the same

motif is also present in the maize *Adh1* and *Sh1* first introns, both of which are known to enhance gene expression. A more recent study of the maize *Sh1* first intron revealed the presence of a redundant 35 bp T-rich motif which enhanced expression (Clancy and Hannah, 2002). It is not clear whether the above-mentioned octameric sequence forms a part of the 35 bp motif or not. The presence of redundant sequence motifs, however, is consistent with the finding that every part of an intron is individually dispensable for enhancement (Rose and Beliakoff, 2000).

Most observations to date suggest that intron-mediated enhancement occurs by cotranscriptional or posttranscriptional mechanisms. An increase in mRNA levels resulting from the presence of an intron has often been observed (Callis *et al*., 1987; Dean *et al*., 1989; Luehrsen and Walbot, 1991; Rethmeier *et al*., 1997; Rose and Last, 1997). This increase in steady state mRNA is not a result of increased transcription (Dean *et al*., 1989; Rose and Last, 1997). Studies have also shown that the half-life of the mRNA was the same with or without the intron (Nash and Walbot, 1992; Rethmeier *et al*., 1997). Extended mRNA persistence is therefore not a defining characteristic of intron-mediated enhancement. Also, increased mRNA levels do not always sufficiently account for increased enzyme activity (Mascarenhas *et al*., 1990; Tanaka *et al*., 1990; Bourdon *et al*., 2001; Rose, 2004). It was therefore suggested that pre-mRNA splicing must somehow improve the quality, as well as the quantity of the mRNA (Mascarenhas *et al*., 1990).

Although splicing seems to be required for intron-mediated enhancement, it alone is not enough, as introns that vary in their ability to enhance expression are all efficiently spliced (Rose, 2002). Reduced splicing efficiency, as a result of the deletion of 5'-exon sequences, mutation of splice junctions or intron deletions which block splicing, causes a decrease in enzyme activity (Luehrsen and Walbot, 1994; Clancy and Hannah, 2002). When splicing of the *Arabidopsis PAT1* intron was prevented, the intron retained some ability to increase mRNA accumulation (Rose, 2002). The simultaneous elimination of branch points and the 5'-splice site, structures involved in the first two steps of spliceosome assembly (Simpson and Filipowicz, 1996), completely abolished enhancement (Rose, 2002). These results suggest that although intron recognition by the splicing machinery is required, splicing *per se* is not enough to enhance expression.

Introns could stimulate expression in several different ways, or by a combination of mechanisms. Several mechanisms have been suggested. Association with the spliceosome may increase mRNA stability by influencing RNA events such as capping, polyadenylation, RNA turnover and transport to the cytoplasm (Simpson and Filipowicz, 1996; Snowden *et al*., 1996). Introns could possibly promote transcript elongation, thereby increasing the probability that full-length transcripts will be produced, leading to increased mRNA accumulation without affecting transcription initiation (Rose, 2002). As different introns could influence expression by different mechanisms, a complete understanding of intron-mediated enhancement will require a detailed analysis of different introns in the context of different promoters and genes, in different species. A better understanding of the role of these sequences can provide new insights into the complex processes that act together to regulate gene expression.

2.5 UDP-GLUCOSE DEHYDROGENASE

2.5.1 Function of UDP-glucose dehydrogenase

UDP-glucose dehydrogenase (EC1.1.1.22) catalyses the oxidation of UDP-glucose to UDP-glucuronic acid with the concomitant reduction of two molecules of NAD^+ (Nelsestuen and Kirkwood, 1971; Turner and Botha, 2002). UDP-glucuronic acid serves as substrate for glycosyltransferases and for nucleotide sugar interconversion enzymes which produce precursors for hemicellulose and pectin, including arabinans, arabinogalactans, glucuronoarabinoxylans, rhamnogalacturonans, xylans and xyloglucans (Carpita, 1996; Bolwell, 2000; Gibeaut, 2000). Both hemicellulose and pectin are key components of plant cell walls, providing a matrix that strengthens the cell wall. It has previously been shown that the enzyme structure of UDP-glucose dehydrogenase is highly conserved between plants and animals, even though the product of the reaction is utilised to produce entirely different polysaccharides in plants (Gibeaut, 2000) and animals (Hempel *et al*., 1994). This suggests strict structural requirements for the correct functioning of the enzyme.

2.5.2 Expression of UDP-glucose dehydrogenase

A general correlation between the expression of UDP-glucose dehydrogenase and a demand for structural polysaccharides in tissues that are actively synthesising cell walls, has been reported

for several species, including sycamore, poplar (Dalessandro and Northcote 1977a; Johansson *et al*., 2002), *Catharanthus roseus* (Amino *et al*., 1985), the liverwort *Reilla helicophylla* (Witt, 1992), French bean (Robertson *et al*., 1995a), soybean (Stewart and Copeland, 1998; Tenhaken and Thulke, 1996), and *Arabidopsis* (Seitz *et al*., 2000). These and other studies have also shown that UDP-glucose dehydrogenase is also often the least active enzyme involved in the nucleotide sugar interconversion pathway, suggesting that this enzyme is rate-limiting for the provision of precursors for the expanding cell wall (Amino *et al*., 1985; Dalessandro and Northcote, 1977c; Robertson *et al*., 1995a; Robertson *et al*., 1995b). In addition, this reaction may represent a control point for the irreversible flow of carbon into the pool of UDP sugars required for the synthesis of structural polysaccharides, as the activity of UDP-glucose dehydrogenase is controlled by feedback inhibition by UDP-xylose, the decarboxylation product of UDP-glucuronic acid (Dalessandro and Northcote1977a; Dalessandro and Northcote1977c; Stewart and Copeland, 1999; Hinterberg *et al*., 2002; Turner and Botha, 2002).

Early investigations of the expression of UDP-glucose dehydrogenase in gymnosperms (Dalessandro and Northcote, 1977b) and angiosperms (Dalessandro and Northcote, 1977a) found that the activity of UDP-glucose dehydrogenase varied during differentiation of cambium to xylem according to the type of polysaccharide synthesised. In the angiosperms, sycamore and poplar, the activity and concentration of UDP-glucose dehydrogenase increased threefold from cambial cells to differentiating and differentiated xylem cells (Dalessandro and Northcote, 1977a), correlating with an increased demand for UDP-glucuronic acid and UDP-xylose during secondary cell wall thickening. In the gymnosperms pine and fir, activity and concentration of UDP-glucose dehydrogenase was much lower than in the angiosperms. Also, in pine a decrease in the activity of UDP-glucose dehydrogenase was observed during differentiation (Dalessandro and Northcote, 1977b). According to the authors, this variation reflects a difference in the composition of the cell walls of angiosperms (more xylan polymers) and gymnosperms (very low amounts of xylan polymers), and the type and amount of polysaccharide formed is controlled by the adjustment of the relevant enzyme activities. This could mean that manipulation of the level of UDP-glucose dehydrogenase expression may permit the modification of cell wall material by changing the availability of monosaccharide precursors.

A study of elicitor-stressed French bean showed a significant induction of UDP-glucose dehydrogenase in response to elicitor treatment (Robertson *et al*., 1995b). The same group subsequently purified a 40 kDa UDP-glucose dehydrogenase with alcohol dehydrogenase activity, from French bean (Robertson *et al*., 1996). An antibody raised against this enzyme localised it to the vascular tissues of French bean hypocotyls, and protein blots confirmed induction of this enzyme by elicitor treatment. It has since been suggested by several authors that this enzyme is not a genuine UDP-glucose dehydrogenase (Tenhaken and Thulke, 1996; Turner and Botha, 2002). Alternatively, it has been suggested that there are two types of dehydrogenases (Bolwell, 2000). However, no further evidence to support the existence of a second type of dehydrogenase has been presented to date.

The highest level of UDP-glucose dehydrogenase gene expression in soybean seedlings was detected in root tips and lateral roots, with moderate expression in the epicotyl and expanding leaves, all actively growing tissues (Tenhaken and Thulke, 1996). Activity of the enzyme was also shown to be maximal during initial stages of nodule growth and development in soybean nodules (Stewart and Copeland, 1998). As observed for soybean, UDP-glucose dehydrogenase activity in *Arabidopsis* was especially high in roots. In addition, many but not all growing tissues showed high activity levels of the enzyme. Hypocotyledons and cotyledons of young seedlings, for instance, did not show significant UDP-glucose dehydrogenase activity. As many seeds contain phytic acid as a storage compound, it has been suggested that the inositol oxidation pathway (discussed in more detail in Chapter 3, p39) is predominantly active in seedlings to metabolise the inositol liberated from the phytic acid (Tenhaken and Thulke, 1996). In poplar, UDP-glucose dehydrogenase was expressed predominantly in differentiating xylem and young leaves with very low levels detected in phloem tissues (Dalessandro and Northcote1977a; Johansson *et al*., 2002). Surprisingly, lower levels of expression were found in leaf meristems than in mature leaves. This finding provides further evidence for the dominance of the inositol oxidation pathway in some actively growing tissues or organs (Johansson *et al*., 2002).

Significant levels of UDP-glucose dehydrogenase have also been detected in some mature tissues, such as the leaf axil meristem of young and old *Arabidopsis* plants, even when these

tissues are considered to be quiescent (Seitz *et al*., 2000). The role of UDP-glucose dehydrogenase in mature tissues is not yet understood.

2.6 REFERENCES

Aarts, M. G. M. and Fiers, M. W. E. J. 2003, What drives plant stress genes?, *Trends in Plant Science*, **8** (3): 99-102.

Ahlandsberg, S., Sun, C., and Jansson, C. 2002, An intronic element directs endosperm-specific expression of the *sbellb* gene during barley seed development, *Plant Cell Reports*, **20**: 864-868.

Amino, S., Takeuchi, Y., and Komamine, A. 1985, Changes in enzyme activities inved in formation and interconversion of UDP-sugars during the cell cycle in a synchronous culture of *Catharanthus roseus*, *Physiologia Plantarum*, **64**: 111-117.

An, G. and Kim, H. 1993, Techniques for isolating and characterizing plant transcription promoters, enhancers, and terminators, in *Methods in plant molecular biology and biotechnology.* Glick, B.R. and Thompson, J.E. (Eds.), ISBN 0-8493-5164-2, pp. 155-166. C.R.C. Press, London.

An, Y. Q., Huang, S., McDowell, J. M., McKinney, E. C., and Meagher, R. B. 1996, Conserved expression of the Arabidopsis ACT1 and ACT3 actin subclass in organ primordia and mature pollen, *Plant Cell*, **8**: 15-30.

Arencibia, A., De la Riva, G., and Selman-Housein, G. 1995, Production of transgenic sugarcane (*Saccharum officinarum* L.) plants by intact cell electroporation, *Plant Cell Reports*, **14** (305) 309.

Arencibia, A., Molina, P., Gutierrez, C., Fuentes, A., Greenidge, V., Menendez, E., De la Riva, G., and Selman-Housein, G. 1992, Regeneration of transgenic sugarcane (*Saccharum officinarum* L.) plants from intact meristematic tissue transformed by electroporation, *Biotecnología Aplicada*, **9**: 156-165.

Arencibia, A., Vazquez, R. I., Prieto, D., Tellez, P., Carmona, E. R., Coego, A., Hernandez, L., De la Riva, G., and Selman-Housein, G. 1997, Transgenic sugarcane plants resistant to stem borer attack, *Molecular Breeding*, **3**: 247-255.

Arencibia, A. D., Carmona, E. R., Cornide, M. T., Castiglione, S., O'Relly, J., Chinea, A., Oramas, P., and Sala, F. 1999, Somaclonal variation in insect-resistant transgenic sugarcane (Saccharum hybrid) plants produced by cell electroporation, *Transgenic Research*, **8** (5): 349-360.

Arencibia, A. D., Carmona, E. R., Tellez, P., Chan, M. T., Yu, S. M., Trujillo, L. E., and Oramas, P. 1998, An efficient protocol for sugarcane (Saccharum s L.) transformation mediated by Agrobacterium tumefaciens, *Transgenic Research*, **7**: 1-10.

Atienzar, F., Evenden, A., Jha, A., Savva, D., and Depledge, M. 2000, Improved analysis of promoter activity in biolistically transformed plant cells, *BioTechniques*, **28** (1): 54-57.

Basu, C., Kausch, A. P., Luo, H., and Chandlee, J. M. 2003, Promoter analysis in transient assays using a GUS reporter gene construct in creeping bentgrass (Agrostis palustris) , *Journal of Plant Physiology*, **160** (10): 1233-1239.

Baulcombe, D. C. 1999, Viruses and gene silencing in plants, *Archives of virology supplementum*, **15**: 189-201.

Benfey, P. N., Ren, L., and Chua, N.-H. 1990, Combinatorial and synergistic properties of CaMV 35S enhancer subdomains, *EMBO Journal*, **9** (6): 1685-1696.

Binet, M. N., Weil, J. H., and Tessier, L. H. 1991, Structure and expression of sunflower ubiquitin genes, *Plant Molecular Biology*, **17** (3): 395-407.

Binns, A. N. and Tomashow, M. F. 1988, Cell biology of *Agrobacterium* infection and transformation of plants, *Annual Review of Microbiology*, **42**: 575-606.

Birch, R. G. 1997, Plant transformation: problems and strategies for practical application, *Annual Review of Plant Physiology and Plant Molecular Biology*, **48**: 297-326.

Bolle, C., Herrmann, R. G., and Oelmüller, R. 1996, Intron sequences are inved in the plastid- and light-dependent expression of the spinach *PsaD* gene, *Plant Journal*, **10**: 919-924.

Bolle, C., Sopory, S., Lübberstedt, T., Herrmann, R. G., and Oelmüller, R. 1994, Segments encoding 5'-untranslated leaders of genes for thylakoid proteins contain *cis*-elements essential for transcription, *Plant Journal*, **6** (4): 513-523.

Bolwell, G. P. 2000, Biosynthesis of plant cell wall polysaccharides, *Trends in Glycoscience and Glycotechnology*, **12** (65): 143-160.

Bourdon, V., Harvey, A., and Lonsdale, D. M. 2001, Introns and their positions affect the translational activity of mRNA in plant cells, *EMBO Reports* **2** (5): 394-398.

Bower, R. and Birch, R. G. 1992, Transgenic sugarcane plants via microprojectile bombardment, *Plant Journal*, **2** (3): 409-416.

Busch, M. A., Bomblies, K., and Weigel, D. 1999, Activation of a floral homeotic gene in *Arabidopsis*, *Science*, **285**: 585-587.

Callis, J., Fromm, M., and Walbot, V. 1987, Introns increase gene expression in cultured maize cells, *Genes and Development*, **1**: 1183-1200.

Carle-Urioste, J. C., Ko, C. H., Benito, M.-I., and Walbot, V. 1994, *In vivo* analysis of intron processing using splicing-dependent reporter gene assays, *Plant Molecular Biology* **26:** 1785-1795.

Carpita, N. C. 1996, Structure and biogenesis of the cell walls of grasses, *Annual Review of Plant Physiology and Plant Molecular Biology,* **47**: 445-476.

Casas, A. M., Kononowicz, A. K., Bressan, R. A., and Hasegawa, P. M. 1995, Cereal transformation through particle bombardment, *Plant Breeding Review*, **13**: 235-264.

Chan, M.-T., Lee, T.-M., and Chang, H.-H. 1993, *Agrobacterium*-mediated production of transgenic rice plants expressing a chimeric α-amylase promoter/β-glucuronidase gene, *Plant Molecular Biology*, **22**: 491-506.

Chan, R. Y. Y., Boudreau-Lariviere, C., Angus, L. M., Mankai, A. M., and Jasmin, B. J. 1999, An intronic enhancer conaining an N-box motif is required for synapse- and tissue-specific expression of the acetylcholinesterase gene in skeletal muscle fibers, *Proceedings of the National Academy of Sciences of USA*, **96**: 4627-4632.

Chen, J., Hayes, P., Roy, K., and Sirotnak, F. M. 2000, Two promoters regulate transcription of the mouse folypolyglutamate synthase gene: three tightly clustered Sp1 sites within the first intron markedly enhance activity of promoter B, *Gene*, **242** (1-2): 257-264.

Chen, R., Silver, D. L., and De Bruijn, F. J. 1998, Nodule parenchyma-specific expression of the sesbania rostrata early nodulin gene SrEnod2 is mediated by its 3' untranslated region, *Plant Cell*, **10** (10): 1585-1602.

Chopra, S., Del-favero, J., Dolferus, R., and Jacobs, M. 1992, Sucrose synthase of *Arabidopsis*: Genomic cloning and sequence characterization, *Plant Molecular Biology*, **18**: 131-134.

Christensen, A. H., Sharrock, R. A., and Quail, P. H. 1992, Maize polyubiquitin genes: structure, thermal perturbation of expression and transcript splicing, and promoter activity following transfer to protoplasts by electroporation, *Plant Molecular Biology*, **18** (4): 675-689.

Clancy, M. and Hannah, L. C. 2002, Splicing of the maize Sh1 first intron is essential for enhancement of gene expression, and a T-rich motif increases expression without affecting splicing, *Plant Physiology*, **130** (2): 918-929.

Clancy, M., Vasil, V., Hannah, L. C., and Vasil, I. K. 1994, Maize *Shrunken-1* intron and exon regions increase gene expression in maize protoplasts, *Plant Science*, **98**: 151-161.

Cornejo, M.-J., Luth, D., Blankenship, K. M., Anderson, O. D., and Blechl, A. E. 1993, Activity of a maize ubiquitin promoter in transgenic rice, *Plant Molecular Biology*, **23**: 567-581.

Dalessandro, G. and Northcote, D. H. 1977a, Changes in enzymatic activities of nucleoside diphosphate sugar interconversions during differentiation of cambium to xylem in sycamore and poplar, *Biochemical Journal*, **162**: 267-279.

Dalessandro, G. and Northcote, D. H. 1977b, Changes in enzymatic activities of nucleoside diphosphate sugar interconversions durring differentiation of cambium to xylem in pine and fir, *Biochemical Journal*, **162**: 281-288.

Dalessandro, G. and Northcote, D. H. 1977c, Possible control sites of polysaccharide synthesis during cell growth and wall expansion of pea seedlings (*Pisum sativum* L.), *Planta*, **134**: 39-44.

De la Riva, G., Gonzalez-Cabrera, J., Vazquez-Padron, R. I., and Ayra-Pardo, C. 1998, *Agrobacterium tumefaciens*: a natural tool for plant transformation, *Electronic Journal of Biotechnology*, **1** (3): 118-133.

Dean, C., Favreau, M., Bond-Nutter, D., Bedbrook, J., and Dunsmuir, P. 1989, Sequences downstream of translation start regulate quantitative expression of two *Petunia rbcS* genes, *Plant Cell*, **1**: 201-208.

DeCleene, M. 1985, The susceptibility of monocotyledons to *Agrobacterium tumefaciens*, *Journal of Plant Physiology*, **113**: 81-89.

Dickey, L. F., Petracek, M. E., Nguyen, T. T., Hansen, E. R., and Thompson, W. F. 1998, Light regulation of Fed-1 mRNA requires an element in the 5' untranslated region and correlates with differential polyribosome association, *Plant Cell*, **10** (3): 475-484.

Donath, M., Mendel, R., Cerff, R., and Martin, W. 1995, Intron-dependent transient expression of the maize *GapA1* gene, *Plant Molecular Biology*, **28**: 667-676.

Dugdale, B., Becker, D. K., Beetham, P. R., Harding, R. M., and Dale, J. L. 2000, Promoters derived from banana bunchy top virus DNA-1 to-5 direct vascular-associated expression in transgenic banana (Musa spp.), *Plant Cell Reports*, **19** (8): 810-814.

Dugdale, B., Becker, D. K., Harding, R. M., and Dale, J. L. 2001, Intron-mediated enhancement of the banana bunchy top virus DNA- 6 promoter in banana (Musa spp.) embryogenic cells and plants, *Plant Cell Reports*, **20** (3): 220-226.

Elliott, A. R., Campbell, J. A., Brettell, R. I. S., and Grof, C. P. L. 1998, *Agrobacterium*-mediated transformation of sugarcane using GFP as a screenable marker, *Australian Journal of Plant Physiology*, **25** (6): 739-743.

Elliott, R. C., Dickey, L. F., White, M. J., and Thompson, W. F. 1989, cis-Acting Elements for Light Regulation of Pea Ferredoxin I Gene Expression Are Located within Transcribed Sequences, *Plant Cell*, **1** (7): 691-698.

Enríques-Obregón, G. A., Vazquez-Padron, R. I., Prieto-Samsonov, D. L., Perez, M., and Selman-Housein, G. 1997, Genetic transformation of sugarcane by Agrobacterium tumefaciens using antioxidant compounds, *Biotecnologia Aplicada*, **14** (3): 69-174.

Enríques-Obregón, G. A., Vásques-Padrón, R. I., Prieto-Samsonov, D. L., De la Riva, G., and Selman-Housein, G. 1998, Herbicide-resistant sugarcane (*Saccharum officinarum* L.) plants by *Agrobacterium*-mediated transformation, *Planta*, **206**: 20-27.

Falco, M. C., Neto, A. T., and Ulian, E. C. 2000, Transformation and expression of a gene for herbicide resistance in a Brazillian sugarcane, *Plant Cell Reports*, **19** (12): 1188-1194.

Fu, H., Kim, S. Y., and Park, W. D. 1995a, A potato *Sus3* sucrose synthase gene contains a context-dependent 3' leader element and a leader intron with both positive and negative tissue-specific effects, *Plant Cell*, **7**: 1395-1403.

Fu, H., Kim, S. Y., and Park, W. D. 1995b, High-level tuber expression and sucrose inducibility of a potato *Sus4* sucrose synthase gene require 5' and 3' flanking sequences and the leader intron, *Plant Cell*, **7**: 1387-1394.

Gallie, D. R. and Young, T. E. 1994, The regulation of gene expression in transformed maize aleurone and endosperm protoplasts. Analysis of promoter activity, intron enhancement, and mRNA untranslated regions on expression, *Plant Physiology*, **106** (3): 929-939.

Gallo-Meagher, M. and Irvine, J. E. 1996, Herbicide resistant transgenic sugarcane plants containing the bar gene, *Crop Science*, **36**: 1367-1374.

Garbarino, J. E., Oosumi, T., and Belknap, W. R. 1995, Isolation of a polyubiquitin promoter and its expression in transgenic potato plants, *Plant Physiology*, **109** (4): 1371-1378.

Gelvin, S. B. 2000, *Agrobacterium* and plant genes involved in T-DNA transfer and integration, *Annual Review of Plant Physiology and Plant Molecular Biology*, **51**: 223-256.

Gelvin, S. B. 2003, *Agrobacterium*-mediated plant transformation: the biology behind the gene-jockeying tool, *Microbiology and Molecular Biology Reviews*, **67** (1): 16-37.

Gibeaut, D. M. 2000, Nucleotide sugars and glycosyltransferases for synthesis of cell wall matrix polysaccharides, *Plant Physiology and Biochemistry*, **38** (1/2): 69-80.

Gidekel, M., Jimenez, B., and Herrera-Estrella, L. 1996, The first intron of the *Arabidopsis thaliana* gene coding for elongation factor 1β contains an enhancer-like element, *Gene*, **170**: 201-206.

Goodall, G. J. and Filipowicz, W. 1991, Different effects of intron nucleotide composition and secondary structure on pre-mRNA splicing in monocot and dicot plants, *EMBO Journal*, **10** (9): 2635-2644.

Gould, J., Devey, M., Hasegawa, O., Ulian, E. C., Peterson, G., and Smith, R. H. 1991, Transformation of *Zea mays* L. using *Agrobacterium tumefaciens* and the shoot apex, *Plant Physiology*, **95** (2): 426-434.

Guilfoyle, T. J. 1997, The structure of plant gene promoters, in *Genetic engineering: principles and methods*, 19 edn, . 19 J. K. Setlow, ed., Plenum Press, New York, pp 15-47.

Hansom, S., Bower, R., Zhang, L., Potier, B., Elliot, A., Basnayake, S., Cordeiro, G., Hogarth, D. M., Cox, M., Berding, N., and Birch, R. G. 1999, Regulation of transgene expression in

sugarcane, *Proceedings of the International Society of Sugarcane Technologists XXIII Congress*, *New Delhi. STAI, New Dehli (Ed. V. Singh)* pp 278-289.

Hawkins, J. D. 1988, A survey of intron and exon lengths, *Nucleic Acids Research*, **16**: 9893-9908.

Hehl, R. and Wingender, E. 2001, Database-assisted promoter analysis, *Trends in Plant Science*, **6** (6): 251-255.

Helliwell, C. A., Webster, C. I., and Gray, J. C. 1997, Light-regulated expression of the pea plastocyanin gene is mediated by elements within the transcribed region of the gene, *Plant Journal*, **12**: (3): 499-506.

Hempel, J., Perozich, J., Romovacek, H., Hinich, A., Kuo, I., and Feingold, D. S. 1994, UDP-glucose dehydrogenase from bovine liver - primary structure and relationship to other dehydrogenases, *Protein Science*, **3**: 1074-1080.

Hiei, Y., Ohta, S., Komari, T., and Kumashiro, T. 1994, Efficient transformation of rice (*Oryza sativa* L.) mediated by *Agrobacterium* and sequence analysis of the boundaries of the T-DNA, *The Plant Journal*, **6** (2): 271-282.

Hinterberg, B., Klos, C., and Tenhaken, R. 2002, Recombinant UDP-glucose dehydrogenase from soybean, *Plant Physiology and Biochemistry*, **40** (12): 1011-1017.

Hoffman, N. E., Ko, K., Milkowski, D., and Pichersky, E. 1991, Isolation and characterization of tomato cDNA and genomic clones encoding the ubiquitin gene ubi3, *Plant Molecular Biology*, **17** (6): 1189-1201.

Hooykaas, P. J. J. and Shilperoort, R. A. 1992, *Agrobacterium* and plant genetic engineering, *Plant Molecular Biology*, **19**: 15-38.

Huang, S., An, Y. Q., McDowell, J. M., McKinney, E. C., and Meagher, R. B. 1997, The Arabidopsis ACTII actin gene is strogly expressed in tissues of the emerging inflorescence, pollen, and developing ovules, *Plant Molecular Biology*, **33** (1): 125-139.

Ingelbrecht, I. L., Irvine, J. E., and Mirkov, T. E. 1999, Posttranscriptional gene silencing in transgenic sugarcane. Dissection of homology-dependant virus resistance in a monocot that has a complex polyploid genome, *Plant Physiology*, **119**: 1187-1197.

Ishida, Y., Saito, H., Ohta, S., Hiei, Y., Komari, T., and Kumashiro, T. 1996, High efficiency transformation of maize (Zea mays L.) mediated by Agrobacterium tumefaciens, *Nature Biotechnology*, **14** (6): 745-750.

Iyer, L. M., Kumpatla, S. P., Chandrasekharan, M. B., and Hall, T. C. 2000, Transgene silencing in monocots, *Plant Molecular Biology*, **43**: 323-346.

Jeon, J.-S., Lee, S., Jung, K.-H., Jun, S.-H., Kim, C., and An, G. 2000, Tissue-preferential expression of a rice a-tubulin gene, *OsTubA1*, mediated by the first intron, *Plant Physiology*, **123**: 1005-1014.

Johansson, H., Sterky, F., Amini, B., Lundeberg, J., and Kleczkowski, L. A. 2002, Molecular cloning and characterization of a cDNA encoding poplar UDP-glucose dehydrogenase, a key gene of hemicellulose/pectin formation, *Biochimica et Biophysica Acta-Gene Structure and Expression*, **1576** (1-2): 53-58.

Keith, B. and Chua, N.-H. 1986, Monocot and dicot pre-mRNAs are processed with different efficiencies in transgenic tobacco, *EMBO Journal*, **5**: 2419-2425.

Komari, T., Hiei, Y., Ishida, Y., Kumashiro, T., and Kubo, T. 1998, Advances in cereal gene transfer, *Current Opinion in Plant Biology*, **1** (2): 161-165.

Komatsu, A., Moriguchi, T., Koyama, K., Omura, M., and Akihama, T. 2002, Analysis of sucrose synthase genes in citrus suggests different roles and phylogenetic relationships, *Journal of Experimental Botany*, **53** (366): 61-71.

Laporte, M. M., Galagan, J. A., Prasch, A. L., Vanderveer, P. J., Hanson, D. T., Shewmaker, C. K., and Sharkey, T. D. 2001, Promoter strength and tissue specificity effects on growth of tomato plants transformed with maize sucrose-phosphate synthase, *Planta*, **212**: 817-822.

Last, D. I., Brettell, R. I. S., Chamberlain, D. A., Chaudhury, A., Larkin, P. J., Marsh, E. L., Peacock, W. J., and Dennis, E. S. 1991, pEmu: an improved vector for gene expression in cereal cells, *Theoretical and Applied Genetics*, **81**: 581-588.

Lefebvre, D. D. and Gellatly, K. S. 1997, Fundamentals of gene structure and control, in *Plant Metabolism*, 2 edn, D. T. Dennis et al., eds., pp. 3-16.

León, P., Planckard, F., and Walbot, V. 1991, Transient gene expression in protoplasts of *Phaseolus vulgaris* isolated from a cell suspension culture, *Plant Physiology*, **95**: 968-972.

Li, Y., Ma, H., Zhang, J., Wang, Z., and Hong, M. 1995, Effects of the first intron of rice Waxy gene on the expression of the foreign genes in rice and tobacco protoplasts, *Plant Science*, **108**: 181-190.

Liu, D., Oard, S. V., and Oard, J. H. 2003, High transgene expression levels in sugarcane (*Saccharum officinarum* L.) driven by the rice ubiquitin promoter RUBQ2, *Plant Science*, **165**: 743-750.

Lorence, A. and Verpoort, R. 2004, Gene transfer and expression in plants, *Methods in Molecular Biology*, **267**: 329-350.

Luehrsen, K. R. and Walbot, V. 1991, Intron enhancement of gene expression and the splicing efficiency of introns in maize cells, *Molecular and General Genetics*, **225** (1): 81-93.

Luehrsen, K. R. and Walbot, V. 1994, Addition of A- and U-rich sequence increases the splicing efficiency of a deleted form of a maize intron, *Plant Molecular Biology* **24** (3): 449-463.

Luo, Z., Camp, S., Mutero, A., and Taylor, P. 1998, Splicing of 5' introns dictates alternative splice selection of acetylcholinesterase pre-mRNA and specific expression during myogenesis, *Journal of Biological Chemistry*, **273** (43): 28486-28495.

Maas, C., Laufs, J., Grant, S., Korfhage, C., and Werr, W. 1991, The combination of a novel stimulatory element in the first exon of the maize Shrunken-1 gene with the following intron 1 enhances reporter gene expression up to 1000-fold, *Plant Molecular Biology*, **16** (2): 199-207.

Maenpaa, P., Gonzalez, E. B., Ahlandsberg, S., and Jansson, C. 1999, Transformation of nuclear and plastomic plant genomes by biolistic particle bombardment, *Molecular Biotechnology*, **13** (1) : 67-72.

Makarevitch, I., Svitashev, S. K., and Somers, D. A. 2003, Complete sequence analysis of transgene loci from plants transformed via microprojectile bombardment, *Plant Molecular Biology*, **52** (2): 421-432.

Manickavasagam, M., Ganapathi, A., Anbazhagan, V. R., Sudhakar, B., Selvaraj, N., Vasudevan, A., and Kasthurirengan, S. 2004, Agrobacterium-mediated genetic transformation and development of herbicide-resistant sugarcane (Saccharum species hybrids) using axillary buds, *Plant Cell Reports*, **23** (3): 134-143.

Marshall, J. S., Stubbs, J. D., Chitty, J. A., Surin, B., and Taylor, W. C. 1997, Expression of the C4 Me1 gene from Flaveria bidentis requires an interaction between 5[prime] and 3[prime] sequences., *Plant Cell*, . **9** (9): 1515-1525.

Martin, T., Frommer, W. B., Salanoubat, M., and Willmitzer, L. 1993, Expression of an *Arabidopsis* sucrose synthase gene indicates a role in metabolization of sucrose both during phloem loading and in sink organs, *The Plant Journal*, **4** (2): 367-377.

Martinez-Trujillo, M., Limones-Briones, V., Chavez-Barcenas, T., and Herrera-Estrella, L. 2003, Functional analysis of the 5' untranslated region of the sucrose phosphate synthase rice gene (*sps1*), *Plant Science*, **165**: 9-20.

Mascarenhas, D., Mettler, I. J., Pierce, D. A., and Lowe, H. W. 1990, Intron-mediated enhancement of heterologous gene expression in maize, *Plant Molecular Biology*, **15**: 913-920.

Matzke, M. A. and Matzke, A. J. M. 1995, How and why do plants inactivate homologous (trans)genes?, *Plant Physiology*, **107**: 679-685.

Matzke, M. A., Aufsatz, W., Kanno, T., Mette, M. F., & Matzke, A. J. M. 2002, Homology-dependent gene silencing and host defense in plants, *Advances in Genetics*, **46**: 235-275.

McElroy, D., Blowers, A. D., Jenes, B., and Wu, R. 1991, Construction of expression vectors based on the rice actin 1 (Act1) 5' region for use in monocot transformation, *Molecular and General Genetics*, **231** (1): 150-160.

McElroy, D., Zhang, W., Cao, J., and Wu, R. 1990, Isolation of an efficient actin promoter for use in rice transformation, *Plant Cell*, **2**: 163-171.

Mendoza, E. M. T. 2002, Molecular strategies to increase and stabilize the sucrose content of sugarcane (Saccharum officinarum L.), *Philippine Agricultural Scientist*, **85** (3): 307-318.

Meredith, J. and Storti, R. V. 1993, Developmental regulation of the *Drosophila* tropomyosin II gene in different muscles is controlled by muscle-type-specific intron enhancer elements and distal and proximal promoter control elements, *Developmental Biology*, **159** (2): 500-512.

Mitsuhara, I., Ugaki, M., Hirochika, H., Oshima, M., Murakami, T., Gotoh, Y., Katayouse, Y., Nakamura, S., Honkura, R., Nishimiya, S., Ueno, K., Mochizuki, A., Tanimoto, H., Tsugawa, H., Otsuki, Y., and Ohashi, Y. 1996, Efficient promoter cassettes for enhanced expression of foreign genes in dicotyledonous and monocotyledonous plants, *Plant Cell Physiology,* **37** (1): 49-59.

Morello, L., Bardini, M., Sala, F., and Breviario, D. 2002, A long leader intron of the *Ostub 16* rice B-tubulin gene is required for high-level gene expression and can autonomously promote transcription both *in vivo* and *in vitro*, *Plant Journal*, **29** (1): 33-44.

Mun, J. H., Lee, S. Y., Yu, H. J., Jeong, Y. M., Shin, M. Y., Kim, H., Lee, I., and Kim, S. G. 2002, Petunia actin-depolymerizing factor is mainly accumulated in vascular tissue and its gene expression is enhanced by the first intron, *Gene*, **292** (1-2): 233-243.

Nash, J. and Walbot, V. 1992, *Bronze-2* Gene Expression and Intron Splicing Patterns in Cells and Tissues of *Zea mays* L., *Plant Physiology*, **100**: 464-471.

Nelsestuen, G. L. and Kirkwood, S. 1971, The mechanism of action of uridine diphosphoglucose dehydrogenase, *Journal of Biological Chemistry*, **246** (12): 3828-3834.

Newell, C. A. 2000, Plant transformation technology. Developments and applications, *Molecular Biotechnology*, **16** (1): 53-65.

Norris, S. R., Meyer, S. E., and Callis, J. 1993, The intron of *Arabidopsis thaliana* polyubiquitin genes is conserved in location and is a quantitative determinant of chimeric gene expression, *Plant Molecular Biology*, **21**: 895-906.

Okkema, P. G., Harrison, S. W., Plunger, V., Aryana, A., and Fire, A. 1993, Sequence requirements for myosin gene expression and regulation in *Caenorhabditis elegans*, *Genetics*, **135** (2): 385-404.

Ono, S., Tanaka, T., Watakabe, Y., and Hiratsuka, K. 2004, Transient Assay System for the Analysis of *PR-1a* Gene Promoter in Tobacco BY-2 Cells, *Bioscience, Biotechnology, and Biochemistry*, **68** (4): 803-807.

Pawlowski, W. P. and Somers, D. A. 1996, Transgene inheritance in plants genetically engineered by microprojectile bombardment, *Molecular Biotechnology*, **6** (1): 17-30.

Pearson, L. and Meagher, R. B. 1990, Diverse soybean actin transcripts contain a large intron in the 5' untranslated leader: structural similarity to vertebrate muscle actin genes, *Plant Molecular Biology* , **14** (4): 513-526.

Plesse, B., Criqui, M. C., Parmentier, Y., Fleck, J., and Genschik, P. 2001, Effects of the polyubiquitin gene Ubi. U4 leader intron and first ubiquitin monomer on reporter gene expression in Nicotiana tabacum, *Plant Molecular Biology*, **45** (6): 655-667.

Potrykus, I. 1990, Gene transfer into cereals: an assessment, *BioTechnology*, **8**: 535-542.

Rakoczy-Trojanowska, M. 2002, Alternative methods of plant transformation - a short review, *Cellular and Molecular Biology Letters*, **7** (3): 849-858.

Rathus, C. and Birch, R. G. 1992, Stable transformation of callus from electroporated sugarcane protoplasts, *Plant Science*, **82** (81): 89.

Rathus, C., Bower, R., and Birch, R. G. 1993, Effects of promoter, intron and enhancer elements on transient gene expression in sugar-cane and carrot protoplasts, *Plant Molecular Biology*, **23** (3): 613-618.

Rethmeier, N., Kramer, E., Van Montagu, M., and Cornelissen, M. 1998, Identification of cat sequences required for intron-dependent gene expression in maize cells, *Plant Journal*, **13** (6): 831-835.

Rethmeier, N., Seurinck, J., Van Montagu, M., and Cornelissen, M. 1997, Intron-mediated enhancement of transgene expression in maize is a nuclear, gene-dependent process, *Plant Journal*, **12** (4): 895-899.

Robertson, D., Beech, I., and Bolwell, G. P. 1995a, Regulation of the enzymes of UDP-sugar metabolism during differentiation of french bean, *Phytochemistry*, **39** (1): 21-28.

Robertson, D., McCormack, B. A., and Bolwell, G. P. 1995b, Cell wall polysaccharide biosynthesis and related metabolism in elicitor-stressed cells of French bean (*Phaseolus vulgaris* L.), *Biochemical Journal*, **306** (3): 745-750.

Robertson, D., Smith, C., and Bolwell, G. P. 1996, Inducible UDP-glucose dehydrogenase from French bean (*Phaseolus vulgaris* L.) locates to vascular tissue and has alcohol dehydrogenase activity, *Biochemical Journal*, **313** (1): 311-317.

Rombauts, S., Florquin, K., Lescot, M., Marchal, K., Rouze, P., and Van de Peer, Y. 2003, Computational Approaches to Identify Promoters and cis-Regulatory Elements in Plant Genomes, *Plant Physiology*, **132** (3): 1162-1176.

Rose, A. B. 2002, Requirements for intron-mediated enhancement of gene expression in Arabidopsis, *Rna*, **8** (11): 1444-1453.

Rose, A. B. 2004, The effect of intron location on intron-mediated enhancement of gene expression in *Arabidopsis*, *Plant Journal*, **40** (5): 744-751.

Rose, A. B. and Beliakoff, J. A. 2000, Intron-mediated enhancement of gene expression independent of unique intron sequences and splicing, *Plant Physiology*, **122** (2): 535-542.

Rose, A. B. and Last, R. L. 1997, Introns act post-transcriptionally to increase expression of the *Arabidopsis thaliana* tryptophan pathway gene *PAT1*, *The Plant Journal*, **11** (3): 455-464.

Santosa, D. A., Hendroko, R., Farouk, A., and Greiner, R. 2004, A rapid and highly efficient method for transformation of sugarcane callus, *Molecular Biotechnology*, 28 (2): 113-119.

Schenk, P. M., Remans, T., Sagi, L., Elliott, A. R., Dietzgen, R. G., Swennen, R., Ebert, P. R., Grof, C. P., and Manners, J. M. 2001, Promoters for pregenomic RNA of banana streak badnavirus are active for transgene expression in monocot and dicot plants, *Plant Molecular Biology*, **47** (3): 399-412.

Schultz, J. R., Tansey, T., Gremke, L., and Storti, R. V. 1991, A muscle-specific intron enhancer required for rescue of indirect flight muscle and jump muscle function regulates *Drosophila* tropomyosin gene expression, *Molecular and Cellular Biology*, **11** (4): 1901-1911.

Schünmann, P. H. D., Llewewllyn, D. J., Surin, B., Boevink, P., De Feyter, R. C., and Waterhouse, P. M. 2003a, A suite of novel promoters and terminators for plant biotechnology, *Functional Plant Biology*, **30**: 443-452.

Schünmann, P. H. D., Surin, B., and Waterhouse, P. M. 2003b, A suite of novel promoters and termionators for plant biotechnology II. The pPLEX series for use in monocots, *Functional Plant Biology*, **30**: 453-460.

Seitz, B., Klos, C., Wurm, M., and Tenhaken, R. 2000, Matrix polysaccharide precursors in Arabidopsis cell walls are synthesized by alternate pathways with organ-specific expression patterns, *Plant Journal*, **21** (6): 537-546.

Setamou, M., Bernal, J. S., Legaspi, J. C., and Mirkov, T. E. 2002, Effects of snowdrop lectin (Galanthus nivalis agglutinin) expressed in transgenic sugarcane on fitness of Cotesia flavipes (Hymenoptera : Braconidae), a parasitoid of the nontarget pest Diatraea saccharalis (Lepidoptera: Crambidae), *Annals of the Entomological Society of America*, **95** (1): 75-83.

Shaw, J. R., Ferl, R. J., Baier, J., StClair, D., Carson, C., McCarty, D. R., and Hannah, L. C. 1994, Structural features of the maize *sus1* gene and protein, *Plant Physiology*, **106**: 1659-1665.

Sijen, T. and Kooter, J. M. 2000, Post-transcriptional gene-silencing: RNAs on the attack or on the defense?, *BioEssays*, **22** (6): 520-531.

Simpson, G. G. and Filipowicz, W. 1996, Splicing of precursors to mRNA in higher plants: mechanism, regulation and sub-nuclear organisation of spliceosomal machinery, *Plant Molecular Biology*, **32**: 1-41.

Sinibaldi, R. M. and Mettler, I. J. 1992, Intron splicing and intron-mediated enhanced expression in monocots, *Progress In Nucleic Acid Research And Molecular Biology*, **42**: 229-257.

Snowden, K. C., Buchholz, W. G., and Hall, T. C. 1996, Intron position affects expression from the *tpi* promoter in rice, *Plant Molecular Biology*, **31** (3): 689-692.

Snyman, S. J., Leibbrandt, N. B., and Botha, F. C. 1998, Buster® resistant sugarcane, *Proc S Afr Sug Technol Ass*, **72**: 138-139.

Sorkin, B. C., Jones, F. S., Cunningham, B. A., and Edelman, G. M. 1993, Identification of the promoter and a transcriptional enhancer of the gene encoding L-CAM, a calcium-dependent cell adhesion molecule, *Proceedings of the National Academy of Sciences of USA*, **90** (23): 11356-11360.

Stewart, D. C. and Copeland, L. 1998, Uridine 5 '-diphosphate-glucose dehydrogenase from soybean nodules, *Plant Physiology*, **116** (1): 349-355.

Stewart, D. C. and Copeland, L. 1999, Kinetic properties of UDP-glucose dehydrogenase from soybean nodules, *Plant Science*, **147** (2): 119-125.

Stitt, M. and Sonnewald, U. 1995, Regulation of metabolism in transgenic plants, *Annual Review of Plant Physiology*, **46**: 341-368.

Tanaka, A., Mita, S., Ohta, S., Kyozuka, J., Shimamoto, K., and Nakamura, K. 1990, Enhancement of foreign gene expression by a dicot intron in rice but not in tobacco is correlated with an increased level of mRNA and an efficient splicing of the intron, *Nucleic Acids Research,* **18**: 6767-6770.

Taylor, N. J. and Fauquet, C. M. 2002, Microparticle bombardment as a tool in plant science and agricultural biotechnology, *DNA and Cell Biology*, **21** (12): 963-977.

Tenhaken, R. and Thulke, O. 1996, Cloning of an enzyme that synthesizes a key nucleotide-sugar precursor of hemicellulose biosynthesis from soybean: UDP-glucose dehydrogenase, *Plant Physiol*, **112** (3): 1127-1134.

Thangavelu, M., Belostotsky, D., Bevan, M. W., Flavell, R. B., Rogers, H. J., and Lonsdale, D. M. 1993, Partial characterization of the *Nicotiana tabacum* actin gene family: Evidence for pollen specific expression of one of the gene family members, *Molecular and General Genetics*, **240**: 290-295.

Tomov, B. W. and Bernal, J. S. 2003, Effects of GNA Transgenic Sugarcane on Life History Parameters of *Parallorhogas pyralophagus* (Marsh) (Hymenoptera: Braconidae), a Parasitoid of Mexican Rice Borer, *Journal of Economic Entomology*, **96** (3): 570-576.

Turner, W. and Botha, F. C. 2002, Purification and kinetic properties of UDP-glucose dehydrogenase from sugarcane, *Archives of Biochemistry and Biophysics*, **407** (2): 209-216.

Tzafrir, I., Torbert, K. A., Lockhart, B. E., Somers, D. A., and Olszewski, N. E. 1998, The sugarcane bacilliform badnavirus promoter is active in both monocots and dicots, *Plant Molecular Biology*, **38**: 347-356.

Tzfira, T. and Citovsky, V. 2000, From host recognition to T-DNA integration: the function of bacterial and plant genes in the *Agrobacterium*-plant cell interaction, *Molecular Plant Pathology*, **1**: 201-212.

Tzfira, T. and Citovsky, V. 2002, Partners-in-infection: host proteins involved in the transformation of plant cells by *Agrobacterium*, *Trends in Cell Biology*, **12**: 121-128.

Tzfira, T., Li, J., Lacroix, B., and Citovsky, V. 2004, Agrobacterium T-DNA integration: molecules and models, *Trends in Genetics*, **20** (8): 375-383.

Vain, P., Finer, K. R., Engler, D. E., Pratt, R. C., and Finer, J. J. 1996, Intron-mediated enhancement of gene expression in maize (Zea mays L) and bluegrass (Poa pratensis L), *Plant Cell Reports*, **15** (7): 489-494.

Vance, V. and Vaucheret, H. 2001, RNA silencing in plants--defense and counterdefense, *Science*, **292**: 2277-2280.

Vasil, I. K. 1994, Molecular improvement of cereals, *Plant Molecular Biology*, **25** (6): 925-937.

Vasil, V., Clancy, M., Ferl, R. J., Vasil, I. K., and Hannah, L. C. 1989, Increased gene expression by the first intron of maize *shrunken-1* locus in grass species, *Plant Physiology*, **91**: 1575-1579.

Vaucheret, H. and Fagard, M. 2001, Transcriptional gene silencing in plants: targets, inducers and regulators, *Trends in Genetics*, **17** (1): 29-35.

Venter, M. and Botha, F. C. 2004, Promoter analysis and transcription profiling: integration of genetic data enhances understanding of gene expression, *Physiologia Plantarum*, **120**: 74-83.

Wang, J., Jiang, J., and Oard, J. H. 2000, Structure, expression and promoter activity of two polyubiquitin genes from rice (*Oryza sativa* L.), *Plant Science*, **156**: 201-211.

Wang, M. L., Goldstein, C. S., Albert, H. H., & Moore, P. H. 2002, Recombinant protein production in sugarcane in Hawaii, *Plant, Animal and Microbe Genomes X Conference*, *January 2002, Town and Country Convention Centre San Diego, CA*

Wang, Y., Zhang, W., Cao, J., McElroy, D., and Wu, R. 1992, Characterization of *cis*-acting elements regulating transcription from the promoter of a constitutively active rice actin gene, *Molecular and Cellular Biology*, **12** (8): 3399-3406.

Wei, H., Albert, H. H., and Moore, P. H. 1999, Differential expression of sugarcane polyubiquitin genes and isolation of promoters from two highly-expressed members of the gene family, *Journal of Plant Physiology*, **155** (4-5): 513-519.

Wei, H., Wang, M. L., Moore, P. H., and Albert, H. H. 2003, Comparitive expression analysis of two sugarcane polyubiquitin promoters and flanking sequences in transgenic plants, *Journal of Plant Physiology*, **160** (10): 1241-1251.

Witt, H.-J. 1992, UDP-glucose metabolism during differentiation and dedifferentiation of *Reilla helicophylla*, *Journal of Plant Physiology*, **140**: 276-281.

Wu, C. and Morris, J. R. 1999, Transvection and other homology effects, *Current Opinion in Genetics and Development*, **9**: 237-246.

Xu, D., McElroy, D., Thornburg, R. W., and Wu, R. 1993, Systemic induction of a potato pin2 promoter by wounding, methyl jasmonate, and abscisic acid in transgenic rice plants, *Plant Molecular Biology*, **22** (4): 573-588.

Xu, J. and Gong, Z. Z. 2003, Intron requirement for AFP gene expression in *Trichoderma viride*, *Microbiology*, **149**: 3093-3097.

Xu, Y., Buchholz, W. G., DeRose, R. T., and Hall, T. C. 1995, Characterization of a rice gene family encoding root-specific proteins, *Plant Molecular Biology*, **27** (2): 237-248.

Xu, Y., Yu, H., and Hall, T. C. 1994, Rice triosephosphate isomerase gene 5' sequence directs β-glucuronidase activity in transgenic tobacco but requires an intron for expression in rice, *Plant Physiology*, **106**: 459-467.

Yamamoto, Y. Y., Kondo, Y., Kato, A., Tsuji, H., and Obokata, J. 1997, Light-responsive elements of the tobacco PSI-D gene are located both upstream and within the transcribed region, *Plant Journal*, **12** (2): 255-265.

Zhang, W., McElroy, D., and Wu, R. 1991, Analysis of rice *Act1* 5' region activity in transgenic rice plants, *Plant Cell*, **3**: 1155-1165.

Zupan, J., Muth, T. R., Draper, O., and Zambryski, P. 2000, The transfer of DNA from *Agrobacterium tumefaciens* into plants: a feast of fundamental insights, *Plant Journal*, **23**: 11-28.

CHAPTER 3

MOLECULAR CLONING AND CHARACTERISATION OF A GENE ENCODING UDP-GLUCOSE DEHYDROGENASE IN SUGARCANE

3.1 ABSTRACT

UDP-glucose dehydrogenase (EC 1.1.1.22) catalyses the oxidation of UDP-glucose to UDP-glucuronic acid, a precursor for structural polysaccharides incorporated into the cell wall. A gene encoding this enzyme was isolated from a sugarcane genomic library. The gene contains an open reading frame (ORF) of 1443 bp, encoding 480 amino acids with a predicted molecular weight of 52 kDa, which is in agreement with that reported for the sugarcane enzyme (Turner and Botha, 2002). The gene contains one large intron (973 bp), located in the 5'-UTR. The derived amino acid sequence showed 88 – 98% identity with UDP-glucose dehydrogenase from other plant species, and contained highly conserved amino acid motifs required for cofactor binding and catalytic activity. Southern blot analysis indicates a low copy number for UDP-glucose dehydrogenase in sugarcane. The possible expression of multiple gene copies or alleles of a single gene was investigated through comparison of sixteen 550 bp fragments, corresponding to the 5' end of the isolated gene, amplified from cDNA prepared from different tissues. Putative Single Nucleotide Polymorphisms (SNP) and small-scale insertion/deletion (INDEL) were identified from the aligned sequences. A total of 5 bi-nucleotide SNPs and one INDEL sequence were identified. Three of these SNPs are located within the coding region of UDP-glucose dehydrogenase and all correspond to synonymous amino-acid substitutions. A one hundred percent identity of the derived amino acid sequences suggests the expression of different alleles of the same gene rather than expression of multiple copies of UDP-glucose dehydrogenase. This study provides the first evidence for the simultaneous expression of multiple alleles of a single gene in the complex polyploid, sugarcane.

3.2 INTRODUCTION

Nucleotide sugar interconversion pathways represent a series of enzymatic reactions by which plants synthesise activated monosaccharides for incorporation into plant cell wall material (Gibeaut, 2000). In general, sugars are activated by conversion into nucleotide sugars, which act as substrates for the generation of other monosaccharides (Bolwell, 2000; Reiter and Vanzin, 2001). The first step in this series of nucleotide sugar interconversion reactions takes place in the cytosol and is catalysed by UDP-glucose dehydrogenase (EC 1.1.1.22), forming UDP-glucuronic acid in an irreversible reaction (Dalessandro and Northcote, 1977; Nelsestuen and Kirkwood, 1971). UDP-glucose dehydrogenase utilises UDP-glucose, produced either by sucrose synthase or UDP-glucose pyrophosphorylase (Quick and Schaffer, 1997), to form UDP-glucuronic acid in a four-electron oxidation reaction (Nelsestuen and Kirkwood, 1971; Turner and Botha, 2002). The latter is subsequently imported to the Golgi (Gerardy-Schahn *et al*., 2001) where it serves as substrate for both glycosyltransferases and for nucleotide sugar interconversion enzymes which produce precursors for hemicellulose and pectin, including arabinans, arabinogalactans, glucuronoarabinoxylans, rhamnogalacturonans, xylans and xyloglucans (Carpita, 1996; Bolwell, 2000; Gibeaut, 2000). Both hemicellulose and pectin are key components of cell walls, providing a matrix that reinforces the cell wall structure.

UDP-glucuronic acid can also be formed by an alternate pathway, which involves the conversion of *myo*-inositol into glucuronic acid in a reaction catalysed by inositol oxygenase. Glucuronic acid is then conjugated to UDP by a monosaccharide kinase and an uridylyltransferase (pathway in plants reviewed by Loewus and Murthy (2000)). Both of these pathways may exist in plants, their importance depending on the plant species and tissue (Seitz *et al*., 2000).

cDNAs encoding UDP-glucose dehydrogenase have been isolated from soybean (Tenhaken and Thulke, 1996), *Arabidopsis* (Seitz *et al*., 2000) and poplar (Johansson *et al*., 2002). These studies all showed that the enzyme structure of UDP-glucose dehydrogenase is highly conserved between plants and animals, even though the product of the reaction is utilised to produce entirely different polysaccharides in plants (Gibeaut, 2000) and animals (Hempel *et al*., 1994). This suggests strict structural requirements for the correct functioning of the enzyme. Based on Southern blots, UDP-glucose dehydrogenase from the three plant species were all believed to

represent a single copy gene. However, a subsequent evaluation of the *Arabidopsis* database after completion of the genome project indicated the presence of three more copies (Reiter and Vanzin, 2001). The four sequences are highly homologous, with amino acid sequence identities between 83% and 93%. Matching sequences in EST databases show that these genes are all transcribed. Though additional gene copies have not been identified for other species, it is unlikely that *Arabidopsis* is the exception in this regard.

The aim of this study was to isolate and characterise a gene encoding UDP-glucose dehydrogenase from sugarcane. Modern sugarcane cultivars (*Saccharum* spp. Hybrids) are derived from early crosses between *Saccharum officinarum* L, a domesticated species, and *Saccharum spontaneum* L, a wild and vigorous relative (reviewed by Grivet and Arruda, 2001). The resulting hybrids have highly complex aneuploid polyploid genomes. *S. officinarum* usually has 2n=80 chromosomes. *S. spontaneum* has been found to have between 2n=40 and 2n=128. Commercial sugarcane cultivars appear to have 2n chromosome numbers of between 100 and 130 (Butterfield *et al*., 2001), 15% to 25% of which has been inherited from *S. spontaneum* (D'Hont *et al*., 1996). Molecular cytogenetics has shown that the basic chromosome number (x) for *S. officinarum* is x=10 and x=8 for *S. spontaneum* (D'Hont *et al.,* 1998). This implies that for each single copy of a gene, up to ten alleles can be present, each of which potentially corresponds to a distinct haplotype. Among the ten alleles, eight to nine should be inherited from *S. officinarum* and one or two from *S. spontaneum*. Knowledge of sugarcane molecular genetics, however, is insufficient to predict whether all these alleles are expressed, or if some of them may have accumulated sequence changes inhibiting their expression. This complexity potentially complicates gene isolation, as many sequences contained in a genomic library, for example, may represent silent copies of a specific gene.

A recent study by Grivet *et al* (2003), in which ESTs are investigated as a source for sequence polymorphism discovery in sugarcane, provides the first evidence for the expression of multiple alleles of a single gene in sugarcane. This study, however, is based on EST sequences obtained from 27 libraries prepared from 10 sugarcane varieties, so whether these alleles are expressed simultaneously in a single plant is not known. In the present study, a sugarcane gene encoding UDP-glucose dehydrogenase was isolated and characterised. The possible expression of multiple

gene copies and alleles of this gene was also investigated. This is the first UDP-glucose dehydrogenase sequence reported from a monocotyledonous plant.

3.3 MATERIALS AND METHODS

3.3.1 Screening of genomic library

A partial cDNA encoding UDP-glucose dehydrogenase (accession number AA525658) was identified by random sequencing of clones from a cDNA library, prepared from the leaf roll of sugarcane cultivar NCo376 (Carson and Botha, 2000). This sequence was used to screen a sugarcane genomic library constructed from sugarcane cultivar N19. The genomic library was prepared in the Lambda Fix II (Stratagene) cloning vector and had a titre of 1.7×10^8 pfu ml^{-1}. Insert sizes ranged from 9 to 23 kb.

The genomic library was prepared in host cells XL1-Blue MRA (P2). These 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 then centrifuged and resuspended in 10 mM $MgSO_4$. Before use, cells were diluted to an OD_{600} of 0.5 with $MgSO_4$. In the first round, 250 000 plaque forming units (pfu) were screened using 10 plates. For each plate, 200 µl of diluted cells were added to the calculated volume of library (containing 25 000 pfu) and incubated at 37 °C for 15 min. Four ml of melted top agarose (1% (w/v) NZ-amine; 0.5% (w/v) NaCl; 0.5% (w/v) yeast extract; 0.2% (w/v) $MgSO_4.7H_2O$; 0.7% (w/v) agarose) was added to the cell-phage suspension and poured onto pre-warmed NZCYM plates (1% (w/v) NZ-amine; 0.5% (w/v) NaCl; 0.5% (w/v) yeast extract; 0.2% (w/v) $MgSO_4.7H_2O$; 1.5% (w/v) agar). Plates were incubated for 16 h at 37 °C, and then cooled at 4 °C for 2 hours before the phages were transferred to nylon membranes. Dry nylon membranes (MagnaGraph nylon membranes; MSI), were cut to size and 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 M NaOH) for 2 min, neutralised in neutralisation solution (1.5 M NaCl; 0.5 M Tris, pH 8.0) for 5 min and rinsed in 2X SSC (1.75% (w/v) NaCl; 0.88% (w/v) sodium citrate), 0.2 M Tris, pH 7.5 for 30 s. Finally, DNA was linked to the membrane by exposing it to UV light for 2.5 min at 120 mJ cm^{-2}.

Prehybridisation and hybridisation was done in RapidHyb buffer (Amersham) at 42 °C, according to the manufacturers instructions. The probe was prepared by random primer incorporation of α^{32}P-dCTP (Prime It II Random Labeling kit, Stratagene). Following hybridisation, membranes were washed twice with 2x SSC, 0.1% (w/v) SDS at 42 °C for 20 min, followed by 2 x SSC, 0.1% (w/v) SDS at 50 °C for 20 min, and finally, 0.5% SSC, 0.1% (w/v) SDS for 20 min. Membranes were then sealed in plastic wrap and exposed to X-ray film (Kodak) in cassettes fitted with intensifying screens (Amersham) for 24 hours. After visualising the hybridisation results, positive plaques were punched out of the agar plates with the back of a Pasteur pipette. These agar plugs were placed in 500 µl of SM-buffer (0.58% (w/v) NaCl; 0.2% (w/v) $MgSO_4.7H_2O$; 50 mM Tris (pH7.5); 0.1% (v/v) gelatin) with 20 µl chloroform. Second and third round screening were performed as described above, except that fewer pfu's were plated on each plate to allow single plaque isolation.

After three rounds of screening, four positive genomic clones were isolated. Lambda DNA of these clones was isolated using the QIAGEN Lambda mini kit. Plate lysates were prepared according to the method described in Sambrook *et al* (1989) and these lysates were then used directly for the isolation of Lambda DNA.

3.3.2 Characterisation of positive genomic library clones

The Deletion Factory ™ system Version 2.0 (GibcoBRL) was used, according to the manufacturer's instructions, to generate nested deletions, in both directions, across the isolated genomic DNA inserts. Deletions are created by intermolecular transposition of an engineered transposon, Transposon γδ, resulting in a set of overlapping deletion derivatives. Deletion derivatives are then recovered by plating on microbiological media and selecting for loss of a contra-selectable marker (the pDELTA2 cosmid vector used in this system contains two pairs of contra-selectable and selectable genes for isolating nested deletions in both directions). This selection in conjunction with screening for a downstream selectable marker, results in nested deletions that extend various distances from one transposon end into the cloned insert DNA. Screening for deletions in this system allows access to both ends of the cloned DNA in the same experiment. 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. Deletion derivatives were sequenced using the Dye Terminator Cycle Sequencing Ready Reaction with Ampli*Taq* DNA Polymerase FS (Perkin Elmer Applied Biosystems) with vector specific primers. DNASIS for Windows version 2.1 (Hitachi Software) was used to identify overlaps in the sequences.

3.3.3 Isolation of the sugarcane UDP-glucose dehydrogenase gene

Sequences comparison with sequences in international databases using the BLAST program at the National Center for Biological Information (NCBI) showed that the entire coding region, and 690 bp of an intron located in the 5'UTR (and therefore only part of the 5'UTR), had been isolated.

To obtain the rest of the gene, a gene specific reverse primer, UGDRev3 (5'-CTCTTCTGGTAGTCGTTGATC; 902-923 bp downstream of the translation initiation codon, ATG), was used with a forward primer, UGDFw3 (5'-ACGCATCGCGCCAAGGAAGA, approximately 110-90 bp upstream of the translation initiation codon on the cDNA sequence), based on a conserved sequence in the 5'UTR of closely related species (EST sequences obtained from international databases), to amplify the intron located within the 5'-UTR from sugarcane genomic DNA. The PCR reaction was performed in a volume of 50 µl using 50 ng of gDNA as template. The PCR mixture also contained: 1 x PCR buffer, 1.5 mM $MgCl_2$, 200 µM each of dATP, dCTP, dGTP and dTTP, 1 U of *Taq* DNA polymerase (all purchased from Promega) and 0.2 µM of each primer. PCR was performed under the following conditions: 94 °C for 45 s (1 cycle), 94 °C for 45 s and 55 °C for 45 s and 72 °C for 45 s (10 cycles), 94 °C for 30 s and 50 °C for 30 s and 72 °C for 30 s (25 cycles), 72 °C for 2 min (1 cycle). Excess nucleotides and primers were removed using a Qiagen PCR Purification system. This fragment was sequenced and then used to screen the sugarcane genomic library, as described above.

One further genomic clone was obtained. Lambda DNA of this clone was isolated as described above, and digested with *Not* I (all restriction enzymes used were obtained from Promega, and used according to the protocols recommended by them) to remove the insert from the Lambda phage arms. Two restriction fragments were obtained and cloned into the pDELTA2 cosmid

vector. The resulting cosmids were subsequently re-digested using *Not* I in combination with other enzymes. Restriction fragments were then separated on a 1.2 % (w/v) agarose gel, followed by capillary blotting and probing the same probe used to screen the library (transfer of DNA, probe labeling, pre-hybridisation and hybridisation were performed as described for Southern blot analysis below). Positive fragments were subcloned into a pCR-Script (SK+) cloning vector (Promega) according to the manufacturers instructions, and sequenced using vector specific primers. A *Not* I *Xba* I fragment of approximately 4600 bp, containing the complete coding region, as well as the 5'-UTR containing an intron of 973 bp, and a further 1600 bp of the promoter region was identified in this way.

3.3.4 Southern blot analysis

Genomic DNA was isolated from 6 g of sugarcane leafroll or young sorghum leaves, according to Dellaporta *et al*. (1983), and quantified with a spectrophotometer. Five micrograms of genomic DNA was digested with *EcoR* I, *EcoR* V and *Xba* I respectively. The digested DNA was separated in a 0.8 % (w/v) agarose gel and transferred overnight to a positively charged nylon membrane by downward capillary blotting in 10x SSC. DNA was cross-linked to the membrane through exposure to UV light for 2.5 min at 120 $mJcm^{-2}$.

Primer UGDRev3 (described above) was used with UGDFw4 (5'-GCTCGATATCTGGTCACAGATCTATCTG, located between 12 – 22 bp upstream of the ATG; the primer contains an *Eco* RV restriction site (underlined) which was subsequently used to clone the promoter) to amplify a probe of 955 bp from the isolated genomic clone containing a sugarcane UDP-glucose dehydrogenase gene. PCR was performed as described above. The probe was labeled using the Prime-It II random primer labeling kit (Stratagene) and [α-^{32}P] dCTP (Amersham).

Pre-hybridisation and hybridisation were performed in ULTRAhyb^{TM} buffer (Ambion) at 42 °C, according to the manufacturer's instructions. Following hybridisation, the membrane was washed twice in 2x SSC, 0.1% (w/v) SDS for 5 min at 42 °C, then twice in 0.1x SSC, 0.1% (w/v) SDS for 15 min at 42 °C, and finally twice in 0.1x SSC, 0.1% (w/v) SDS for 15 min at 65 °C. The washed membranes were exposed to a Multi Purpose Phosphor Screen for 16 hours and

visualised using a phospho-imager and analysis system (Packard Cyclone; Packard Instrument Company Inc, USA).

3.3.5 Isolation of RNA

Total RNA was isolated from the leafroll, internode 7, roots and buds (from young and old internodes) of mature field grown sugarcane. RNA was extracted from 5 g of each tissue sample according to a method modified from Bugos *et al*. (1995). Tissues were ground to a fine powder in liquid nitrogen and added to 25:24:1 phenol:chloroform:isoamyl alcohol. After vortexing, an equal volume of homogenisation buffer (0.1 M Tris-HCl (pH 7.5), 1 mM EDTA, 0.1 M NaCl and 1% (w/v) SDS) was added. Sodium acetate (pH 5.2) was then added to a final concentration 0.1 M. The emulsion was mixed, incubated on ice for 15 min and centrifuged at 4 °C (12 000 *g*, 15 min). The aqueous phase was subsequently transferred to a new tube and RNA precipitated through the addition of one volume of isopropanol followed by incubation at – 70 °C for at least 30 min. Precipitated RNA was recovered by centrifugation at 4 °C (10 000 g, 10 min). Excess salts were removed from the pellet by washing with 70% (v/v) ethanol. The pellet was air-dried and resuspended in 750 µl of diethyl pyrocarbonate (DEPC) treated water. Insoluble particles were removed by centrifugation (10 000 g, 5 min). The supernatant was transferred to a microcentrifuge tube, and RNA precipitated again using LiCl at a final concentration of 2 M. To maximize RNA precipitation, samples were incubated overnight at 4 °C. RNA was pelleted by centrifugation at 4 °C (12 000 *g*, 15 min). The pellet was again washed with 70 % (v/v) ethanol, and resuspended in DEPC treated water. Remaining insolubles were removed by centrifugation (10 000 *g*, 5 min). RNA concentration was determined spectrophotometrically.

3.3.6 Preparation of cDNA

First strand cDNA synthesis was performed from 5 µg of total RNA using SUPERSCRIPT^{TM} II Rnase H^{-} Reverse Transcriptase (GibcoBRL) according to the manufacturers instructions. The resulting cDNA was further purified using a Qiagen PCR Purification kit.

3.3.7 Amplification of UDP-glucose dehydrogenase from cDNA from different tissues

Primers UGDFw3 and UGDRev3 were used to amplify a 1042 bp fragment of the UDP-glucose dehydrogenase gene from 1 ng of total cDNA prepared from field-grown sugarcane leafroll,

internode 7, roots and buds (taken from young and old internodes). Four duplicate reactions were performed for each tissue (PCR was performed as described above). Amplified fragments from the same tissues were pooled, cloned using a pGEM®-T Easy Vector System (Promega) and used to transform DH10B competent cells (GibcoBRL).

Plasmid DNA was isolated from 4 separate positive colonies per tissue using the Qiagen High Pure Plasmid Purification kit. Sixteen cloned amplified fragments were purified, and the orientation of the inserts was determined by restriction digestion. Clones were all sequenced once from the 5'-end using vector specific primers and a Dye Terminator Cycle Sequencing Ready Reaction with Ampli*Taq* DNA Polymerase FS (Perkin Elmer Applied Biosystems).

3.3.8 Analysis of UDP-glucose dehydrogenase from cDNA from different tissues

DNASIS for Windows version 2.1 (Hitachi Software) was used to align and compare the first-pass 5' sequences. To avoid poor quality sequences only 550 bp, including 114 bp of the 5'-UTR and the first 436 bp of the ORF, of each sequence was retained. Putative Single Nucleotide Polymorphisms (SNP) and small-scale insertion/deletion (INDEL) were identified from the aligned sequences. A SNP or INDEL site was only retained if the surrounding adjacent sequences were perfectly homologous for all the sequences, and the least frequent variant at the site occurred at least twice.

Homologous sugarcane ESTs from different sugarcane varieties and libraries, encoding UDP-glucose dehydrogenase genes, were identified using the BLAST program at the NCBI. The threshold was a BLASTN score $\geq$ 800 with the query sequence. Sequences that met this criteria were retrieved from GenBank and investigated for the presence or absence of the SNPs and INDELs identified in the ESTs produced from sugarcane variety N19 leafroll, internode 7, roots and buds.

3.4 RESULTS

3.4.1 Isolation of the sugarcane UDP-glucose dehydrogenase gene

Initial screening of a sugarcane genomic library using a partial cDNA encoding UDP-glucose dehydrogenase as a probe led to the isolation of four clones. Sequence analysis of these clones revealed that only one of them contained the complete coding region of the gene with an additional 720 bp of the upstream sequence. The clone contained an open reading frame (ORF) of 1443 bp, encoding 480 amino acids with a predicted molecular weight of 52 kDa, similar to that reported for sugarcane UDP-glucose dehydrogenase (Turner and Botha, 2002). Comparison to sequences in international databases confirmed that this sequence showed extensive homology to UDP-glucose dehydrogenases from other species. Near perfect homology was observed between the query sequence and a partial sorghum sequence. This sequence is part of BAC clone 170F8 (GenBank accession number AF503433) and only contains the first 650 nt of the coding region. In this area, an identity of 97% on a nucleotide level and 100% at an amino acid level was observed. Identity to other species for which the full-length sequences were available is shown in Table 3.1.

Table 3.1 Identity of UDP-glucose dehydrogenase from sugarcane to other plant species at a nucleotide (nt) (ORF only) and amino acid (aa) level.

UGDH source	GenBank accession number	% Identity (similarity)	
		nt	aa
Zea Mays	AY103689	95	98 (99)
Oryza Sativa	AK103919	87	95 (97)
Glycine Max	GMU53418	78	91 (95)
Arabidopsis Thaliana	AYO88902	77	89 (95)
Populus Tremula	AF053973	76	88 (94)

% Identity refers to identical residues, % similarity refers to residues with similar properties

Analysis of the deduced amino acid sequence revealed that all the structural features, initially identified for the bovine enzyme (Hempel *et al*., 1994), were present. These include the NAD cofactor binding site, the catalytic site with a centered Cys residue, and the two Pro residues believed to represent the main chain bends in the protein structure (Figure 3.1).

The right border of an intron was found 24 bp upstream of the translation initiation site. Although the complete sequence of the coding region was obtained, the left border of the intron, as well as the rest of the 5'-UTR sequence, was not contained in the genomic clone. To obtain the missing sequence, a primer was designed based on a conserved sequence, located approximately 110 – 90 bp upstream of the translation initiation site in the 5'UTR region of the closely related species maize and sorghum, and used in conjunction with a sugarcane gene specific primer to amplify a fragment of approximately 1200 bp from sugarcane genomic DNA. Sequence analysis confirmed that this fragment contained an intron sequence of 973 bp (located within the 5 'UTR) as well as a part of the 5' end of a sugarcane UDP-glucose dehydrogenase gene. This fragment was used to re-screen the sugarcane genomic library. One further genomic clone was obtained. From this clone a restriction fragment of approximately 4600 bp, which contained the complete coding region, 5'-UTR and intron, as well as 1600 bp of upstream sequence, was cloned. Figure 3.1 shows the nucleotide and deduced amino acid sequences of the gene.

3.4.2 Southern blot analysis

The genomic complexity of sugarcane and sorghum UDP-glucose dehydrogenase was investigated by digesting genomic DNA with selected restriction enzymes, followed by analysis on a Southern blot using a sugarcane UDP-glucose dehydrogenase probe. Restriction enzymes used to prepare the genomic DNA all had a restriction site outside of, but close to the coding region. A very simple and very similar banding pattern was observed for both of the species investigated, which probably reflects a low copy number (Figure 3.2).

3.4.3 Analysis of UDP-glucose dehydrogenase from cDNA prepared from different tissues

Gene specific primers were used to amplify a fragment of the UDP-glucose dehydrogenase gene from cDNA prepared from sugarcane leafroll (LR), internode 6 (I6), roots (RT) and buds (BD). Four cloned sequences per source tissue were sequenced from the 5' end. DNASIS for Windows version 2.1 (Hitachi Software) was used to align and compare 550 bp, including 114 bp of the 5'-UTR and the first 436 bp of the ORF, of the first-pass 5' sequences.

TGCTTCTGGTTTCTGACCAGCCTTGACGCGGTTACTATTTTTTTTTTCTACGGGACACGGCACGGTTATTTTTTG -80

AAGCCGAGGGGTAGATCTGGTTCGGTGGGTTTGCGGGCCGGGGTATTTAAGGACTCGTGGCCTCGCTTCTCCTCT -15

GCACACACACAAAGCGAGCTCCTCTTCTCCCACGCATCGCGCCAAGGAAGAGAGAGATCGCCGCTCCTCTTCTCT 60

TCGTCGCCTCTACTGGTGTGGGTGCCCGAGATCTTGCTGCAGTTCCTCGATATCTGGTGAGACTCGCCCACGCTG 135

TTACCGCCCTGATTTCAATTTTTTTTCCTTTCTGATGAGTCCCGCGCCTGATCATG c t a g t t g a g g a t c g t g a t t 210

t c a t t c t g g g g c t a g t t t g t t c c t c g g c t t c g t t c t g c t g c a a g c a g t t g t c t c c t c g a g t t c g t t c t c t g a g g a 285

c g a c c c t g t c t t c a g a t c t g g a a g t t t c g t g g t t t g g a t c t g a t a g g a g t t a c g g t t a c g c g t t g t a g t g g c t g a 360

a g a g g a t t t g c g c c c c c c t a a t g t t g c c t t t g c a c c c c t a c t g g a t c a g a g t t t c a g a t c t t t c t g t c t g g t c g c 435

a g t t t t a a a c g c c t t g a c g g t g c t c a t a t a c g a g t a a a t t t c a t t t t a t a a t t c g g t g g t c a t g g a t c t g t a t c c 510

t g c g a t t c c t g c c c t t c a c t c t g t t t a g g t g g t g t t t g t g t c a g g t a a c g g c a g t g c a c g t a a t g g g a a t g g g a t 585

c a c a g c t g c c t g g t t g t g t a a t a g g t g g t t g c t t a a t t t g t c t g g t t t t t c t t a g a a a a g g t t t t a g a a a g c a t t 660

a a t t t c a c a t g t a t t c g g t g g a g a a g a g g g g c c a t c c c a a t c t t g a t c a g c a g a a a a c c t g a g a t g t g c c g t g c c 735

t g c t t t a a c a a a t a t g a g t t c t g a t g a a t g t g t c c t a a c a c t t c c a c c c a a a a c a a a c c t t t t t a g a t g t g g t t t 810

g a c c t g g t c g g t a g a t g t t g c c a a t t g c t a t a c c t a g t a t a t c c g a a c t c c t a t t t a t t c a g t t c a c t t a a g t a a 885

c c t t c t c g g a a c a t g c t t a g c t a t g c t c t g t t a t t t c a g t c g t g c a c t a g t a g t c t a g t g t t c a g g a t c t g g t g c 960

t t t t g c t t g t g a a t c c t g a t t g c t t c t g t t a a c t t g g g a a a c t t g g t t g a c a t g t c t c a t c t g t c a g a t a t t c t c 1035

a g g g t a t g t g c t c t a t t t t a c a a t c c a c t a a t g a a g c a c c t t t t g c a a t t a c a g GTCACAGATCTATCTGACAAG 1110

ATGGTGAAGATCTGCTGCATCGGTGCTGGCTATGTCGGTGGCCCGACCATGGCCGTCATTGCCCTCAAGTGCCCA 1185
M V K I C C I G A G Y V G G P T M A V I A L K C P 25

GACATTGAGGTTGTTGTTGTTGACATCTCCAAGCCCCGCATTGAGGCCTGGAACAGCGACACCCTCCCGATCTAC 1260
D I E V V V V D I S K P R I E A W N S D T L P I Y 50

GAGCCTGGCCTCGATGATGTTGTGAAGCAGTGCAGGGGGAAGAACCTCTTCTTCAGCACTGATGTTGAAAAGCAT 1335
E P G L D D V V K Q C R G K N L F F S T D V E K H 75

GTCGCTGAGGCTGACATTATTTTCGTCTCGGTGAACACCCCCACCAAGACCCGTGGGCTTGGAGCTGGCAAGGCT 1410
V A E A D I I F V S V N T P T K T R G L G A G K A 100

GCCGACCTCACCTACTGGGAGAGTGCTGCTCGTATGATCGCTGATGTCTCCAAGTCTGACAAGATTGTTGTTGAG 1485
A D L T Y W E S A A R M I A D V S K S D K I V V E 125

AAGTCCACCGTCCCTGTCAAGACTGCTGAGGCCATTGAGAAGATCTTGACCCACAACAGCAAGGGCATCAACTAC 1560
K S T V P V K T A E A I E K I L T H N S K G I N Y 150

CAGATCCTTTCCAACCCAGAGTTCCTTGCGGAGGGCACTGCTATTGAGGACCTGTTCAAGCCTGACAGAGTGCTC 1635
Q I L S N P E F L A E G T A I E D L F K P D R V L 175

ATCGGTGGTCGGGAGACCCCTGAGGGCAGGAAGGCCGTCCAGGCTCTCAAGGATGTGTACGCTCACTGGGTTCCC 1710
I G G R E T P E G R K A V Q A L K D V Y A H W V P 200

GAGGACAGGATCCTCACCACCAACCTGTGGTCTGCTGAGCTCTCCAAGCTCGCTGCCAACGCGTTCTTGGCACAA 1785
E D R I L T T N L W S A E L S K L A A N A F L A Q 225

AGGATCTCCTCTGTGAATGCCATCTCCGCCCTCTGCGAAGCAACTGGTGCCAATGTGTCTGAGGTGGCTTACGCC 1860
R I S S V N A I S A L C E A T G A N V S E V A Y A 250

GTGGGCAAGGACACCAGGATTGGCCCCAAGTTCCTGAACGCCAGTGTTGGGTTCGGTGGCTCTTGCTTCCAGAAG 1935
V G K D T R I G P K F L N A S V G F G G S C F Q K 275

GACATCCTGAACTTGGTGTACATCTGCGAGTGCAATGGCCTGCCCGAGGTCGCCAACTACTGGAAACAGGTGATC 2010
D I L N L V Y I C E C N G L P E V A N Y W K Q V I 300

AAGATCAACGACTACCAGAAGAGCCGGTTCGTTAACCGCGTTGTGTCCTCCATGTTCAACACCGTCGCCGGCAAG 2085
K I N D Y Q K S R F V N R V V S S M F N T V A G K 325

AAGATCGCTGTCCTCGGCTTCGCCTTCAAGAAGGACACCGGTGACACCAGGGAGACCCCGGCCATTGACGTCTGC 2160
K I A V L G F A F K K D T G D T R E T P A I D V C 350

```
AAGGGCCTGTTGGGTGACAAGGCCCAGATCAGCATCTACGACCCCCAGGTGACGGAKGACCAGATYCAGCGGGAC  2235
K  G  L  L  G  D  K  A  Q  I  S  I  Y  D  P  Q  V  T  X  D  Q  I  Q  R  D     375

CTGGCCATGAACAAGTTCGACTGGGACCACCCGATGCACCTGCAGCCAACGAGCCCCACGGCCGTGAAGCAGGTG  2310
L  A  M  N  K  F  D  W  D  H  P  M  H  L  Q  P  T  S  P  T  A  V  K  Q  V     400

AGATGCGTGTGGGACGCGTACGAGGCCACCAAGGGTGCCCACGGGCTGTGCATCCTGACCGAGTGGGACGAGTTC  2385
R  C  V  W  D  A  Y  E  A  T  K  G  A  H  G  L  C  I  L  T  E  W  D  E  F     425

AAGACCCTGGACTACCAGAAGATCTTCGACAACATGCAGAAGCCTGCCTTCGTCTTCGATGGCCGCAACATCGTC  2460
K  T  L  D  Y  Q  K  I  F  D  N  M  Q  K  P  A  F  V  F  D  G  R  N  I  V     450

GACGCCGAGAAGCTGAGGGAGATCGGCTTCATCGTCTACTCCATCGGCAAGCCGCTGGACGCCTGGCTCAAGGAC  2535
D  A  E  K  L  R  E  I  G  F  I  V  Y  S  I  G  K  P  L  D  A  W  L  K  D     475

ATGCCCGCGGTCGCTTAATTCCACCCTTCATCCGAGGTGCTCCATGGATTGAATTCGGGGAACGAAGAGGAACTG  2610
M  P  A  V  A  *                                                              480

GTTGACCATTCTTTATTGCAGTTTGTTTTTGCAGGCTACGCTACGATTTCTCTCGTGTCAGGCATAAAAGTAAAA  2685

GTTGGACGGCGCTGCTAGTATTCCCTGTTCGGTTGGTGTATTTCATGGTTGGAGGACGTCTGTAGATGTAACAAT  2760

CCTCAGGCCCTCGTGCTCTGCTTGAGGAAATTTCCGTGTACGGTTGTACCCTGGACCTGCTATAGCCGGTTGATT  2835

CTTCAATTGTATTCCTAAAAGTTATACGATCGATATGTTTCTTTTTTGCAGAATAACAATTCTTGCTCACGTTTC  2910

ACCCTTTT                                                                      2918
```

Figure 3.1 Nucleotide (nt) and derived amino acid (aa) sequence of a sugarcane gene encoding UDP-glucose dehydrogenase. The gene contains a single intron located in the 5'-UTR (lower case in figure), followed by a continuous open reading frame encoding 480 aa. A putative transcription initiation site, based on homology to known full length cDNA sequences and ESTs, is numbered as nt 1, and underlined with an arrow. The stop codon is indicated by an asterisk at base pair 2551. The highly conserved NAD cofactor binding site (aa 8-18) is boxed. Pro residues located at 89 and 156 (boxed) have been reported to represent main bends in the protein structure (Hempel *et al*., 1994). The catalytic site (aa 267-278) with a centered Cys residue (aa 273) is shaded in gray. Two conserved Lys residues (circled) are located at positions 216 and 335. The above-mentioned features are perfectly conserved for all known plant UDP-glucose dehydrogenases (Johansson *et al.*, 2002).

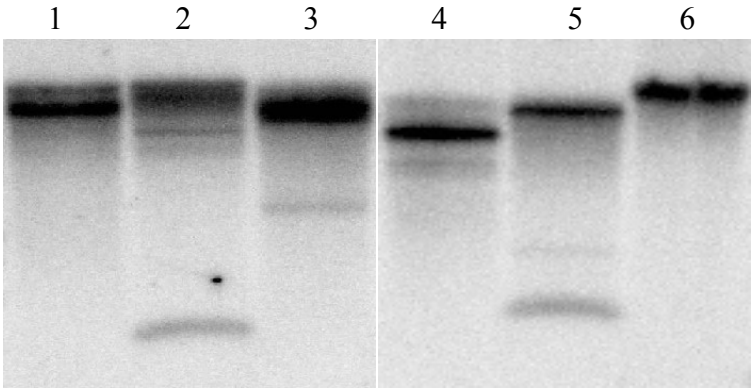


Figure 3.2 Southern analysis of sugarcane (1-3) and sorghum (4-6) UDP-glucose dehydrogenase. A 944 bp fragment of the 5' end of the sugarcane gene (PCR product of primers UGD Fw4 and UGD Rev3) was used to probe 5 µg of completely digested gDNA: (1, 4) *Eco* RI; (2, 5) *Eco* RV; (3, 6) *Xba* I.

Using the criteria described under materials and methods, putative SNPs and INDELs were identified from the aligned sequences. A total of 4 bi-nucleotide SNPs were identified in this way. Three of these SNPs are located within the coding region of UDP-glucose dehydrogenase and all correspond to synonymous amino-acid substitutions. The fourth SNP is located in the non-coding 5'-UTR. Another SNP, which did not meet the selection criteria as it occurred in only one of the sequences, LR2, was also retained. This putative SNP, located in the 5'-UTR, 30 bp upstream of the translation initiation site (ATG), can either be a 'G' (15/16 sequences) or a T (1/16 sequences). A 'T' at this position creates an *Eco* RV restriction site. This restriction site was present in the genomic sequence isolated after screening a sugarcane genomic library, and was therefore retained as a legitimate SNP for further investigation. Only one INDEL of 6 bp, located in the 5'-UTR, was detected. This INDEL was present in half of the sequences (i.e. 8) and only one variant was observed. Combining the five SNPs and the INDEL leads to the identification of seven distinct haplotypes. These results are summarised in Table 3.2.

To investigate whether these alleles were also expressed in other sugarcane varieties, the BLAST program at the NCBI was used to identify homologous sequences from the Brazilian Sugarcane EST Project (SUCEST). This project recently produced about 230 000 randomly cloned cDNAs from 27 libraries prepared from 10 sugarcane varieties (**http://sucest.lad.ic.unicamp.br/en/**). Sequences with a BLASTN score $\geq$ 800 with the query sequence, were retrieved from GenBank and investigated for the presence or absence of the SNPs and INDEL described above. Seventy-one sequences that covered the 550 bp of the UDP-glucose dehydrogenase gene originally aligned were retrieved. With regards to the four SNPs and the INDEL, sixty-nine of these sequences matched one of the seven putative alleles (Table 3.2). One allele, initially represented by only one sequence, was not matched by any of the ESTs. Though the probability that a specific SNP, only retained if the surrounding adjacent sequences were perfectly homologous for all the sequences, and the least frequent variant at the site occurred at least twice, is quite high, the appearance of an allele, i.e. the combination of several SNPs and INDELS, may still be the result of a sequencing error as all the aligned sequences were first-pass 5' sequences. Also, the reliability of SNP detection, and therefore detection of specific alleles, increases with the volume of sequences available. Sugarcane ESTs identical to the other 6 putative alleles (with regards to the SNPs and INDEL) support the validity of the results obtained.

Table 3.2 Sequence polymorphisms inside the sugarcane UDP-glucose dehydrogenase gene, observed in partial cDNA sequences derived from the leafroll (LR), internode 6 (I6), roots (RT) and buds (BD) of sugarcane cultivar N19. Putative SNPs and an INDEL were identified after alignment of the 16 sequences identified in column 1 (selection criteria described in the text). The position of each site is given relative to the translation initiation site. The consensus line gives the majority variant observed at each site (only one variant of the INDEL sequence was observed therefore this sequence is given in the consensus line). A '*' indicates that the cDNA sequence is identical to the consensus line. The presence of the INDEL is indicated by a '+' and it's absence by a '-'. Seven distinct haplotypes are identified in column eight. Homologous sugarcane [1]ESTs from different sugarcane varieties and libraries were identified using the BLAST program at the NCBI. Seventy-one highly homologous sequences (BLASTN score ≥ 800) were retrieved from GenBank and investigated for the presence or absence of the observed SNPs and INDEL. Sixty-nine of these matched one of the seven identified haplotypes (shown in the last column).

Sequence	**INDEL**	**SNP**						
Consensus	**CTCTAC**	**C**	**G**	**C**	**A**	**T**	**Haplotype**	**Number of matching ESTs retrieved from GenBank**
Position	-65	-30	-27	180	210	247		
LR 1	-	T	*	*	G	C	1	40
LR 2	+	*	T	*	*	*	2	7
LR 3	+	*	*	G	*	*	3	3
LR 4	+	*	*	*	*	*	4	9
I6 1	+	*	*	G	*	*	3	3
I6 2	-	T	*	G	*	*	5	0
I6 3	+	*	*	*	*	*	4	9
I6 4	-	T	*	*	G	C	1	40
RT 1	-	T	*	*	G	C	1	40
RT 2	+	*	*	*	*	*	4	9
RT 3	-	T	*	*	G	C	1	40
RT 4	+	*	*	*	*	*	4	9
BD 1	+	*	*	G	*	*	3	3
BD 2	-	*	*	*	G	C	6	4
BD 3	-	T	*	*	G	C	1	40
BD 4	-	*	*	*	*	*	7	6

[1]ESTs retrieved from GenBank all came from the Brazilian Sugarcane EST Project (SUCEST), which recently produced about 230 000 randomly cloned cDNAs from 27 libraries prepared from 10 sugarcane varieties (http://sucest.lad.ic.unicamp.br/en/).

3.5 DISCUSSION

UDP-glucose dehydrogenase catalyses the oxidation of UDP-glucose to UDP-glucuronic acid, a precursor for sugar nucleotides, which are incorporated into pectin and hemicelluloses. Enzyme activity of UDP-GlcDH is correlated with growing and expanding tissue that have a demand for hemicellulose precursors. UDP-glucose dehydrogenase was recently purified from sugarcane (Turner and Botha, 2002). Characterisation of the kinetic properties of this enzyme revealed significant differences between UDP-glucose dehydrogenase from sugarcane (Turner and Botha, 2002), the first UDP-glucose dehydrogenase isolated from a monocot, and those previously isolated from dicotyledonous plant species (Robertson *et al*., 1996; Stewart and Copeland, 1998). It was suggested that these differences may represent divergence between monocotyledonous and dicotyledonous plant species. In the current study a gene encoding UDP-glucose dehydrogenase was isolated from sugarcane, representing the first sequence of this gene isolated from a monocotyledonous species. The gene contains a single intron of 973 bp located in the 5'-UTR, followed by a continuous open reading frame encoding 480 aa. A predicted molecular mass for the derived amino acid sequence of 52 kDa is in agreement with that found for sugarcane UDP-glucose dehydrogenase (Turner and Botha, 2002). Introns located in the 5'-UTR have been documented in many cases to have a large positive effect on gene expression, especially in monocots (Christensen *et al*., 1992; Luehrsen and Walbot, 1991; Maas *et al*., 1991; McElroy *et al*., 1991; Rethmeier *et al*., 1997; Rethmeier *et al*., 1998). As all plant sequences that encode UDP-glucose dehydrogenase isolated to date have been cDNA sequences, an intron in this position has not been reported for other plant species. However, PCR amplification from genomic DNA of sugarcane, maize and sorghum, using primers located on either side of the intron (UGDRev3 and UGDFw3, as described in materials and methods), resulted in a product of similar size for all three species (results not shown). This indicates that UDP-glucose dehydrogenase genes from these closely related monocots contain this 5'-UTR intron. Also, an investigation of sequences available in Genbank confirms the presence of an intron in this position for UDP-glucose dehydrogenase from sorghum, rice and *Arabidopsis* (GenBank accession numbers AF503433, AL731873 and AL391143, respectively). All these species contain a large intron (sugarcane, 973 bp, sorghum 962 bp, rice 651 bp, and *Arabidopsis* 773 bp) located 20 to 30 bp upstream of the translation initiation site.

Within the coding region, the isolated sequence contains all of the conserved motifs previously described for UDP-glucose dehydrogenases from plant (Johansson *et al*., 2002; Tenhaken and Thulke, 1996) and animal sources (Franzen *et al*., 1981; Hempel *et al*., 1994). The highly conserved GXGXXGG pattern of the NAD cofactor binding site is found near the N-terminus (aa 8-14). This cofactor binding site and the following four amino acids (aa 15-18) are conserved for all plant species examined (Johansson *et al*., 2002). Similarly conserved are two Pro residues, located at positions 89 and 156 for sugarcane, believed to represent main bends in the protein structure (Hempel *et al*., 1994). Two Lys residues (aa 216 and 335), which correspond to Lys 219 and Lys 338 of bovine UDP-glucose dehydrogenase, are also found in the sugarcane sequence. One of these Lys residues is probably catalytically involved in the first half-reaction of the enzyme (conversion of UDP-glucose to UDP-aldehydoglucose) (Hempel *et al*., 1994). The main catalytic site (Figure 3.1, aa 267-278) of UDP-glucose dehydrogenase was initially identified for the bovine enzyme by chemical modifications of amino acids and subsequent peptide sequencing (Franzen *et al*., 1981). For bovine UDP-glucose dehydrogenase, a centered Cys residue (aa 273 in sugarcane) supposedly involved in the second-half of the reaction catalysed by UDP-glucose dehydrogenase (conversion of UDP-aldehydoglucose to UDP-glucuronate) is essential for the functioning of the enzyme (Hempel *et al*., 1994). A recent study of soybean UDP-glucose dehydrogenase (Hinterberg *et al*., 2002) showed that this enzyme is strongly inhibited by reagents that modify cystein groups in proteins, confirming the essential role of this residue in plant UDP-glucose dehydrogenases.

Besides these conserved motifs, the entire sugarcane sequence is highly homologous to sequences isolated from other plant species (Table 3.1). Identity of 91% (with a similarity of 95%) between the sugarcane and soybean sequences suggests that the differences observed in the kinetic properties of UDP-glucose dehydrogenase purified from these two species (Stewart and Copeland, 1999; Turner and Botha, 2002), is probably not a result of structural differences in the protein.

The genomic complexity of sugarcane and sorghum UDP-glucose dehydrogenase was investigated by Southern blot analysis. Sorghum appears to be an excellent model for sugarcane, as it is a closely related diploid with a monoploid genome size very close to that predicted for

sugarcane (Butterfield *et al*., 2001), and its chromosomes are collinear with sugarcane (Dufour *et al*., 1997). Restriction enzymes used to prepare the genomic DNA all had a restriction site outside of, but close to the coding region. A very simple and very similar banding pattern was observed for both of the species investigated, which probably reflects a low copy number (Figure 3.2). Based on Southern blots, UDP-glucose dehydrogenase from soybean (Tenhaken and Thulke, 1996), *Arabidopsis* (Seitz *et al*., 2000) and poplar (Johansson *et al*., 2002) were all believed to represent a single copy gene. However, a subsequent evaluation of the *Arabidopsis* database after completion of the genome project indicated the presence of three more copies (Reiter and Vanzin, 2001). The four sequences are highly homologous (83% - 93% identity) and are all transcribed. This means that Southern blot analysis is not an accurate technique for determining the exact copy number of a gene within a genome. It can at best indicate a relatively high or low copy number.

Commercial sugarcane cultivars appear to have 2n chromosome numbers of between 100 and 130 (Butterfield *et al*., 2001). This implies that for each single copy of a gene, up to ten alleles can be present, each of which potentially corresponds to a distinct haplotype. Not enough is known about sugarcane molecular genetics to predict whether all these alleles are expressed. To investigate the possibility that sugarcane expresses multiple alleles or more than one homologous copy of UDP-glucose dehydrogenase, a 550 bp fragment of the 5' end of the gene, including 114 bp of the 5'-UTR and the first 436 bp of the ORF, was amplified from cDNA prepared from different tissues. In total, sixteen sequences amplified from four different tissues were compared. Though differences were observed between sequences, a 100% identity of the derived amino acid sequences of all the clones indicated allelic variation within a single gene rather than multiple different homologous gene copies. As it is possible, even likely, that sugarcane, like *Arabidopsis*, contains more than one gene copy of UDP-glucose dehydrogenase, it could be that the primers used to amplify the 5' end of the gene are specific for one of these copies. Also, an inducible UDP-glucose dehydrogenase or highly tissue specific copy may not be represented in the cDNA pools used. Sugarcane EST projects are currently underway in South Africa (Carson and Botha, 2000; Carson and Botha, 2002; Carson *et al*., 2002), Australia (Casu *et al*., 2003) and Brazil (**http://sucest.lad.ic.unicamp.br/en/**). More than 230 000 sugarcane EST sequences are

currently available at the NCBI. A search of these sequences will probably reveal additional expressed copies of UDP-glucose dehydrogenase.

Putative Single Nucleotide Polymorphisms (SNP) and small-scale insertion/deletion (INDEL) were identified from the aligned sequences. A total of 5 bi-nucleotide SNPs and one INDEL sequence were identified (Table 3.2). Three of these SNPs are located within the coding region of UDP-glucose dehydrogenase and all correspond to synonymous amino-acid substitutions. The mean density of SNPs in the part of the coding region examined was therefore one every 145 bp, similar to that reported by Grivet *et al.* (2003) for sugarcane *Adh* genes (one every 122 bp). This is much higher than the one every 1500 (Deutsch *et al*., 2001) or 3330 bp (Garg *et al*., 1999) reported for human ESTs. The other two SNPs are located in the non-coding 5'-UTR. One of the SNPs in the 5'-UTR creates or removes an *Eco* RV restriction site. Although only one amplified sequence contained this restriction site, the presence of this SNP gives a plausible explanation for the presence of an additional fainter band (one allele rather than one gene copy) observed on the Southern blot when genomic DNA is digested using *Eco* RV. Only one INDEL of 6 bp, located in the 5'-UTR, was detected. This INDEL was present in half of the sequences (i.e. 8) and only one variant was observed. The particularly high density of polymorphisms detected within the 5'-UTR, 3 in 114 bp of 5'-UTR, is consistent with the findings of Grivet *et al*. (2003). The two SNPs and INDEL occur within a 40 bp fragment of the 5'-UTR. Based on results obtained for sugarcane *Adh* genes and other unpublished results, Grivet *et al.* (2003) suggest that the high level of variation, detected over short sequence of the 5'-leader, may be a common feature of sugarcane genes. Combining the five SNPs and the INDEL leads to the identification of seven distinct haplotypes. It is, of course, necessary to consider that it is possible to confuse alleles of the same gene, occupying the same locus on a particular chromosome, with recently diverged paralogous loci. However, the fact that all of these SNPs located within the coding region correspond to synonymous amino-acid substitutions, therefore encoding an identical protein (at least for the first 145 amino acids) supports that these haplotypes represent different alleles of the same gene.

To investigate whether these alleles were also expressed in other sugarcane varieties, homologous sequences from the Brazilian Sugarcane EST Project (SUCEST) were identified. Seventy-one

sequences that covered the 550 bp of the UDP-glucose dehydrogenase gene originally aligned, were retrieved. With regards to the four SNPs and the INDEL, sixty-nine of these sequences matched one of the seven putative alleles (Table 3.2). Almost two thirds of these sequences (40/69) corresponded to a single haplotype. In the current study, this haplotype was also represented by the most sequences (5/16) and amplified from every tissue. There are two possible explanations for the frequent occurrence of a specific sequence. It could be that this allele occurs more frequently than others. Another possibility is that the expression of this allele is regulated by a stronger promoter than that of other alleles. Further studies of allelic variation and expression of these alleles in sugarcane are required to fully explain this phenomenon.

This study provides the first evidence for the simultaneous expression of distinct haplotypes in one sugarcane plant. As the reliability of detected polymorphisms, and therefore detection of specific alleles, increases with the volume of sequences available, the presence in the SUCEST database of expressed sequences homologous to the identified haplotypes (with regards to the SNPs and INDEL), supports the validity of the results obtained. Whether this is a common feature of sugarcane genes is not known at present. However, the possibility that multiple alleles are expressed to provide the required levels of a specific enzyme, rather than the increased expression of one dominant allele is encouraging for sugarcane gene and promoter isolation, as many sequences contained in a genomic library, for example, may represent expressed copies of a specific gene adjacent to active promoters.

3.6 REFERENCES

Bolwell, G. P. 2000, Biosynthesis of plant cell wall polysaccharides, *Trends in Glycoscience and Glycotechnology*, **12** (65): 143-160.

Bugos, R. C., Chiang, V. L., Zhang, X. H., Campbell, E. R., Podila, G. K., and Campbell, W. H. 1995, RNA isolation from plant tissues recalcitrant to extraction in guanidine. *BioTechniques*, **19**: 734-737.

Butterfield, M., D'Hont, A., and Berding, N. 2001 The sugarcane genome: A synthesis of current understanding, and lessons for breeding and biotechnology. *Proc Soc Afr Sugarcane Technol Ass*, **75**: 1-5.

Carpita, N. C. 1996, Structure and biogenesis of the cell walls of grasses, *Annual Review of Plant Physiology and Plant Molecular Biology,* **47**: 445-476.

Carson, D. L. and Botha, F. C. 2000, Preliminary analysis of expressed sequence tags for sugarcane, *Crop Science*, **40** (6): 1769-1779.

Carson, D. L. and Botha, F. C. 2002, Genes expressed in sugarcane maturing internodal tissue, *Plant Cell Reports*, **20** (11): 1075-1081.

Carson, D. L., Huckett, B. I., and Botha, F. C. 2002, Sugarcane ESTs differentially expressed in immature and maturing internodal tissue, *Plant Science*, **162** (2): 289-300.

Casu, R. E., Grof, C. P. L., Rae, A. L., McIntyre, C. L., Dimmock, C. M., and Manners, J. M. 2003, Identification of a novel sugar transporter homologue strongly expressed in maturing stem vascular tissues of sugarcane by expressed sequence tag and microarray analysis, *Plant Molecular Biology*, **52**: 371-386.

Christensen, A. H., Sharrock, R. A., and Quail, P. H. 1992, Maize polyubiquitin genes: structure, thermal perturbation of expression and transcript splicing, and promoter activity following transfer to protoplasts by electroporation, *Plant Molecular Biology*, **18** (4): 675-689.

D'Hont, A., Grivet, L., Feldmann, P., Rao, S., Berding, N., and Glaszmann, J. C. 1996, Characterisation of the double genome structure of modern sugarcane cultivars (*Saccharum* spp) by molecular cytogenetics, *Molecular and General Genetics*, **250**: 405-413.

D'Hont, A., Ison, D., Alix, K., Roux, C., and Glaszmann, J.C. 1998, Determination of basic chromosome numbers in the genus *Saccharum* by physical mapping of ribosomal RNA genes. *Genome,* **41**: 221-225.

Dalessandro, G. and Northcote, D. H. 1977, Possible control sites of polysaccharide synthesis during cell growth and wall expansion of pea seedlings (*Pisum sativum* L.), *Planta*, **134**: 39-44.

Dellaporta, S. L., Wood, J., and Hicks, J. B. 1983, A plant DNA minipreparation, *Plant Molecular Biology Reporter*, **1**: 19-21.

Deutsch, S., Iseli, C., Bucher, P., Antonarakis, S. E., and Scott, H. S. 2001, A cSNP map and database for human chromosome 21, *Genome Research*, **11**: 300-307.

Dufour, P., Grivet, L., D'Hont, A., Deu, M., Trouche, G., Glaszmann, J. C., and Hamon, P. 1997, Construction of a composite sorghum genome map and comparison with sugarcane, a related complex polyploid, *Theoretical and Applied Genetics*, **94**: 409-418.

Franzen, B., Carrubba, C., Feingold, D. S., Ashcom, J., and Franzen, J. S. 1981, Amino acid sequence of the tryptic peptide containing the catalytic-site thiol group of bovine liver uridine diphosphate glucose dehydrogenase, *Biochemical Journal*, **199** (3): 599-602.

Garg, K., Green, P., and Nickerson, D. A. 1999, Identification of candidate coding region single nucleotide polymorphisms in 165 human genes using assembled expressed sequence tags, *Genome Research*, **9**: 1087-1092.

Gerardy-Schahn, R., Oelmann, S., and Bakker, H. 2001, Nucleotide sugar transporters: biological and functional aspects, *Biochimie*, **83**: 775-782.

Gibeaut, D. M. 2000, Nucleotide sugars and glycosyltransferases for synthesis of cell wall matrix polysaccharides, *Plant Physiology and Biochemistry*, **38**: 69-80.

Grivet, L. and Arruda, P. 2001, Sugarcane genomics: depicting the complex genome of an important tropical crop, *Current Opinion in Plant Biology*, **5** (2): 122-127.

Grivet, L., Glaszmann, J. C., Vincentz, M., da Silva, F., and Arruda, P. 2003, ESTs as a source for sequence polymorphism discovery in sugarcane: example of the Adh genes, *Theoretical and Applied Genetics* , **106** (2): 190-197.

Hempel, J., Perozich, J., Romovacek, H., Hinich, A., Kuo, I., and Feingold, D. S. 1994, UDP-glucose dehydrogenase from bovine liver - primary structure and relationship to other dehydrogenases, *Protein Science*, **3**: 1074-1080.

Hinterberg, B., Klos, C., and Tenhaken, R. 2002, Recombinant UDP-glucose dehydrogenase from soybean, *Plant Physiology and Biochemistry*, **40** (12): 1011-1017.

Johansson, H., Sterky, F., Amini, B., Lundeberg, J., and Kleczkowski, L. A. 2002, Molecular cloning and characterization of a cDNA encoding poplar UDP-glucose dehydrogenase, a key gene of hemicellulose/pectin formation, *Biochimica et Biophysica Acta-Gene Structure and Expression*, **1576**: 53-58.

Loewus, F. A. and Murthy, P. P. N. 2000, *myo*-Inositol metabolism in plants, *Plant Science*, **150**: 1-19.

Luehrsen, K. R. and Walbot, V. 1991, Intron enhancement of gene expression and the splicing efficiency of introns in maize cells, *Molecular and General Genetics*, **225** (1): 81-93.

Maas, C., Laufs, J., Grant, S., Korfhage, C., and Werr, W. 1991, The combination of a novel stimulatory element in the first exon of the maize Shrunken-1 gene with the following intron 1 enhances reporter gene expression up to 1000-fold, *Plant Molecular Biology*, **16** (2): 199-207.

McElroy, D., Blowers, A. D., Jenes, B., and Wu, R. 1991, Construction of expression vectors based on the rice actin 1 (Act1) 5' region for use in monocot transformation, *Molecular and General Genetics*, **231** (1): 150-160.

Nelsestuen, G. L. and Kirkwood, S. 1971, The mechanism of action of uridine diphosphoglucose dehydrogenase, *Journal of Biological Chemistry*, **246** (12): 3828-3834.

Quick, W. P. and Schaffer, A. A. 1997, Sucrose metabolism in sources and sinks, in *Photoassimilate distribution in plants and crops: source-sink relationships*, E. Zamski and A. A. Schaffer, eds., Marcel-Dekker Inc., New York, pp. 115-156.

Reiter, W.-D. and Vanzin, G. F. 2001, Molecular genetics of nucleotide sugar interconversion pathways in plants, *Plant Molecular Biology*, **47**: 95-113.

Rethmeier, N., Kramer, E., Van Montagu, M., and Cornelissen, M. 1998, Identification of cat sequences required for intron-dependent gene expression in maize cells, *The Plant Journal*, **13** (6): 831-835.

Rethmeier, N., Seurinck, J., VanMontagu, M., and Cornelissen, M. 1997, Intron-mediated enhancement of transgene expression in maize is a nuclear, gene-dependent process, *The Plant Journal* **12** (4): 895-899.

Robertson, D., Smith, C., and Bolwell, G. P. 1996, Inducible UDP-glucose dehydrogenase from French bean (*Phaseolus vulgaris* L.) locates to vascular tissue and has alcohol dehydrogenase activity, *Biochemical Journal*, **313**: 311-317.

Sambrook, J., Fritsch, E. F., and Maniatis, T. 1989, *Molecular Cloning - A Laboratory Manual*, 2 edn, Cold Spring Habour Laboratory Press, New York.

Seitz, B., Klos, C., Wurm, M., and Tenhaken, R. 2000, Matrix polysaccharide precursors in Arabidopsis cell walls are synthesized by alternate pathways with organ-specific expression patterns, *The Plant Journal*, **21** (6): 537-546.

Stewart, D. C. and Copeland, L. 1998, Uridine 5 '-diphosphate-glucose dehydrogenase from soybean nodules, *Plant Physiology*, **116** (1): 349-355.

Stewart, D. C. and Copeland, L. 1999, Kinetic properties of UDP-glucose dehydrogenase from soybean nodules, *Plant Science*, **147** (2): 119-125.

Tenhaken, R. and Thulke, O. 1996, Cloning of an enzyme that synthesizes a key nucleotide-sugar precursor of hemicellulose biosynthesis from soybean: UDP-glucose dehydrogenase, *Plant Physiology*, **112** (3): 1127-1134.

Turner, W. and Botha, F. C. 2002, Purification and kinetic properties of UDP-glucose dehydrogenase from sugarcane, *Archives of Biochemistry and Biophysics*, **407** (2): 209-216.

CHAPTER 4

TISSUE SPECIFIC EXPRESSION OF UDP-GLUCOSE DEHYDROGENASE IN SUGARCANE

4.1 ABSTRACT

Most hemicelluloses and pectins that are incorporated into the cell wall are derived from a common precursor, namely UDP-glucuronic acid. UDP-glucuronic acid can be produced by the enzyme UDP-glucose dehydrogenase (EC 1.1.1.22). In this study the contribution of UDP-glucose dehydrogenase to pentan synthesis, as well as the expression pattern and subcellular localisation of the enzyme in mature sugarcane plants was studied at the tissue and cellular level. Radiolabelling with positionally labelled glucose was used to investigate the relative contributions of glycolysis, the oxidative pentose phosphate pathway and pentan synthesis to glucose catabolism. Significantly (P=0.05) more radiolabel was released as CO^2 from [6-^{14}C]-glucose than [1-^{14}C]-glucose in internodes 3, 4 and 5. This demonstrates a significant contribution of UDP-glucose dehydrogenase to glucose oxidation in the younger internodes. In addition, there was significantly (P=0.05) more radiolabel in the cell wall (fiber) component when the tissue was labelled with [1-^{14}C]-glucose rather than [6-^{14}C]-glucose. This also demonstrates a selective decarboxylation of glucose in position 6 prior to incorporation into the cell wall and is consistent with a major role for UDP-glucose dehydrogenase in cell wall synthesis in the younger internodes. This implies that the production of UDP-glucuronic acid through the UDP-glucose dehydrogenase reaction is only important in the developing internodes. High levels of expression of both the UDP-glucose dehydrogenase transcript and protein were detected in the leafroll, roots and young internodes of the sugarcane culm. *In situ* hybridisation showed that the UDP-glucose dehydrogenase transcript is present in virtually all cell types present in the sugarcane internode, while immunolocalisation in internodal sections from different developmental stages showed that the abundance of the protein declined in all cell types as maturity increased. The abundance of UDP-glucose dehydrogenase in developing tissues confirms that this enzyme plays an important role in the provision of hemicellulose precursors in most developing tissues of the sugarcane plant.

4.2 INTRODUCTION

Plant cells are surrounded by a network of interwoven polysaccharides that make up the cell wall. This cell wall, involved in the control of cell growth, cell signalling and defence, is essential for every aspect of plant life and represents a major carbon sink. The carbohydrate components of the cell wall are derived from a common precursor, namely UDP-glucose (Carpita, 1996). UDP-glucose is probably present in all plant cells and can be produced by either sucrose synthase (EC 2.4.1.13) or UDP-glucose pyrophosphorylase (EC 2.7.7.9) (Quick and Schaffer, 1997). Nucleotide sugar interconversion pathways represent a series of enzymatic reactions by which plants synthesise activated monosaccharides for incorporation into plant cell wall material (Gibeaut, 2000). In general, sugars are activated by conversion into nucleotide sugars, which act as substrates for the generation of other monosaccharides (Bolwell, 2000; Reiter and Vanzin, 2001).

The first step in this series of nucleotide sugar interconversion reactions takes place in the cytosol and produces UDP-glucuronic acid. UDP-glucuronic acid can either be produced directly from UDP-glucose by UDP-glucose dehydrogenase (EC 1.1.1.22) (Nelsestuen and Kirkwood,1971; Dalessandro and Northcote, 1977b), or can be formed by an alternate pathway, which involves the conversion of *myo*-inositol into glucuronic acid in a reaction catalysed by inositol oxygenase (reviewed by Loewus and Murthy, 2000). Both of these pathways may exist in plants, their importance depending on the plant species and tissue (Seitz *et al*., 2000). UDP-glucuronic acid is imported to the Golgi (Gerardy-Schahn *et al*., 2001) where it serves as substrate for both glycosyltransferases and for nucleotide sugar interconversion enzymes which produce precursors for hemicellulose and pectin, including arabinans, arabinogalactans, glucuronoarabinoxylans, rhamnogalacturonans, xylans and xyloglucans (Carpita, 1996; Bolwell, 2000; Gibeaut, 2000).

Significant levels of UDP-glucose are present in the sugarcane culm (Whittaker and Botha 1997). The cycling of carbon between sucrose and hexoses (Batta and Singh, 1986), resulting from the simultaneous synthesis and degradation of sucrose (Whittaker and Botha, 1997, Vorster and Botha, 1999) probably ensures that a significant pool of UDP-glucose, not only the precursor for the synthesis of structural polysaccharides, but also a respiratory substrate and a substrate for

sucrose synthesis, is maintained. As sugarcane is cultivated for its sugar-rich stalks, most carbon partitioning research in this plant has focussed on the accumulation of sucrose and partitioning within the sugar pool. Relatively little attention has been paid to the allocation of carbon to structural components of the cell such as the cell wall.

Studies of enzymes involved in the nucleotide sugar interconversion pathway found that UDP-glucose dehydrogenase is often the least active enzyme of the pathway suggesting that this enzyme may be rate-limiting for the provision of precursors for the expanding cell wall (Amino *et al*., 1985; Dalessandro and Northcote, 1977b; Robertson, Beech and Bolwell, 1995; Robertson, McCormack and Bolwell, 1995). Also, control of the activity of UDP-glucose dehydrogenase by feedback inhibition by UDP-xylose, the decarboxylation product of UDP-glucuronic acid, supports the idea that the reaction catalysed by this enzyme may represent a control point for the irreversible carbon flux into the pool of UDP sugars required for the biosynthesis of matrix polysaccharides (Dalessandro and Northcote 1977a; Dalessandro and Northcote1977c; Hinterberg *et al*., 2002; Stewart and Copeland, 1999; Turner and Botha, 2002).

A strong correlation between the expression of UDP-glucose dehydrogenase and a demand for structural polysaccharides in tissues that are actively synthesising cell walls, has been reported for several plant species (Dalessandro and Northcote, 1977a; Amino *et al*., 1985; Witt, 1992; Robertson, Beech, and Bolwell, 1995; Tenhaken and Thulke, 1996; Stewart and Copeland, 1998; Seitz *et al*., 2000; Johansson *et al*., 2002). In soybean seedlings the highest level of UDP-glucose dehydrogenase gene expression was detected in actively growing tissues such as root tips and lateral roots, with moderate expression in the epicotyl and in expanding leaves (Tenhaken and Thulke, 1996). The activity of the enzyme was also shown to be maximal during initial stages of nodule growth and development in soybean nodules (Stewart and Copeland, 1998). Promoter-reporter gene fusions, as well as protein blots, activity assays and histochemical activity staining in *Arabidopsis* showed that UDP-glucose dehydrogenase expression is regulated at a transcriptional level (Seitz *et al*., 2000). Also, many but not all growing tissues showed high activity levels of the enzyme. Hypocotyledons and cotyledons of young seedlings, for instance, did not show significant UDP-glucose dehydrogenase activity. In these tissues UDP-glucuronic acid is synthesised through the alternative inositol oxidation pathway. As observed for soybean,

UDP-glucose dehydrogenase activity in *Arabidopsis* was especially high in roots (Seitz *et al*., 2000). In poplar UDP-glucose dehydrogenase was expressed predominantly in differentiating xylem and young leaves with very low levels detected in phloem tissues (Dalessandro and Northcote, 1977a; Johansson *et al*., 2002).

UDP-glucose dehydrogenase has previously been purified from rapidly expanding culm tissues of sugarcane (Turner and Botha, 2002). Although the kinetic properties of the sugarcane enzyme were studied, no information is currently available about distribution of the enzyme in the plant. Here we report that UDP-glucose dehydrogenase plays an important role in pentan synthesis in younger compared to the more mature internodes. In addition, the data clearly show that the enzyme is expressed in all cell types present in the sugarcane internode and that expression is highest in actively growing tissues.

4.3 MATERIALS AND METHODS

4.3.1 Plant material

Mature, non-flowering, field-grown (in Stellenbosch, South Africa) sugarcane of cultivar N19, was used the isolation of nucleic acids and proteins and for the preparation of tissue sections. RNA and proteins were isolated from the same tissue samples. Tissues were sampled from roots, leaves and internodes. The internode above the natural break point of the sugarcane stalk was defined as internode 3. Internodes below this point were numbered sequentially through to 17 as maturity increased. Young and mature leaves were defined as actively growing, and actively photosynthesising, respectively.

4.3.2 ^{14}C Labelling studies

For analysis, selected internodes were excised and longitudinal cores were sectioned mid-way between the core and periphery of the internode, using a cork borer 6 mm in diameter. One mm slices were sectioned, using a hand microtome and immediately placed in 50 ml buffer containing 25 mM K-MES (pH 5.7), 250 mM mannitol and 1 mM $CaCl_2$ (Lingle, 1989) and washed for 15 min. Excess buffer was blotted from tissue discs, and the discs transferred to 1.5 ml buffer

containing 25 mM K-MES (pH 5.7), 250 mM mannitol, 5 mM glucose and 5mM fructose in 250 ml Erlenmeyer flasks. In all the labelling experiments the specific activity of glucose was 23 Bq $nmol^{-1}$ (Vorster and Botha, 1999). Discs were vacuum-infiltrated for 5 seconds and the flasks sealed with rubber stoppers. Samples were incubated for 3 h on a rotary shaker at 115 rpm. $^{14}CO_2$ released over the incubation period was collected in 500 µl 12% (w/v) KOH contained in a central well. After incubation, discs were transferred to 15 ml ice-cold 1% (m/v) $CaCl_2$. Unincorporated sugars were removed in three consecutive 2 min washes. The first two washes removed 98% of unincorporated label. Discs were transferred to 20 ml 80% (v/v) EtOH in sealed 50 ml centrifuge tubes, and incubated in an 80ºC water-bath overnight. The extracts were centrifuged at 12 000 g in a Sorvall SLA-600TC rotor for 15 min at 25°C. The EtOH-soluble supernatant was removed and discarded.

The insoluble component was fractionated as described previously (Dickson, 1979). Pellets were homogenised in a mortar and pestle at room temperature, transferred to 2 ml Eppendorf tubes and dried down in a vacuum centrifuge to remove any remaining EtOH. Proteins and starch were removed as described previously (Bindon and Botha, 2002). The final insoluble material, representing the cell wall components (fibre), was added to 1 ml Soluene®-350 (Packard) and vortexed. Samples were left at room temperature for 24 h, after which the ^{14}C cpm were determined.

4.3.3 Production of anti-body

UDP-glucose dehydrogenase was purified to homogeneity as previously described (Turner and Botha, 2002). Rabbits were immunised through injection of eight fractions containing 100 µg of purified protein, according to the method described by Bellstedt *et al.* (1987). Polyclonal antiserum was obtained from whole blood, collected after 38 days.

4.3.4 Enzyme extraction

The rind was removed from excised sugarcane internodes and the tissue ground to a fine powder in liquid nitrogen. Protein was extracted at 4°C using approximately 2 ml extraction buffer per gram tissue. The extraction buffer contained 50 mM Tris-Cl (pH 8.2), 10% (v/v) ethanediol, 8 mM DTT, 1 mM EDTA, 50 mM KCl and 2% (w/v) polyvinylpolypyrrolidone (PVPP). The

extract was filtered through muslin cloth and centrifuged (20 000 *g*, 15 min) to remove coarse material and then desalted on a prepacked PD-10 desalting column equilibrated with the same buffer.

UDP-glucose dehydrogenase was assayed routinely by monitoring NAD^+ reduction at 340 nm in a BioTech (CA, USA) microtitre plate reader maintained at 25ºC. UDP-glucose dehydrogenase activity was linear with respect to enzyme concentration. Reaction mixtures contained Tris-Cl (50 mM, pH 8.4), NAD^+ (4 mM), KCl (50 mM), ethanediol (10 % v/v), and UDP-glucose (5 mM) in a total volume of 250 µl. Reactions were started by the addition of NAD^+. One unit of enzyme activity was defined as the amount of enzyme resulting in the production of 1 µmol UDP-glucose per min at 25ºC.

4.3.5 Immuno-inactivation of UDP-glucose dehydrogenase activity

The effect of anti-UDP-glucose dehydrogenase IgG on the sugarcane enzyme activity was determined in a 150 µl reaction mixture containing crude protein extract (0.015 units of enzyme) and 2.5 to 50 µl of antiserum. After incubation for 45 min at 22ºC, 10 µl of insoluble Protein A was added, and the extract incubated for a further 30 min (Cawood *et al*., 1988). The sample was then centrifuged and residual enzyme activity measured as described above. Percentage inactivation by the antibody was expressed as a percentage of the UDP-glucose dehydrogenase activity in extracts, which were similarly incubated, but treated with pre-immune serum and insoluble Protein A.

4.3.6 Protein extraction and protein blot analysis

Total soluble protein was extracted from sugarcane leafroll, young leaves, mature leaves, internodes 3, 6, 9, 12, 15, 17, and roots. Tissues were ground to a fine powder in liquid nitrogen and soluble proteins extracted in approximately 1:4 (w/v) extraction buffer containing 100 mM Hepes (*N*-2-hydroxyethylpiperazine-*N'*-2-ethanesulfonic acid, pH 7), 1 mM EDTA, 10 mM DTT, 0.5 mM Pefabloc SC (4-(2-aminoethyl)-benzenesulfonyl fluoride, hydrogenchloride), 10% (v/v) glycerol and 2% (w/v) insoluble PVPP. Samples were briefly vortexed and then centrifuged to remove insoluble cell components (16 000 *g*, 10 min). Protein concentrations were determined

spectrophotometrically using the technique described by Bradford (1976), adjusted for microtiter plate application, using immunoglobulun G (IgG) as a standard.

Proteins representing sugarcane leafroll, young leaves, mature leaves, internodes 3, 6, 9, 12, 15, 17, and roots were separated using SDS-PAGE electrophoresis. Twenty micrograms of each sample were resolved in a discontinuous 12 % (v/v) gel, followed by a 4 % (v/v) stacking gel (Laemmli, 1970). A premixed protein molecular weight marker (Roche) was used as a standard. The gel was then equilibrated in transfer buffer (48 mM Tris, 39 mM (v/v) glycine, 20 % (v/v) methanol and 0.0375 % (v/v) SDS) for 20 min at 4^{o}C. The protein was transferred onto a Hybond-C (Amersham) membrane using a Transfer blot, Semi-dry transfer cell (BioRad) at 10-15 V for an hour. The membrane was stained with Ponceau-S solution (0.2 % (w/v) in 3 % (v/v) trichloroacetic acid), and destained with TBST buffer (20 mM Tris (pH 7.6), 137 mM (w/v) NaCl and 0.1 % (v/v) Tween-20). Blocking was performed overnight with 4% (w/v) BSA in TBST buffer.

The primary antibody against sugarcane UDP-glucose dehydrogenase was diluted 1: 2000, inoculated into the blocking buffer and incubated for 6 h at room temperature. The membrane was then washed three times for 15 min with TBST buffer. The secondary antibody, Anti-rabbit IgG (whole molecule) conjugated to alkaline phosphatase (Roche), was diluted 1: 2000 in 3 % (w/v) fat free milk in TBST and added to the membrane for an hour at room temperature. Thereafter the membrane was washed thoroughly, once in TBST and twice for 5 min each in TBST containing 10 % (w/v) SDS. The membrane was washed in TBST again and the signal was developed using a detection buffer (1 Nitroblue tetrazolium/5-Bromo-4-chloro-indolylphosphate (NBT/BCIP) tablet per 10 ml Milli-Q water).

4.3.7 RNA extraction

Total RNA was isolated from sugarcane leafroll, young leaves, mature leaves, internodes 3, 6, 9, 12, 15, 17, and roots. RNA was extracted from 5g of each tissue sample according to a method modified from Bugos *et al*. (1995). Tissues were ground to a fine powder in liquid nitrogen and added to 25:24:1 phenol:chloroform:isoamyl alcohol. After vortexing, an equal volume of homogenisation buffer (0.1 M Tris-HCl (pH 7.5), 1 mM EDTA, 0.1 M NaCl and 1% (w/v) SDS)

was added. Sodium acetate (pH 5.2) was then added to a final concentration 0.1 M. The emulsion was mixed, incubated on ice for 15 min and centrifuged at 4 °C (12 000 *g*, 15 min). The aqueous phase was subsequently transferred to a new tube and RNA precipitated through the addition of one volume of isopropanol followed by incubation at – 70 °C for at least 30 min. Precipitated RNA was recovered by centrifugation at 4 °C (10 000 *g*, 10 min). Excess salts were removed from the pellet by washing with 70% (v/v) ethanol. The pellet was air-dried and resuspended in 750 µl of diethyl pyrocarbonate (DEPC) treated water. Insoluble particles were removed by centrifugation (10 000 *g*, 5 min). The supernatant was transferred to a microcentrifuge tube, and RNA precipitated again using LiCl at a final concentration of 2 M. To maximise RNA precipitation, samples were incubated overnight at 4 °C. RNA was pelleted by centrifugation at 4 °C (12 000 *g*, 15 min). The pellet was again washed with 70 % (v/v) ethanol, and resuspended in DEPC treated water. Remaining insolubles were removed by centrifugation (10 000 *g*, 5 min). RNA concentration was determined spectrophotometrically.

4.3.8 Northern blot analysis

Six micrograms of total RNA from each tissue sample were separated in a 1.2 % (w/v) agarose gel. The samples were denatured through the addition of 50 % (v/v) formaldehyde and 1x MOPS buffer (200mM 3-[N-morpholino] propanesulphonic acid; 50 mM NaOAC; 5 mM EDTA) and subsequent incubation at 65 °C for 10 min before loading. The RNA was then transferred overnight to a positively charged Nylon membrane (Boehringer Mannheim) by upward capillary blotting in 10x SSC. RNA was further linked to the membrane by UV cross-linking at 120 mJ cm^{-2} for 2.5 min.

Primers UGD Fw4 and UGD Rev3 were used to amplify a fragment of 944 bp of the 5' end of the gene. This PCR amplified probe was labelled using the Prime-It II random primer labelling kit (Stratagene) and [α-^{32}P] dCTP (Amersham). Pre-hybridisation and hybridisation were performed in ULTRAhyb^{TM} buffer (Ambion) at 42 °C, according to the manufacturer's instructions. Following hybridisation, the membrane was washed twice in 2x SSC, 0.1% (w/v) SDS for 5 min at 42 °C, then twice in 0.1x SSC, 0.1% (w/v) SDS for 15 min at 42 °C, and finally twice in 0.1x SSC, 0.1% (w/v) SDS for 15 min at 65 °C. The washed membranes were exposed

to a Multi Purpose Phosphor Screen for 16 hours and visualised using a phospho-imager and analysis system (Packard Cyclone; Packard Instrument Company Inc, USA).

4.3.9 In situ hybridisation

Cellular localisation of the UDP-glucose dehydrogenase transcript was investigated using *in situ* hybridisation with DIG-labelled sense and anti-sense RNA riboprobes. Briefly, hand sectioned tissue slices of internode 7 were prepared as described above. Internodal tissue sections were fixed in 4 % (w/v) paraformaldehyde, pre-treated to reduce non-specific binding and dehydrated through a graded ethanol series. Pre-treatment included exposure of the tissue sections to 0.2 M HCl for 15 min; 0.125 $mg.ml^{-1}$ pronase (Sigma-Aldrich Chemie, GmbH Steinheim, Germany) in 50 mM Tris-HCl (pH 7.5), 0.5 mM EDTA for 10 min; 0.2 % (v/v) glycine for 2 min and 1 % (v/v) acetic anhydride in triethanolamine (pH 8) for 10 min before a final dehydration through a graded ethanol series.

A 561 bp Bam H1-Pst 1 restriction fragment was excised from the UDP-glucose dehydrogenase cDNA, and used to generate single stranded DIG-labelled sense and antisense probes by in vitro transcription of linearised template DNAs in the presence of DIG-labelled dUTP as described by the manufacturer (Boehringer Mannheim, GmbH Mannheim, Germany). These probes were diluted in hybridisation buffer (Sigma-Aldrich Chemie, GmbH Steinheim, Germany) to a final concentration of 200 $ng.ml^{-1}$. Following overnight hybridisation at 37 °C and washing at 50 °C in 2 x SSC in 50 % formamide, sections were treated with 1 % blocking agent (Roche Molecular Biochemicals, GmbH Mannheim, Germany) in 100 mM Tris-HCl (pH 7.5) in 150 mM NaCl for 1 hour before incubation for an additional hour with the antibody (Anti-DIG Fab fragments, Roche Molecular Biochemicals, GmbH Mannheim, Germany) diluted in the blocking solution (1:3000). Tissue sections were finally subjected to a detection buffer containing one tablet of NBT/BCIP tablet (Roche) in 10 % (w/v) polyvinyl alcohol (MW=70 000 –100 000) for 30-90 min. Sections were studied using a Nikon Eclipse E400 Microscope and photographed with a Nikon Coolpix 990 digital camera.

4.3.10 Immunohistochemistry

Internodes 3, 7, 10 and 13, were sampled as representatives of young, immature, mature and old tissues respectively. The selected internodes were excised and longitudinal cores were sectioned mid-way between the core and periphery of the internode, using a cork borer 6 mm in diameter. These cylindrical tissue sections were bisected lengthways and incubated overnight at 4 °C in a fixative (4% (w/v) paraformaldehyde in phosphate buffered saline (PBS; 130 mM (w/v) NaCl, 7 mM (w/v) Na_2HPO_4, 3 mM (w/v) $NaH_2PO_4.2H_2O$), pH 7, with 0.1 % (v/v) Triton X 100 and 0.1 % (v/v) Tween 20). The following day transverse sections were sliced by hand at a thickness of approximately 0.5 mm, and rinsed in PBS buffer for at least 15 min. Tissue sections were blocked in 100 mM Tris (pH 7.5) containing 150 mM NaCl, 15 $mg.ml^{-1}$ gelatin, 10 $mg.ml^{-1}$ BSA and 0.1 % (w/v) sodium azide for 2 hours at 37 °C with gentle agitation (all further incubation steps were performed at this temperature).

The primary antibody against sugarcane UDP-glucose dehydrogenase was diluted 1: 2000 in the blocking buffer and incubated for one hour. Pre-immune serum was used as a negative control at a dilution of 1:3000. Tissue sections were then washed three times for 15 min with PBS buffer (pH 7.5) containing 0.5 µl Tween 20 per ml buffer. A commercial secondary antibody, Anti-rabbit IgG (whole molecule) conjugated to alkaline phosphatase (Roche), diluted 1: 2000 in blocking buffer, was added and sections were incubated for one hour. Sections were washed as described above and the bound antibodies detected after incubation in a staining buffer, comprising of one NBT/BCIP tablet (Roche) per 10 ml of 10% (w/v) polyvinyl alcohol (MW=70 000-100 000) solution (Sigma), for 30 min (or until significant colouration occurred). The reaction was stopped in tap water containing 0.1 M EDTA. Sections were studied with a Nikon Eclipse E400 microscope and photographed with a Nikon Coolpix 990 digital camera.

4.4 RESULTS

4.4.1 Carbon partitioning of [1-^{14}C] glucose and [6-^{14}C] glucose

To estimate the relative contributions of glycolysis, the oxidative pentose phosphate (OPP) pathway and pentan synthesis to glucose catabolism we used two approaches. Firstly, $^{14}CO_2$ production was measured from tissues specifically labelled with [1-^{14}C] glucose and [6-^{14}C] glucose. The ratio of CO_2 evolution from carbon 1 and carbon 6 of glucose was used as an indicator of the relative importance of the different pathways. Secondly, the partitioning of label to the fiber component of the cell wall was analysed. Evidently the C-6/C-1 ratio exceeded unity in the young internodes, which strongly suggested contribution of pentan synthesis to CO_2 production (Table 4.1). In the older internodes the C-6/C-1 ratio was below unity and this indicated an increased importance of the OPP pathway relative to glycolysis and pentan synthesis. A larger contribution of carbon 1 than carbon 6 was evident in the insoluble matter for the young internodes (Table 4.1). This data clearly demonstrate a larger contribution by pentan synthesis to cell wall synthesis in the younger internodes compared to the more mature internodes 7 and 9.

Table 4.1 Incorporation of ^{14}C in CO_2 production and the cell wall (fiber) component of internodal tissue slices from N19 supplied with [1-^{14}C]-glucose and [6-^{14}C]-glucose for 3 h. Each value represents the average of three labelling experiments.

Internode	CO2		C6/C1	Fiber	
	C1	C6		C1	C6
	kBq mg^{-1} protein			*kBq mg^{-1} protein*	
3	1.33 ± 0.15	3.20 ± 0.34	2.40	4.27 ± 0.50	3.30 ± 0.31
4	1.03 ± 0.12	2.37 ± 0.27	2.29	3.41 ± 0.39	2.40 ± 0.31
5	0.97 ± 0.09	1.56 ± 0.12	1.61	3.20 ± 0.37	2.40 ± 0.27
7	0.72 ± 0.10	0.61 ± 0.10	0.85	2.40 ± 0.30	2.12 ± 0.21
9	0.52 ± 0.08	0.41 ± 0.09	0.79	2.10 ± 0.31	1.97 ± 0.19

4.4.2 Immuno-inactivation of UDP-glucose dehydrogenase activity

UDP-glucose dehydrogenase was previously purified from sugarcane (Turner and Botha, 2002). This protein was used to produce an anti-body by immunising rabbits. Polyclonal antiserum was

obtained from whole blood, collected after 38 days. A serial dilution of anti-UDP-glucose dehydrogenase was used to immunoprecipitate UDP-glucose dehydrogenase from a crude protein extract. Following precipitation, residual enzyme activity was measured. The enzyme activity could be completely precipitated by the polyclonal antibody (Figure 4.1).

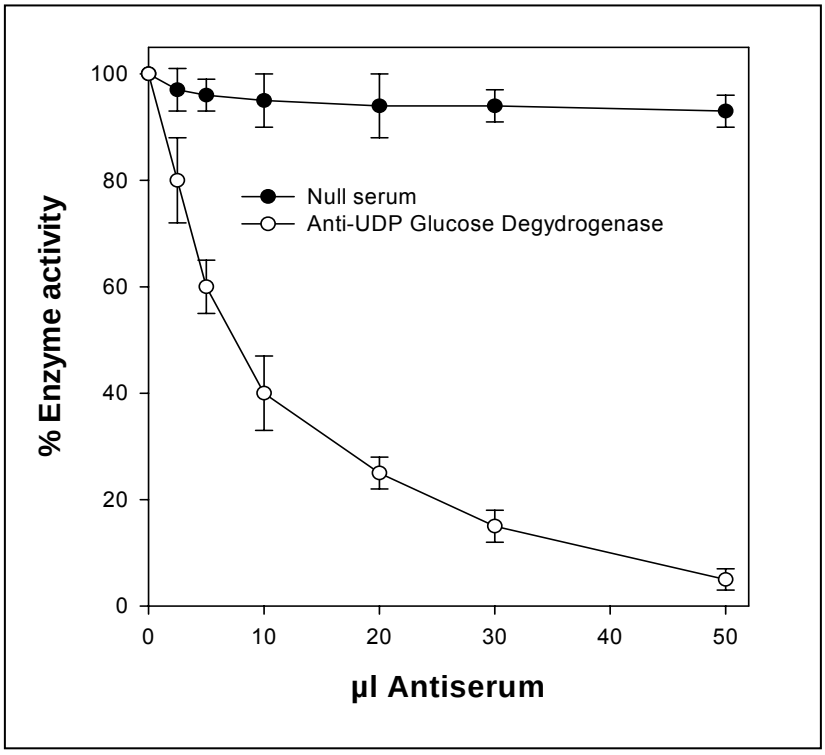


Figure 4.1 Immuno-removal of UDP glucose dehydrogenase. Percentage inactivation by the antibody is expressed as a percentage of the UDP-glucose dehydrogenase activity in extracts, which were similarly incubated, but treated with pre-immune serum and insoluble Protein A.

4.4.3 Expression analysis of sugarcane UDP-glucose dehydrogenase

The abundance of the UDP-glucose dehydrogenase transcript and protein was studied in different sugarcane tissues at various stages of development. Transcript levels detected in young leaves, mature leaves and internodes 3-18, showed that expression levels of UDP-glucose dehydrogenase correlated with growing and expanding tissues (Figure 4.2 B). The highest levels of transcript were detected in the leafroll and internode 3. Expression rapidly declined down the culm, with

almost no transcript detected in internode 18. Expression also declined in older leaves, but a signal was still visible. A relatively high level of expression was observed in root tissue. Protein levels followed a similar trend, with the highest levels of protein detected in young developing tissues (Figure 4.2 C). The observed molecular mass of the protein corresponds to that previously reported (Turner and Botha, 2002) for sugarcane UDP-glucose dehydrogenase (ca. 52 kDa).

4.4.4 Cellular localisation of UDP-glucose dehydrogenase

The cellular localisation of the UDP-glucose dehydrogenase transcript and protein was investigated by *in situ* hybridisation and immunolocalisation, respectively. The presence of both the transcript and the protein was visualised on tissue sections as dark blue staining. *In situ* hybridisation was used to investigate the distribution of the UDP-glucose dehydrogenase transcript in tissue sections prepared from internode 7 (Figure 4.3). Results showed high transcript levels in the vascular tissues of the stem, specifically in the companion cells of the phloem, in the xylem parenchyma and in cells of the vascular parenchyma. Strong labelling was also detected in the sucrose storing parenchyma cells, where blue staining was restricted to the cytoplasm.

Immunolocalisation was used to study abundance of the protein in internodal sections from different developmental stages (Figure 4.4). Distribution of the protein was similar to that of the transcript, i.e. the protein was detected in virtually all cell types found in internodal tissues, specifically the vascular tissues and sucrose storing parenchyma cells. Tissue sections prepared from different developmental stages showed that abundance of the protein rapidly declined down the stem (as maturity increased), with almost no protein detected in tissues from internode 13. In control experiments (sense RNA probe and pre-immune serum), only weak background colour was detected.

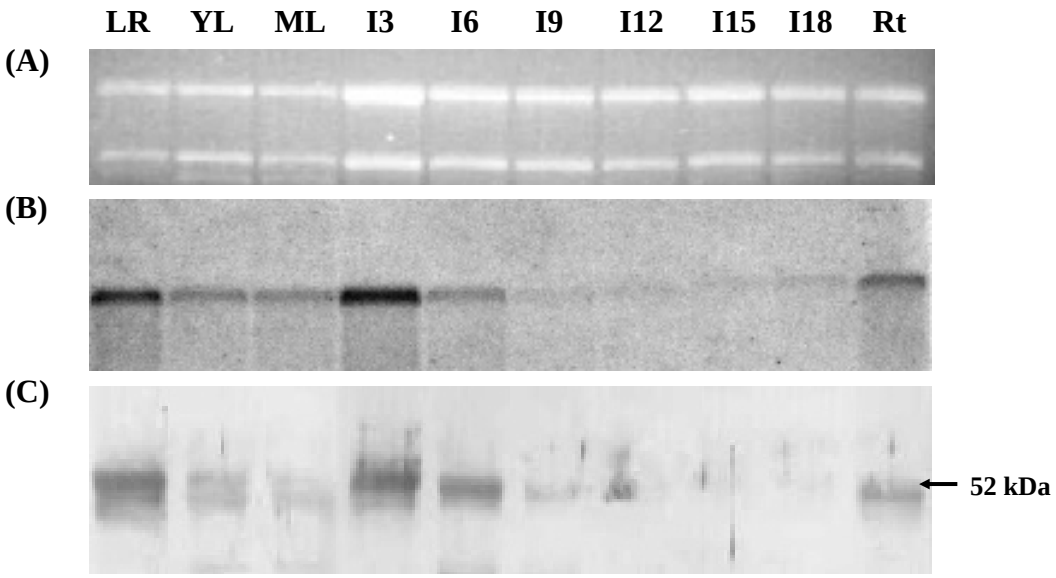


Figure 4.2 Expression of UDP-glucose dehydrogenase in sugarcane tissues. EtBr stain of extracted RNA used for membrane preparation (A). UDP-glucose dehydrogenase expression at a transcript (B) and protein (C) level, demonstrated by northern and western blot techniques, respectively. Each lane represents a different tissue: LR: leaf roll; YL: young leaf; ML: mature leaf; I3 - I18: internodal tissue; and Rt: root.

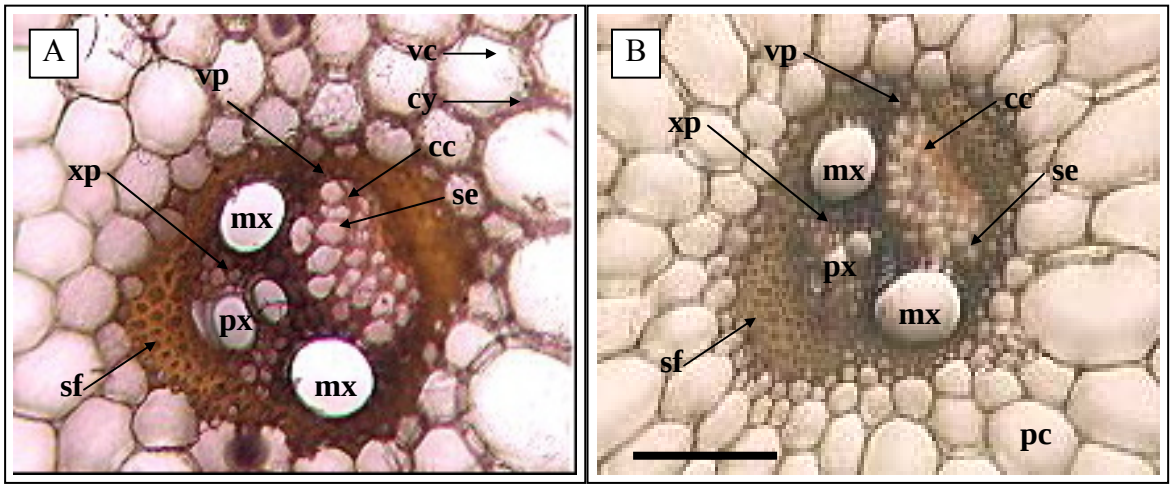


Figure 4.3 Investigation of cellular location of UDP-glucose dehydrogenase transcripts by *in situ* hybridisation on sections of sugarcane culm (internode 7). In a section hybridised to the antisense probe (A), transcripts were detected in the xylem parenchyma (xp), vascular parenchyma (vp), the phloem companion cells (cc), and the cytosol (cy) of the sucrose storing parenchyma cells (pc). A similar section hybridised to the sense control probe (B) showed no labelling (vc, vacuole; mx, metaxylem; px, protoxylem; se, sieve elements; sf, sclerenchymatous fibres; bar equals 150 µm)

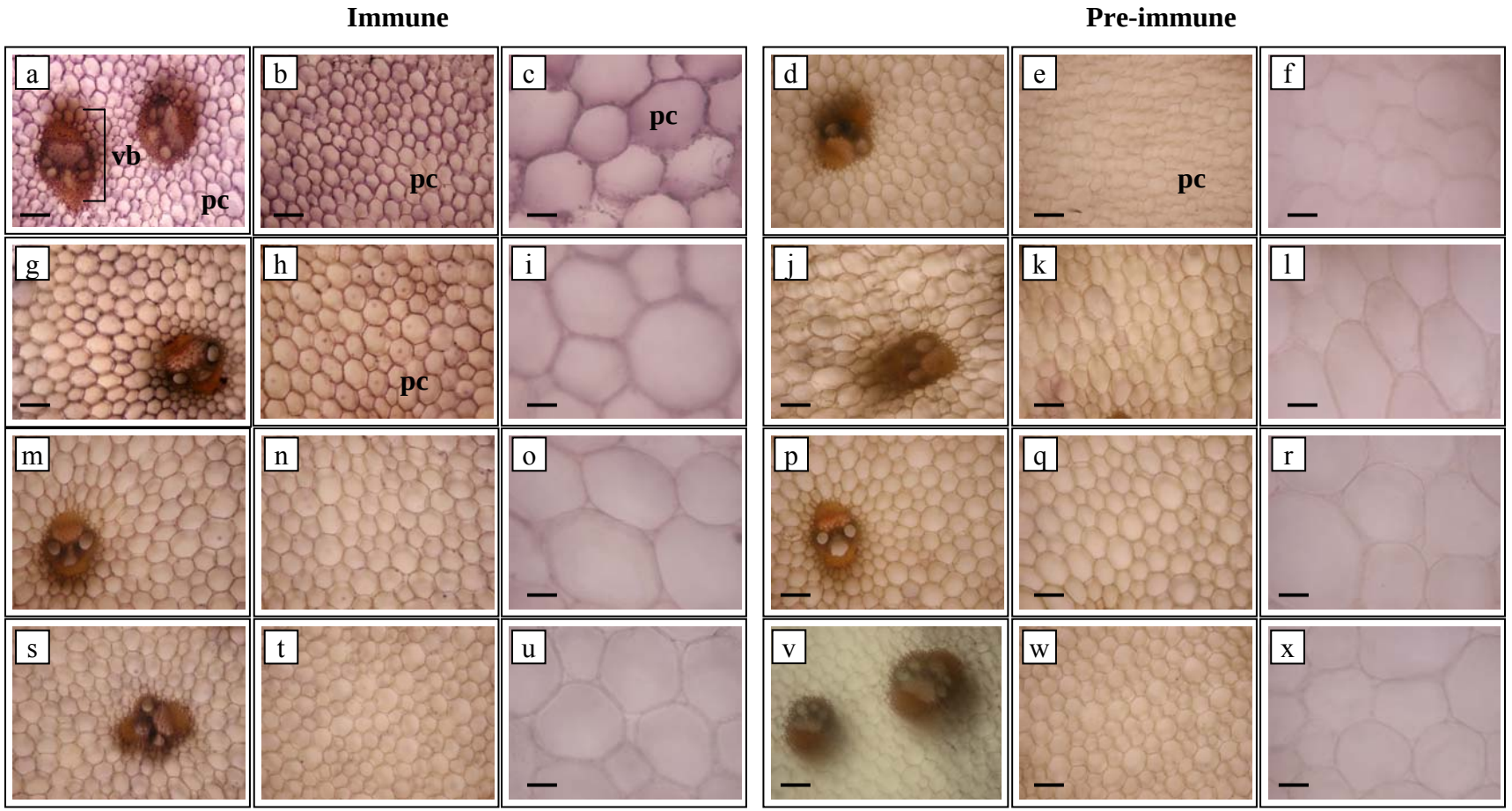


Figure 4.4 Distribution of UDP-glucose dehydrogenase, detected by antibody binding, on sections of sugarcane culm from internodes 3 (a-c: anti-UDP-glucose dehydrogenase; d-f: pre-immune serum), 7 (g-i: anti-UDP-glucose dehydrogenase; j-l: pre-immune serum); 10 (m-o: anti-UDP-glucose dehydrogenase; p-r: pre-immune serum) and 13 (s-u: anti-UDP-glucose dehydrogenase; v-x: pre-immune serum). In sections exposed to anti-UDP-glucose dehydrogenase, the protein was present in the vascular bundles (vb) and the sucrose storing parenchyma cells (pc). Bar equals 150 µm (a,b,d,e,g,h,j,k,m,n,p,q,s,t,v,w); 60 µm (c,f,i,l,o,r,u,x).

4.5 DISCUSSION

Hemicelluloses and pectins that are incorporated into the cell wall are derived from a common precursor, namely UDP-glucuronic acid. Two alternative pathways can supply UDP-glucuronic acid. UDP-glucose dehydrogenase utilises UDP-glucose to form UDP-glucuronic acid (Nelsestuen and Kirkwood, 1971; Turner and Botha, 2002). Alternatively, UDP-glucuronic acid can also be formed by the conversion of *myo*-inositol into glucuronic acid, which is subsequently conjugated to UDP in a reaction catalysed by inositol oxygenase (Loewus and Murthy, 2000).

In sugarcane the cycling of carbon between sucrose and hexoses results in a significant pool of UDP-glucose (Whittaker and Botha 1997). Radiolabelling was used to investigate the relative contributions of glycolysis, the oxidative pentose phosphate pathway and pentan synthesis to glucose catabolism. Selective decarboxylation of carbon 6 from glucose is an important source of released CO_2 in younger internodes and this leads to the C-6/C-1 ratios higher than unity. This most probably indicates that pentan synthesis is a more important component of metabolism in the young tissue. The net conversion of UDP-glucuronic acid to UDP-xylose results in the release of CO_2 from carbon number 6 (Davies *et al*., 1964). Usually, a C-6/C-1 ratio of unity is the predicted maximum value and is indicative of the sole contribution of glycolysis, since an increasing contribution of the oxidative pentose phosphate pathway to CO_2 release merely serves to decrease the ratio below unity (Davies *et al*.,1964). Ratios (C-6/C-1) of CO_2 release exceeding unity are indicative of increased pentan synthesis associated with growth (Hill and ap Rees, 1994).

UDP-glucose dehydrogenase was recently purified from rapidly expanding culm tissues of sugarcane (Turner and Botha, 2002). Although the kinetic properties of the sugarcane enzyme were studied, no information is available about the distribution of the enzyme in the plant. In this study the expression pattern and subcellular localisation of UDP-glucose dehydrogenase in mature sugarcane plants was studied at the tissue and cellular levels.

Earlier studies demonstrated that the demand for hemicellulose precursors was highest in actively dividing and rapidly expanding cells (Dalessandro and Northcote, 1977a; Amino *et al*., 1985;

Witt 1992; Robertson, Beech, and Bolwell, 1995; Tenhaken and Thulke, 1996; Stewart and Copeland, 1998; Seitz *et al*., 2000; Johansson *et al*., 2002). A similar correlation between the expression of UDP-glucose dehydrogenase and a demand for structural polysaccharides was observed for sugarcane in the present study. High levels of expression of both the UDP-glucose dehydrogenase transcript and protein were detected in the leafroll, roots and young internodes of the sugarcane culm. The expression of both the transcript and protein decreased along the length of the culm (with increasing maturity) to undetectable levels in older internodes. A similar trend was observed in *Arabidopsis* (older plants) (Seitz *et al*., 2000) and soybean (Tenhaken and Thulke, 1996; Stewart and Copeland, 1998) where UDP-glucose dehydrogenase was active in young growing tissues but not in the mature tissues. The absence of UDP-glucose dehydrogenase in mature tissues is thought to be correlated with the lack of demand for UDP-glucuronic acid derived sugars in differentiated cells (Tenhaken and Thulke, 1996).

The sugarcane culm is composed of storage parenchyma tissue permeated by numerous vascular bundles. *In situ* hybridisation showed that the UDP-glucose dehydrogenase transcript is present in virtually all cell types present in the sugarcane internode. As in other plant species (Stewart and Copeland, 1998) the sugarcane enzyme is only detected in the cytosol. Immunolocalisation in internodal sections from different developmental stages showed that abundance of the protein declined in all cell types as maturity increased, with almost no protein detected in internode 13. The abundance of UDP-glucose dehydrogenase in developing tissues confirms that this enzyme plays an important role in the provision of hemicellulose precursors in most developing tissues of the sugarcane plant.

Based on these results it is likely that UDP-glucose dehydrogenase dominates over the *myo*-inositol pathway in the provision of UDP-glucuronic acid as precursor for structural polysaccharides in most actively growing sugarcane tissues. To further investigate the importance of UDP-glucose dehydrogenase in the formation of cell walls in sugarcane, various transgenic plants in which UDP-glucose dehydrogenase is expressed in an anti-sense orientation, are currently being analysed. It is currently not understood why cell wall polysaccharides from grasses are so different from those of dicotyledonous plants (Carpita, 1996). Transgenic plants

with modified hemicellulose contents resulting from altered levels of UDP-glucose dehydrogenase, could provide some novel insights into this phenomenon.

4.6 ACKNOWLEDGEMENTS

I would like to thank Prof Frikkie Botha for technical assistance with carbon labelling work, production of the UDP-glucose dehydrogenase antibody and the immuno-inactivation of UDP-glucose dehydrogenase activity. The *in situ* hybridisation work discussed in this chapter was performed by Rakeshnie Ramoutar and formed part of her M.Sc. study.

4.7 REFERENCES

Amino, S., Takeuchi, Y., and Komamine, A. 1985, Changes in enzyme activities involved in formation and interconversion of UDP-sugars during the cell cycle in a synchronous culture of *Catharanthus roseus*, *Physiologia Plantarum*, **64:** 111-117.

Batta, S. K. and Singh, R. 1986, Sucrose metabolism in sugar cane grown under varying climatic conditions: synthesis and storage of sucrose in relation to the activities of sucrose synthase, sucrose phosphate synthase and invertase, *Phytochemistry*, **25**: 2431-2437.

Bellstedt, D. U., Human, P. A., Rowland, G. F., and Van der Merwe, K. J. 1987, Acid-treated, naked bacteria as immune carriers for protein antigens, *Journal of Immunological Methods*, **98** (2): 249-255.

Bindon, K. A. and Botha, F. C. 2002, Carbon allocation to the insoluble fraction, respiration and triose-phosphate cycling in the sugarcane culm, *Physiologia Plantarum*, **116**: 12-19.

Bolwell, G. P. 2000, Biosynthesis of plant cell wall polysaccharides, *Trends in Glycoscience and Glycotechnology*, **12** (65): 143-160.

Bradford, M. M. 1976, A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding, *Analytical Biochemistry*, **7** (72): 248-254.

Bugos, R. C., Chiang, V. L., Zhang, X. H., Campbell, E. R., Podila, G. K., and Campbell, W. H. 1995, RNA isolation from plant tissues recalcitrant to extraction in guanidine, *BioTechniques*, **19**: 734-737.

Carpita, N. C. 1996, Structure and biogenesis of the cell walls of grasses, *Annual Review of Plant Physiology and Plant Molecular Biology*, **47**: 445-476.

Cawood, M. E., Botha, F. C., and Small, J. G. C. 1988, Molecular properties of the ATP: D-fructose-6-phosphate 1-phosphotransferase isoenzyme from cucumis sativus, *Plant Cell Physiology*, **29** (1): 195-199.

Dalessandro, G. and Northcote, D. H. 1977a, Changes in enzymatic activities of nucleoside diphosphate sugar interconversions during differentiation of cambium to xylem in sycamore and poplar, *Biochemical Journal*, **162**: 267-279.

Dalessandro, G. and Northcote, D. H. 1977b, Possible control sites of polysaccharide synthesis during cell growth and wall expansion of pea seedlings (*Pisum sativum* L.), *Planta*, **134**: 39-44.

Davies, D. D., Giovanelli, J., and ap Rees, T. 1964, The carbohydrates, in *Plant Biochemistry*, W. James, ed., Blackwell Scientific, Oxford, pp. 85-155.

Dickson, R. E. 1979, Analytical procedures for the sequential extraction of ^{14}C-labeled constituents from leaves, bark and wood of cottonwood plants, *Physiologia Plantarum*, **45**: 480-488.

Gerardy-Schahn, R., Oelmann, S., and Bakker, H. 2001, Nucleotide sugar transporters: biological and functional aspects, *Biochimie*, **83**: 775-782.

Gibeaut, D. M. 2000, Nucleotide sugars and glycosyltransferases for synthesis of cell wall matrix polysaccharides, *Plant Physiology and Biochemistry*, **38:** 69-80.

Hill, S. A. and ap Rees, T. 1994, Fluxes of carbohydrate metabolism in ripening bananas, *Planta*, **192**: 52-60.

Hinterberg, B., Klos, C., and Tenhaken, R. 2002, Recombinant UDP-glucose dehydrogenase from soybean, *Plant Physiology and Biochemistry*, **40** (12): 1011-1017.

Johansson, H., Sterky, F., Amini, B., Lundeberg, J., and Kleczkowski, L. A. 2002, Molecular cloning and characterization of a cDNA encoding poplar UDP-glucose dehydrogenase, a key gene of hemicellulose/pectin formation, *Biochimica et Biophysica Acta-Gene Structure and Expression*, **1576** (1-2): 53-58.

Laemmli, U. K. 1970, Cleavage of structural proteins during the assembly of the head of bacteriophage T4, *Nature,* **227** (259): 680-685.

Lingle, S. E. 1989, Evidence for the uptake of sucrose intact into sugarcane internodes, *Plant Physiology*, **90**: 6-8.

Loewus, F. A. and Murthy, P. P. N. 2000, *myo*-Inositol metabolism in plants, *Plant Science*, **150**: 1-19.

Nelsestuen, G. L. and Kirkwood, S. 1971, The mechanism of action of uridine diphosphoglucose dehydrogenase, *Journal of Biological Chemistry*, **246** (12): 3828-3834.

Quick, W. P. and Schaffer, A. A. 1997, "Sucrose metabolism in sources and sinks," in *Photoassimilate distribution in plants and crops: source-sink relationships*, E. Zamski and A. A. Schaffer, eds., Marcel-Dekker Inc., New York, pp. 115-156.

Reiter, W.-D. and Vanzin, G. F. 2001, Molecular genetics of nucleotide sugar interconversion pathways in plants, *Plant Molecular Biology*, **47**: 95-113.

Robertson, D., Beech, I., and Bolwell, G. P. 1995, Regulation of the enzymes of UDP-sugar metabolism during differentiation of french bean, *Phytochemistry*, **39** (1): 21-28.

Robertson, D., McCormack, B. A., and Bolwell, G. P. 1995, Cell wall polysaccharide biosynthesis and related metabolism in elicitor-stressed cells of French bean (Phaseolus vulgaris L.), *Biochemical Journal*, **306** (3): 745-750.

Seitz, B., Klos, C., Wurm, M., and Tenhaken, R. 2000, Matrix polysaccharide precursors in Arabidopsis cell walls are synthesized by alternate pathways with organ-specific expression patterns, *Plant Journal*, **21** (6): 537-546.

Stewart, D. C. and Copeland, L. 1998, Uridine 5 '-diphosphate-glucose dehydrogenase from soybean nodules, *Plant Physiology*, **116** (1): 349-355.

Stewart, D. C. and Copeland, L. 1999, Kinetic properties of UDP-glucose dehydrogenase from soybean nodules, *Plant Science*, **147** (2): 119-125.

Tenhaken, R. and Thulke, O. 1996, Cloning of an enzyme that synthesizes a key nucleotide-sugar precursor of hemicellulose biosynthesis from soybean: UDP-glucose dehydrogenase, *Plant Physiology*, **112** (3): 1127-1134.

Turner, W. and Botha, F. C. 2002, Purification and kinetic properties of UDP-glucose dehydrogenase from sugarcane, *Archives of Biochemistry and Biophysics*, **407** (2): 209-216.

Vorster, D. J. and Botha, F. C. 1999, Sugarcane internodal invertases and tissue maturity, *Journal of Plant Physiology*, **155**: 470-476.

Whittaker, A. and Botha, F. C. 1997, Carbon partitioningduring sucrose accumulation in sugarcane internodal tissue, *Plant Physiology*, **115**: 1651-1659.

Witt, H.-J. 1992, UDP-glucose metabolism during differentiation and dedifferentiation of *Reilla helicophylla*, *Journal of Plant Physiology*, **140**: 276-281.

CHAPTER 5

ISOLATION AND EVALUATION OF A DEVELOPMENTALLY REGULATED SUGARCANE PROMOTER

5.1 ABSTRACT

The young internodes of sugarcane are ideal targets for altering metabolism, through genetic manipulation, to potentially control known fungal diseases such as Smut or to increase sucrose yields in these regions that are currently being discarded. At present, no regulatory sequences that specifically drive transgene expression in young developing sugarcane tissues are available. The objective of this study was therefore to isolate and evaluate such a sequence. The approach followed was firstly, to identify an endogenous gene expressed in the desired pattern, and then to isolate the corresponding promoter from the sugarcane genome. The promoter of a gene encoding UDP-glucose dehydrogenase was selected for isolation, based on what is known about the function of the enzyme it encodes. The promoter region (1700 bp), including an intron (973 bp) located in the 5'-untranslated region (UTR) of this gene, was isolated and subsequently fused to the GUS reporter gene for transient expression analysis and plant transformation. Transient expression analysis showed that the presence of the intron was essential for strong GUS expression. Analysis of stably transformed transgenic sugarcane plants, grown in a green house, indicates that the promoter is able to drive GUS expression in a tissue specific manner under these conditions.

5.2 INTRODUCTION

When sugarcane is harvested, the top internodes are traditionally discarded due to low juice purity. The top of the cane is also the point of infection for Smut, the most important fungal disease of sugarcane in South Africa. Genetic manipulation has the potential to alter metabolism in these tissues to increase the sucrose yield, or to control Smut, possibly through the insertion of

a single gene. Although the transformation of sugarcane is well established (Arencibia *et al*., 1995; Arencibia *et al*., 1998; Bower and Birch, 1992), a major obstacle limiting progress in this area is the availability of promoters.

Efficient genetic manipulation is dependant on the availability of promoter elements that allow the control of transgene expression in different tissues and at different developmental stages. The shortage of these regulatory elements, as well as patent considerations (Birch, 1997), has made it necessary to isolate specific promoters from sugarcane. Several promoters that direct near-constitutive expression in monocots have been isolated. These include promoters isolated from plants, such as the maize polyubiquitin (*ubi-1*) promoter (Christensen and Quail, 1996) and the rice actin (*Act1*) promoter (McElroy *et al*., 1990), viral promoters such as the cauliflower mosaic virus (CaMV) 35S promoter (Benfey *et al*., 1990; Terada and Shimamoto, 1990), sugarcane bacilliform badnavirus promoter (Tzafrir *et al*., 1998), and promoters isolated from the banana streak badnavirus (Schenk *et al*., 2001). Though constitutive expression of a transgene may sometimes be required, targeting the expression to a specific tissue where the action of the transgene is required will greatly decrease the metabolic load resulting from transformation. At present, no regulatory sequences are available that specifically drive transgene expression in developing sugarcane tissues. The aim of this study was to isolate and evaluate a developmentally regulated sugarcane promoter.

One possible approach to obtain promoters which direct specific levels and distribution of expression is to identify endogenous genes already expressed in the desired pattern in the organism targeted for transformation, in this instance, sugarcane. The corresponding promoter can then be isolated from the genome of the target organism. The promoter of a gene encoding UDP-glucose dehydrogenase was selected as a potential target for promoter isolation, based on what is known about the function of the enzyme that it encodes.

UDP-glucose dehydrogenase catalyses the oxidation of UDP-glucose to UDP-glucuronic acid (Nelsestuen and Kirkwood, 1971), a precursor for sugar nucleotides, which are incorporated into pectin and hemicelluloses. Both pectin and hemicellulose are key components of cell walls, providing a matrix that strengthens the cell wall structure (Gibeaut, 2000). As UDP-glucose

dehydrogenase is required for growth and development, the promoter of this gene could possibly be used to drive transgene expression in young developing tissues.

In this study the promoter and first intron, located in the 5'-untranslated region, of an UDP-glucose dehydrogenase gene was isolated from sugarcane. The isolated sequence was evaluated for its ability to drive transgene expression in a transient system and stably transformed sugarcane. Results indicate that the UDP-glucose dehydrogenase promoter can drive highly tissue specific expression in transgenic sugarcane. Also, the presence of the intron is essential for promoter activity.

5.3 MATERIALS AND METHODS

5.3.1 Isolation of UDP-glucose dehydrogenase promoter

As described in Chapter 3, a genomic clone containing the UDP-glucose dehydrogenase gene and upstream sequence was isolated after screening a sugarcane genomic library, prepared from sugarcane variety N19. From this genomic clone, a *Not* I-*Xba* I fragment of approximately 4600 bp, containing the complete coding region, as well as the 5'-UTR containing an intron of 973 bp, and a further 1600 bp of the promoter region was sub-cloned (Figure 5.1). An *Eco* RV restriction site was located in the 5'-UTR, 5 bp upstream of the left border of the intron. This site was used to further digest the cloned sequence to yield two fragments of approximately 1700 (named

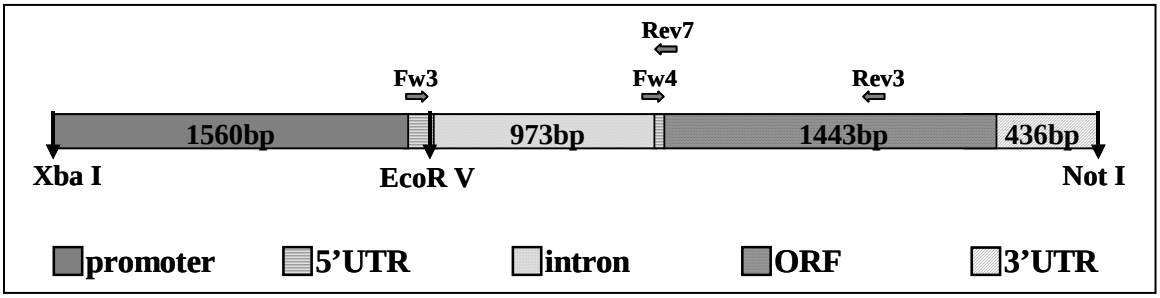


Figure 5.1 Graphic representation of a *Not* I *Xba* I fragment containing the sugarcane UDP-glucose dehydrogenase promoter and gene. Gene specific primers and restriction sites used to construct the promoter cassettes are also shown.

UGD10SLL, containing the promoter) and 2900 bp (named UGD10SS, containing the 5'-UTR intron and coding sequence of the gene), respectively.

5.3.2 Construction of UDP-glucose dehydrogenase promoter and chimeric GUS reporter gene constructs

Reporter constructs were made by placing the bacterial *uidA* gene, coding for β-glucuronidase (GUS), and the *Agrobacterium* nopaline synthase (*nos*) terminator (Jefferson *et al*., 1987), under the control of the sugarcane UDP-glucose dehydrogenase promoter, with or without the 5'-intron. Two promoter-reporter gene constructs were prepared: one containing the UDP-glucose dehydrogenase promoter, 5'-UTR and intron, namely pBGUS UGDip, and one that only contained the promoter and 5'-UTR, pBGUS UGDp. A *BamH* I-*Eco* RI fragment containing the GUS gene and adjacent *nos* terminator, was removed from the construct, pAHC27 (Christensen and Quail, 1996) and cloned into a pBluescript II SK + cloning vector (Stratagene). The resulting promoterless vector was named pBGUS 000.

To prepare pBGUS UGDip, the promoter sequence (approximately 1700 bp) was PCR amplified from cloned fragment UGD10SLL, using the vector-specific primers, T3 and T7. The intron, 5'-UTR and first 3 codons of the UDP-glucose dehydrogenase gene (approximately 1000 bp) were PCR amplified from the cloned fragment UGD10SS using one vector specific primer, T7, and one gene-specific primer that included the translation initiation codon, ATG, and contained a *BamH* I restriction site, UGDRev7 (5'-GCACGGATCCTTCAC**CAT**CTTGTCAGATAG, restriction site underlined, ATG in bold). The PCR reaction was performed in a volume of 50 µl using 2ng of plasmid DNA as template. The PCR mixture also contained: 1 x PCR buffer, 1.5 mM $MgCl_2$, 200 µM each of dATP, dCTP, dGTP and dTTP, 1 U of *Taq* DNA polymerase (all purchased from Promega) and 0.2 µM of each primer. PCR was performed under the following conditions: 94 °C for 45 s (1 cycle), 94 °C for 45 s and 55 °C for 45 s and 72 °C for 45 s (10 cycles), 94 °C for 30 s and 50 °C for 30 s and 72 °C for 30 s (25 cycles), 72 °C for 2 min (1 cycle). For each fragment, four identical PCR reactions were performed, subsequently pooled and excess nucleotides and primers were removed using a Qiagen PCR Purification system. Equimolar amounts of the two purified PCR products (1.5 µg of PCR product of UGD10SLL and 0.75 µg of PCR product of UGD10SS) were pooled and digested with *Eco* RV for 2 hours at 37

°C. The restriction enzyme was removed using a Qiagen quick spin column. The two fragments were then ligated through the addition of T4 DNA Ligase and a Rapid Ligation Buffer (Promega). The ligation reaction was performed at room temperature and allowed to proceed for 2 hours. The reaction mixture was again purified using a Qiagen quick spin column. The ligation product was then double digested using *Xba* I and *BamH* I and cloned into pBGUS 000 (prepared using the same restriction enzymes).

To prepare the intronless promoter construct, pBGUS UGDp, the promoter sequence (approximately 1700 bp) was again amplified from cloned fragment UGD10SLL, as described above. To remove the intron at the intron splice sites, thereby keeping the 5'-UTR intact, two gene specific primers were used in a PCR reaction to amplify a second fragment containing the first 955 bp of the open reading frame, using UGD10SS as template. Primer UGDRev3 (5'-CTCTTCTGGTAGTCGTTGATC; 902-923 bp downstream of the translation initiation codon, ATG) was used with UGDFw4 (5'-GCTCGATATCT**GG**TCACAGATCTATCTG, located between 12 – 22 bp upstream of the ATG). UGDFw4 is homologous to the 5'-UTR as it occurs in the cDNA, including the intron splice site (bold), and contains the *Eco* RV restriction site (underlined) located in the 5'-UTR, 5 bp upstream of the left border of the intron. The two amplified fragments were digested with *Eco* RV and ligated to each other as described above. The ligation product was then used as a template for a third PCR reaction. Vector primer T3 (used to amplify the PCR product of UGD10SLL), was used with the gene-specific primer UGDRev7, described above. This PCR product was then double digested using *Xba* I and *BamH* I and cloned into pBGUS 000 (prepared using the same restriction enzymes).

5.3.3 Particle bombardment of 5 day old maize coleoptiles for transient expression analysis

Transient expression assays were done to verify whether an active promoter had been isolated. Two sugarcane promoter constructs, pBGUS UGDip and pBGUS UGDp, and a control plasmid, pAHC27 (Christensen, Sharrock, and Quail, 1992), which contains the maize polyubiquitin (*ubi-1*) promoter and first intron, fused to the GUS reporter gene, were used in transient expression assays. Maize seeds were germinated in the dark on wet tissue paper. Five days after germination, coleoptiles were removed with a scalpel and placed in petri dishes containing wet filter paper. The constructs were delivered by microprojectile bombardment using DNA-coated

tungsten particles. Bombarded tissues were kept in the dark on wet filter paper and assayed for GUS activity after 24 hours.

5.3.4 Sugarcane tissue culture

Standard tissue culture protocols (Bower and Birch, 1992; Snyman *et al*., 1996) were used to produce and maintain embryogenic sugarcane callus from mature field grown plants of variety NCo310. Briefly, sugarcane callus was initiated in the dark from leafroll sections on callus initiation medium, MSC_3 (MS basal medium (Sigma) with 2 % (w/v) sucrose, 0.05% (w/v) casein hydrolysate, 0.6 % (w/v) agargel, and 3 mg l^{-1} 2.4-dichloro-phenoxyacetic acid (2.4-D) (Sigma)). Callus was transferred to fresh medium every 2 weeks.

5.3.5 Sugarcane transformation

Actively growing embryogenic calli (predominantly globular pro-embryoids) was transformed by particle bombardment, using a particle inflow gun (PIG) (Snyman *et al*., 1996) and DNA-coated tungsten particles (Grade M-10, Bio-Rad Laboratories). Four hours prior to bombardment, the callus was placed onto osmoticum medium (MSC_3 supplemented with 3.64 % (w/v) sorbitol and 3.64 % (w/v) mannitol). A $CaCl_2$/spermidine coprecipitation method (Birch and Franks, 1991) was used to bind DNA to tungsten particles. Calli were co-transformed with either pBGUS UGDip or pBGUS UGDp, and the pEmuGN plasmid (Last *et al*., 1991) containing the neomycin phosphotransferase gene, *nptII* (the ratio of the two plasmids was 1:1 and the total amount of DNA 10 μg). Four hours after bombardment, callus was transferred from the osmoticum medium to MSC_3 medium. Two days later callus was placed on a selection medium (MSC_3 containing 50 $mg.l^{-1}$ geneticin (G418)). After a selection period of 8 – 12 weeks in the dark, geneticin resistant callus was transferred onto antibiotic-containing regeneration medium (MS basal medium (Sigma) supplemented with 2 % (w/v) sucrose, 0.05% (w/v) casein hydrolysate, 0.6 % (w/v) agargel, with 50 mg l^{-1} geneticin) and incubated in the light. Geneticin resistant plantlets were then rooted and hardened off on regeneration medium without antibiotic, and subsequently grown under green house conditions.

5.3.6 Analysis of GUS activity

The same method was used for histochemical staining of maize seedling for transient assays (Section5.3.3), and for transformed sugarcane plants (Section 5.3.5). Transgenic plant tissues that were assayed include whole (tissue culture-grown) plants, sectioned leafroll tissue (transverse hand sections approximately 1 cm in length) and transverse stem sections (sliced by hand at a thickness of approximately 0.5 mm) from green house grown plants. Bombarded tissues, whole transgenic plants, or transgenic plant tissue sections were incubated for 24 hours at 37 °C (Jefferson *et al*., 1987) in substrate-containing GUS assay solution (100 mM sodium phosphate buffer (pH 7.0); 5 mM potassium ferrocyanide; 5 mM potassium ferricyanide; 2 mM X-Gluc (5-bromo-4-chloro-3-indolyl-β-glucuronic acid); 0.3 % (v/v) Triton X-100) which was introduced into the cells by a brief vacuum infiltration. The incubation was followed by dehydration in an ethanol series or clearing with chlorallactophenol (CLP, 2:1:1 chloral hydrate: lactic acid: phenol) as described by Beeckman and Engler (1994). Tissues were studied using a Leica Fluo III (whole plant and sectioned leafroll assays) or Nikon Eclipse E400 (stem sections) Microscope and photographed with a Nikon Coolpix 990 digital camera.

5.3.7 PCR amplification and Southern blot analysis of transgenic plants

Genomic DNA was isolated from young leaves of green house-grown transgenic sugarcane plants according to Dellaporta *et al*. (1983). The presence of the promoter and GUS reporter gene, as well as the *nptII* selectable marker gene was determined by PCR. For plants transformed with pBGUS UGDip and pBGUS UGDp, a promoter-specific forward primer, UGDFw3 (5'-ACGCATCGCGCCAAGGAAGA, located approximately 100-80 bp upstream of the left border of the intron), was used in combination with the reporter gene specific reverse primer, GUSp (5'-GCTTTCCCACCAACGCTGATC). Successful amplification from plants transformed with pBGUS UGDip should result in an amplified fragment of 1220 bp, while plants transformed with the intronless pBGUS UGDp should yield an amplification product of 250 bp. In addition, primers Npt II F (5'-ACCATGGTTGAACAAGATGGATTG) and Npt II R (5'-CTCAGAAGAACTCGTCAAGAAGG) were used to amplify an 799 bp fragment of the *nptII* gene from all transgenic plants.

For Southern blot analysis, 5 µg of genomic DNA was digested with *Xba* I, which has a single restriction site in both of the transformation vectors. The digested DNA was separated in a 0.8 % (w/v) agarose gel and transferred overnight to a positively charged nylon membrane (Roche) by downward capillary blotting in 10x SSC. DNA was cross-linked to the membrane through exposure to UV light for 2.5 min at 120 $mJcm^{-2}$. A *Sac* I-*Eco* RV restriction digest of pBGUS UGDip was used to isolate a fragment of 1074 bp of the GUS reporter gene. This fragment was labeled using the Prime-It II random primer labeling kit (Stratagene) and [α-^{32}P] dCTP (Amersham). Pre-hybridisation and hybridisation were performed in ULTRAhyb^{TM} buffer (Ambion) at 42 °C, according to the manufacturer's instructions. Following hybridisation, the membrane was washed twice in 2x SSC, 0.1% (w/v) SDS for 5 min at 42 °C, then twice in 0.1x SSC, 0.1% (w/v) SDS for 15 min at 42 °C, and finally twice in 0.1x SSC, 0.1% (w/v) SDS for 15 min at 65 °C. The washed membranes were exposed to a Multi Purpose Phosphor Screen for 16 hours and visualized using a phospho-imager and analysis system (Packard Cyclone; Packard Instrument Company Inc, USA).

5.4 RESULTS

5.4.1 Isolation and characterisation of the sugarcane UDP-glucose dehydrogenase promoter

After several rounds of screening a genomic library, a clone containing the UDP-glucose dehydrogenase gene and upstream sequence was isolated. A 2.7 kb fragment upstream of the translation initiation site was isolated. Sequence analysis of this region revealed the presence of an intron of 973 bp located within the 5'-UTR, 21 bp upstream of the translation initiation site, and preceded by an untranslated exon of approximately 122 bp (The exact location of the transcription initiation site was not determined experimentally, but putatively assigned based on comparison with other full-length cDNAs and homologous ESTs). A further 1700 bp of putative promoter sequence was retained for further analysis.

Examination of the nucleotide composition of the untranslated leader exon revealed two C/T-elements with the core sequence CTCCTCTTCTC (Figure 5.2), separated by 32 bp. A 15 bp

G/A-rich sequence (i.e. C/T in the non-coding strand of the DNA) is located between the C/T elements. A partial inverted repeat of the core sequence of the C/T sequence is also found in the promoter region (-12 - -21, relative to the putative transcription initiation site). Although the core sequence of the C/T-element does not correspond to a known transcription factor binding site, C/T rich sequences located in the 5'-UTR are believed to play an important role in enhancing the expression of some genes (Mun *et al*., 2002; Pearson and Meagher, 1990).

Computational analysis of the promoter region (Higo *et al*., 1999) revealed the presence of many putative transcription factor binding sites (TFBS). A TATA box-like motif (TATTTAA) is found between 43 and 37 bp upstream of the transcription initiation site (Figure 5.2). Although this sequence does not correspond exactly to the eukaryotic TATA box consensus sequence TATA(T/A)A(T/A), a TATA box of the same sequence has previously been shown to be functional in the rice *PAL* promoter (Zhu *et al*., 1995). An 5'-AAAG-3' motif is repeated six times within 400 bp of the transcription start site, i.e. the proximal promoter region. This motif has previously been described as the core binding site for a group of plant-specific transcription factors, Dof proteins (for DNA binding with one finger) (Yanagisawa, 1996; Yanagisawa and Schmidt, 1999). The arrangement of the motifs includes one tandem repeat and a further two binding sites seperated by 7 bp. The grouping of these motifs close together and their location close to the transcription initiation site in the absence of other obvious TFBS, suggests a possible role for Dof transcription factors in the regulation of the expression of UDP-glucose dehydrogenase.

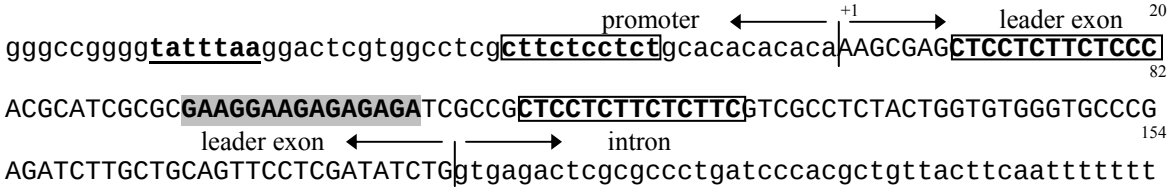


Figure 5.2 Nucleotide sequence around the transcription initiation site of the sugarcane UDP-glucose dehydrogenase gene. The putative transcription start site is designated as +1. A putative TATA box is underlined. Two C/T elements located in the leader exon and one in the promoter region are boxed. The G/A rich sequence found between the C/T elements is shaded.

5.4.2 Transient expression analysis

To directly assay promoter activity, two different expression cassettes containing the GUS reporter gene fused to the putative UDP-glucose dehydrogenase promoter with (pBGUS UGDip) and without (pBGUS UGDp) the intron, were prepared. These two constructs and a control plasmid, pAHC27 (Christensen *et al*., 1992), which contains the maize polyubiquitin (*ubi-1*) promoter and first intron fused to the GUS reporter gene, were delivered to 5 day old maize coleoptiles by microprojectile bombardment using DNA-coated tungsten particles. Transient gene expression, as indicated by blue spots, was observed for all the constructs (Figure 5.3). pAHC27 was used as a reference as *ubi-1* has previously been shown to be one of the most efficient promoters for driving GUS expression in sugarcane (Gallo-Meagher and Irvine, 1993). Similar (high) levels of transient GUS expression were observed resulting from bombardment with the intron-containing construct, pBGUS UGDip, and pAHC27. Bombardment with the intronless construct, pBGUS UGDp, resulted in very weak transient expression.

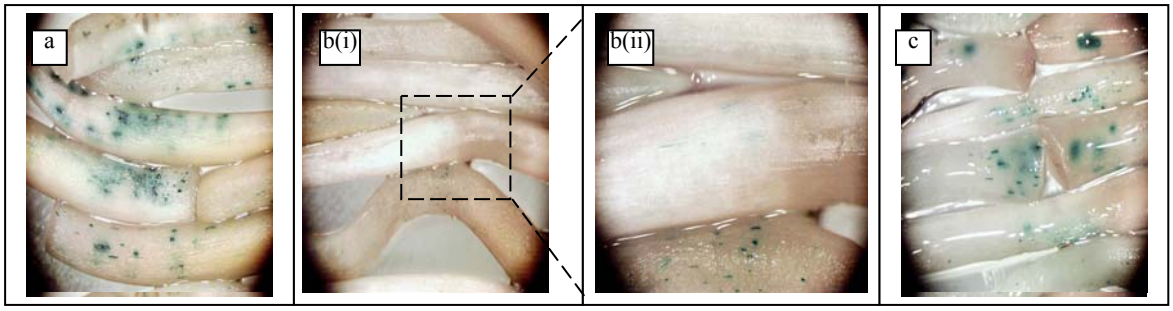


Figure 5.3 Transient expression analysis following particle bombardment of 5 day old maize coleoptiles. Two constructs, pBGUS-UGDip (a), and pBGUS-UGDp (b(i) and b(ii), were tested for their ability to drive transgene expression. pAHC27 (c), which contains the maize polyubiquitin (*ubi-1*) promoter and first intron, was included as a positive control.

5.4.3 Sugarcane transformation

To determine whether the sugarcane UDP-glucose dehydrogenase promoter could produce high, stable gene expression, transgenic sugarcane plantlets, expressing pBGUS UGDip or pBGUS UGDp, were regenerated from geneticin resistant callus. Transformation of embryogenic sugarcane calli with pBGUS UGDip and pBGUS UGDp resulted in 4 and 6 transgenic lines, respectively. PCR was carried out to confirm the transformation status of these plants. Results

showed that all the putative transgenic plants contained the selection gene, *nptII*, and that all but one of the plants contained the relevant promoter reporter-gene construct (Figure 5.4). One line transformed with pBGUS UGDip only contained the selection gene. This plant line, ip2.1, was included as a negative control in all histochemical assays. As a result, only three transgenic plant lines for pBGUS UGDip were available for further analysis.

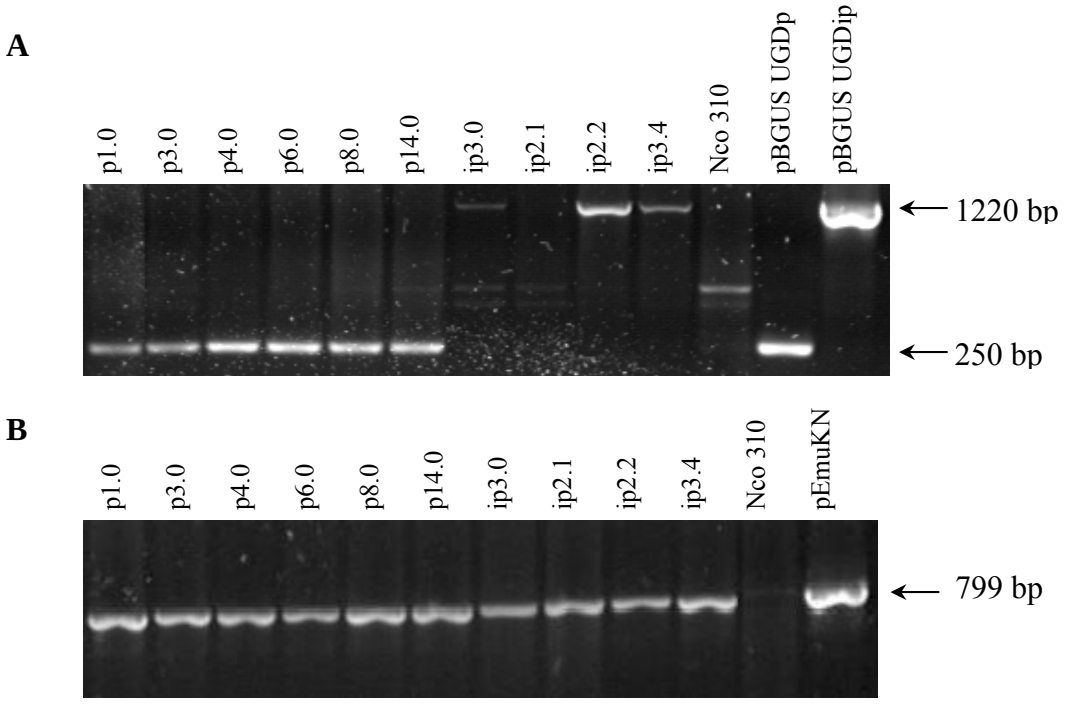


Figure 5.4 Confirmation of the presence of the promoter, GUS reporter gene and *nptII* selectable marker gene by PCR amplification from genomic DNA isolated from transgenic sugarcane plants. A: For plants transformed with pBGUS UGDip and pBGUS UGDp, a promoter specific forward primer, UGDFw3 was used in combination with a reporter gene specific reverse primer, GUSp. Successful amplification from plants transformed with the intronless pBGUS UGDp should give an amplification product of 250 bp, while plants transformed with pBGUS UGDip should result in an amplified fragment of 1220 bp. B: Primers Npt II F and Npt II R were used to amplify an 799 bp fragment of the nptII gene from all transgenic plants.

Regenerated plantlets with shoots of approximately 10 cm, still growing on media in tissue culture, were histochemically assayed for GUS activity, followed by clearing through an ethanol series. Shoots developed from callus bombarded with the intronless pBGUS UGDp showed no

detectable GUS staining. In the three lines transformed with pBGUS UGDip, however, similar, highly tissue specific GUS expression was observed. Blue staining was observed in cells at the base of developing leaves, on the midrib, just above the leaf sheath (Figure 5.5 a, b). Gus expression was also seen in the developing leafroll, increasing in intensity towards the centre of the leafroll. Under a microscope, the highest level of GUS expression in this tissue is visible in the guard cells of the developing stomata (Figure 5.5 d). Although GUS expression was clearly limited to the same cell- and tissue- type for all three the lines, a difference was observed in the intensity of blue staining. In one line, ip3.4 (Figure 5.5 a), GUS expression was consistently stronger than that observed for the other two lines, ip2.2 and ip3.0, respectively (Figure 5.5 b left and right, respectively). Blue staining of similar intensity was always observed for these two lines. No GUS expression was detected for the negative control, ip2.1 (Figure 5.5 c).

Histochemical GUS assays of tissue-culture grown plantlets (with shoots of approximately 10 cm) were repeated for the transgenic line containing pBGUS UGDip which showed the most intense blue staining after clearing through an ethanol series, namely ip3.4, one line transformed with pBGUS UGDp, clone p14, and the negative control, ip2.1. Stained tissues were then cleared with CLP. Clearing with CLP revealed high levels of background blue staining in the negative control (Figure 5.5 e, on right), and expression above background level could not be detected for clone p14 (results not shown). Clearing with CLP did, however, unmask significant GUS activity not previously detected for transgenic plant line ip3.4 (Figure 5.5 e, on left). Blue staining was observed in most tissues types of ip3.4, including the leaves, culm and roots.

Transgenic plants were then hardened off on regeneration medium without antibiotic, and subsequently grown under green house conditions. Green house grown sugarcane plants with between 10 and 12 internodes, were examined for GUS expression in the leafroll and young internodes. Assayed tissues were passed through an ethanol series and examined for blue staining. As before, no GUS expression was observed for transgenic lines transformed with the intronless pBGUS UGDp and the negative control, ip2.1 (Figure 5.5 f (right), h, j). In three transgenic lines produced by bombardment with pBGUS UGDip blue staining was observed in both tissue types that were assayed. In leafroll sections, blue staining in clone ip3.4 (Figure 5.5 f, left) was again more intense than the other two lines (ip2.2 Figure 5.5 f, middle). GUS activity

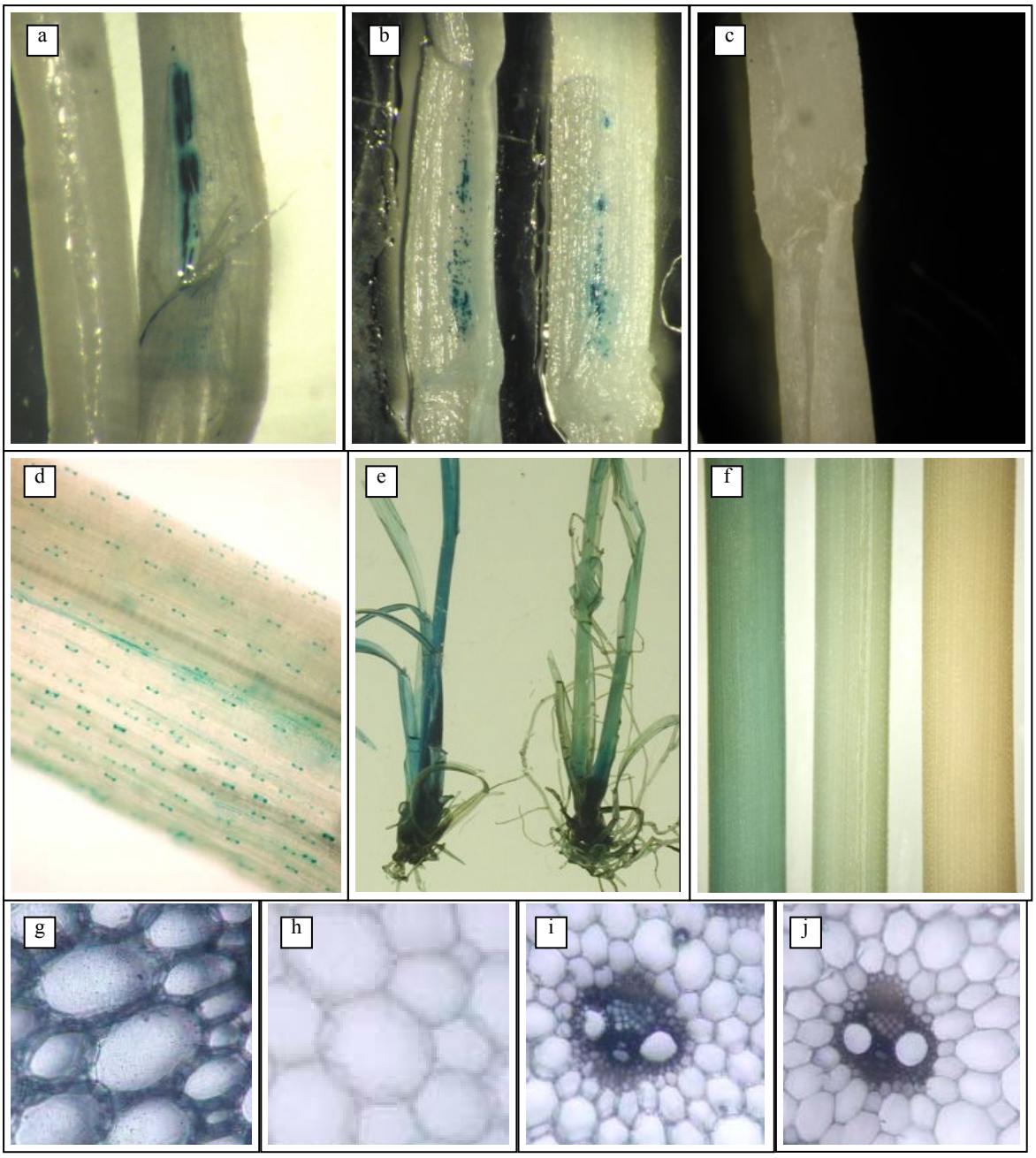


Figure 5.5 Histochemical assays of GUS expression in transgenic sugarcane transformed with pBGUS UGDip. When whole plants (taken from tissue culture) were assayed for GUS activity followed by clearing in an ethanol series, blue staining was observed at the base of developing leaves on the midrib just above the leaf sheath, for clones ip3.4 (a), ip2.2 (b, left) and ip3.0 (b, right). No such expression was observed in a negative control (c). GUS activity was also detected in developing leafroll of these clones, specifically in the guard cells of the stomata (ip3.4, d). Clearing of clone ip3.4 with CLP unmasked more GUS activity (e, left). CLP clearing also unmasked background staining in a negative control (e, right). In the leafroll of green house-grown clones ip3.4 (f, left) and ip2.2 (f, middle) GUS activity was detected, but not in a negative control (f, right). Stem sections of the same clones showed high levels of GUS activity in parenchyma cells (g) and vascular bundles (i). No blue staining was detected in a negative control (h, j).

was also observed in transverse culm sections taken from internode 4 of these lines. As expected, blue staining was observed in the cytosol of the sucrose storing parenchyma cells (Figure 5.5 g), as well as in the vascular bundles (Figure 5.5 i) (only results for clone ip3.4 are shown).

5.4.4 Southern blot analysis

To confirm stable integration of the transgenes, transgenic plants were subjected to Southern blot analysis. Southern blot hybridisation confirmed the PCR, which showed the presence of the reporter gene in three and six independent lines of pBGUS UGDip, and pBGUS UGDp, respectively (Figure 5.6). The enzyme used for digestion of genomic DNA has a single restriction site in both of the transformation vectors. The presence of multiple hybridisation bands is therefore consistent with integration of the transgene at multiple locations and/or a complex transgene array at a single locus. Based on the different hybridisation profiles observed, all the plants were derived from independent transformation events.

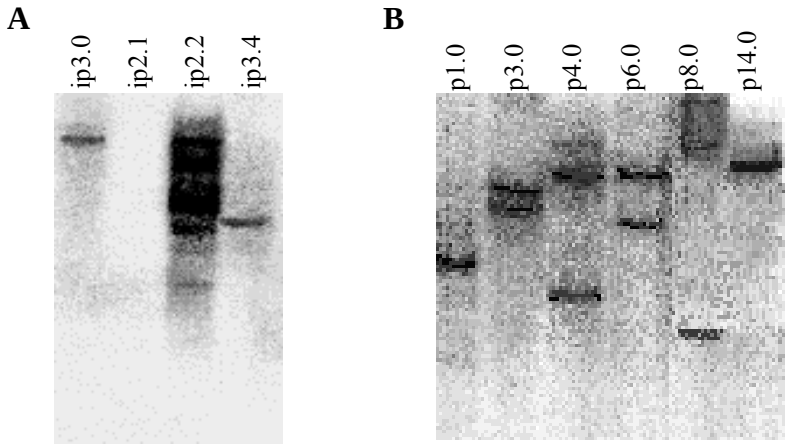


Figure 5.6 Southern blot analysis of transgenic sugarcane plants resulting from particle bombardment with (A) pBGUS UGDip and (B) pBGUS UGDp. A 1074 bp restriction fragment of the GUS gene was used to probe 5 µg of gDNA completely digested with *Xba* I, which has a single restriction site in both of the transformation vectors. The number of the clone represented is given at the top of each lane.

5.4 DISCUSSION

In the current study a 2.7 kb fragment upstream of a UDP-glucose dehydrogenase gene was isolated through screening a genomic library, and subsequently fused to a GUS reporter gene. The fragment was found to contain a large intron located 24 bp upstream of the translation start site, preceeded by an untranslated leader exon, and a further 1700 bp of putative promoter sequence. Examination of the nucleotide composition of the untranslated leader exon and surrounding sequences showed the presence of several C/T-rich sequences. C/T rich sequences located in the 5'-untranslated leaders of genes have been shown to play an important role in enhancing the expression of some genes (Bolle *et al*., 1994; Martinez-Trujillo *et al*., 2003; Zhang *et al*., 1991). Interestingly, C/T repeat sequences have also been detected in the untranslated leader exon, directly upstream of a 5'-leader intron, of most plant actin genes (An *et al*., 1996; McElroy *et al*., 1990; Nairn *et al*., 1988; Pearson and Meagher, 1990; Wang *et al*., 1992). Although most of these authors have suggested that the C/T repeats may act as transcriptional enhancers of plant actin genes, this has not been demonstrated experimentally.

Computational analysis of the promoter region (Higo *et al*., 1999) revealed the presence of many putative transcription factor binding sites (TFBS), providing further evidence that an active promoter sequence was indeed isolated. A putative TATA box is found between 43 and 37 bp upstream of the transcription initiation site. Six possible Dof protein binding sites were identified within the proximal promoter region. Dof proteins are a family of plant-specific transcription factors whose actions are believed to be related to biological processes unique to plants (Yanagisawa, 1996). These transcription factors have been shown to contribute to the regulation of genes involved in photosynthesis (Yanagisawa and Sheen, 1998), response to stress and hormone signals (Mena *et al*., 2002) and carbon metabolism (Yanagisawa, 2000). A possible role for Dof proteins in the regulation of UDP-glucose dehydrogenase, an enzyme involved in a process unique to plants, i.e. cell-wall synthesis, is therefore not surprising. Interestingly, Dof transcription factors have also been shown to regulate guard cell-specific gene expression (Plesch *et al*., 2000; 2001). As discussed later, in the current study most of the GUS expression detected in the developing leafroll of transgenic plants containing a promoter-GUS fusion was limited to the guard cells. Though several other consensus binding sites for known plant transcription

factors were also identified within the promoter sequence, the location of these motifs in the distal promoter region indicates a less significant role in the regulation of UDP-glucose dehydrogenase. It must be noted that the highly conserved core sequences of TFBS are relatively short and thus occur at a statistically predictable frequency in any give sequence. The functionality of the observed motifs was not tested as a part of this study. However, as more promoter sequences become available, a comparison of putative TFBS located in the promoters of genes with similar expression profiles could confirm the possible involvement of Dof transcription factors in the regulation of genes involved in cell-wall synthesis, and thereby provide more refined targets for deletion- or mutational analysis

Transient GUS expression assays showed that the 2.7 kb fragment had strong promoter activity in 5 day old maize coleoptiles, comparable to activity of the maize *ubi-1* promoter. *Ubi-1* has previously been shown to be one of the most efficient promoters for driving GUS expression in sugarcane (Gallo-Meagher and Irvine, 1993). This result indicated that all the major elements required for high gene expression levels were present in the isolated fragment. Deletion of the intron from the 2.7 kb fragment, leaving the untranslated exon and the 24 bp of exon 2 preceeding the translation initiation site intact, reduced expression of the reporter gene to barely detectable levels. Introns are documented in many cases to have a large positive effect on gene expression, especially in monocotyledonous species. Examples include introns from the maize *AdhI*, *Bz1*, (Callis *et al*., 1987) and *GapAI* (Donath *et al*., 1995) genes, and the rice *SalT* (Rethmeier *et al*., 1997), *Wx* (Li *et al*., 1995), *tpi* (Xu *et al*., 1994), and *Ostub 16* (Morello *et al*., 2002) genes.

An investigation of sequences available in Genbank confirms the presence of an intron in this position for UDP-glucose dehydrogenase from sorghum, rice and *Arabidopsis* (GenBank accession numbers AF503433, AL731873 and AL391143, respectively). All these species contain a large intron (sugarcane, 973 bp, sorghum 962 bp, rice 651 bp, and *Arabidopsis* 773 bp) located 20 to 30 bp upstream of the translation initiation site. Also, PCR amplification (described in Chapter 3) indicates that UDP-glucose dehydrogenase from sorghum also contains this 5'-UTR intron. Conservation in different species of an intron in this possition suggests an important role for this intron in the regulation of the expression of UDP-glucose dehydrogenase in plants.

As discussed in Chapter 2, other examples of the conservation of large introns in the 5'-UTR between a non-coding first exon and a coding second exon have also been found. Introns are found in this position in plant sucrose synthase (Fu *et al*., 1995a, b; Vasil *et al*., 1989; Shaw *et al*., 1994; Chopra *et al*., 1992; Komatsu *et al*., 2002), actin (McElroy *et al*., 1990; Pearson and Meagher, 1990; An *et al*., 1996; Huang *et al*., 1997; Thangavelu *et al*., 1993) and polyubiquitin (Norris *et al*., 1993; Christensen *et al*., 1992; Binet *et al*., 1991; Hoffman *et al*., 1991; Garbarino *et al*., 1995; Plesse *et al*., 2001; Wei *et al*., 1999; Wang *et al*., 2000) genes. Many of these introns have been shown to contribute to the regulation of the level and pattern of the expression of the associated genes (Christensen *et al*., 1992; Fu *et al*., 1995a,b; McElroy *et al*., 1990; Norris *et al*., 1993; Vasil *et al*., 1989; Wei *et al*., 1999).

Analysis of GUS expression in transgenic sugarcane plants showed that the 2.7 kb fragment retained promoter activity when reintroduced into plants. Deletion of the intron (as described above) resulted in a loss of all GUS expression (detectable above background GUS activity) in transgenic plants. When immature plants (with shoots of approximately 10 cm), transformed with the 2.7 kb fragment fused to a GUS reporter gene, were assayed for GUS expression, blue staining was observed in cells at the base of developing leaves, on the midrib, just above the leaf sheath. Transgene expression seemed to be limited to a specific cell type. A possible explanation is that certain sugarcane varieties, including the Nco310, the variety used in this study, develop a group of hairs in this position (Artschwager, 1940), which would require the deposition of a thick secondary cell wall. As expected, GUS expression was also seen in the developing leafroll, increasing in intensity towards the centre of the leafroll. Under a microscope, the highest level of GUS expression in this tissue is visible in the guard cells of the developing stomata where thick cell walls are actively being synthesised. GUS expression was clearly limited to the same cell- and tissue- types for all three the lines, confirming that the tissue specificity of the promoter was retained after transformation.

In whole plant preparations, the pigments present in plant tissues can partially or completely mask the sites of GUS activity. Stained tissues were passed through an ethanol series, which is known to remove pigments such as chloroform from the tissues without affecting the ClBr-indigo. Ethanol, however, makes the inherently opaque cell membranes of the plant tissue even

more opaque, thereby hindering GUS detection. To unmask more GUS activity, clearing with chlorallactophenol (CLP), which has been shown to make plant tissues transparent while not affecting the ClBr-indigo precipitate (Beeckman and Engler, 1994), was used as an alternative technique. Clearing with CLP unmasked significant GUS activity not previously detected in a transgenic plant where GUS expression was regulated by the 2.7 kb fragment. As can be expected from a promoter associated with actively growing tissues, blue staining was observed in most tissues types of this clone, including the leaves, stem and roots. Clearing with CLP, however, leaves tissues transparent, making it difficult to distinguish between areas of high and low GUS activity. Without the first intron no GUS expression above background could be visualized even after clearing with CLP.

Histochemical assays of leafroll material and transverse stem sections confirmed that the reporter gene was still actively expressed in mature plants. Blue staining was observed in the cytosol of the sucrose storing parenchyma cells as well as in the vascular bundles. UDP-glucose dehydrogenase expression in these tissues was demonstrated in chapter 4. This provides further evidence for the retention of tissue specificity after reinsertion of the transgene into the sugarcane genome.

According to the Southern blot analysis, the clone for which the highest levels of GUS expression were detected contains a single copy of the transgene. Two other clones which expressed similar levels of GUS, but at a lower level, contained 1 and 5 copies of the transgene respectively. Although only three transgenic lines containing the UDP-glucose dehydrogenase promoter were considered, at least for these lines there seems to be no correlation between GUS activity in transgenic plants and the copy number of the transgene. Similar findings were previously reported for the maize polyubiquitin (*ubi-1*) promoter (Hansom *et al.*, 1999).

Transgenic plant lines that showed GUS expression are currently being evaluated in field trials. Although the results of this study are only valid for green house grown plants, this is the first demonstrated isolation of a developmentally regulated promoter from sugarcane. Given that many promoters are silenced when introduced into sugarcane (Birch *et al*., 1996; Hansom *et al*.,

1999), retention of promoter activity and specificity driving foreign gene expression when reintroduced into sugarcane is a very encouraging result.

5.6 REFERENCES

Ahlandsberg, S., Sun, C., and Jansson, C. 2002, An intronic element directs endosperm-specific expression of the *sbellb* gene during barley seed development, *Plant Cell Reports*, **20**: 864-868.

An, Y. Q., Huang, S., McDowell, J. M., McKinney, E. C., and Meagher, R. B. 1996, Conserved expression of the Arabidopsis ACT1 and ACT3 actin subclass in organ primordia and mature pollen, *Plant Cell*, **8**: 15-30.

Arencibia, A.D., Carmona, E. R., Tellez, P., Chan, M. T., Yu, S. M., Trujillo, L. E., and Oramas, P. 1998, An efficient protocol for sugarcane (Saccharum spp. L.) transformation mediated by Agrobacterium tumefaciens, *Transgenic Research*, **7**: 1-10.

Arencibia, A., De la Riva, G., and Selman-Housein, G. 1995, Production of transgenic sugarcane (*Saccharum officinarum* L.) plants by intact cell electroporation., *Plant Cell Reports*, **14:** 305-309.

Artschwager, E. 1940, Morphology of the vegetative organs of sugarcane, *Journal of Agricultural Research*, **60** (8): 503-549.

Beeckman, T. and Engler, G. 1994, An easy technique for the clearing of histochemically stained plant tissue. *Plant Molecular Biology Reporter,* **12** (1): 37-42.

Benfey, P. N., Ren, L., and Chua, N.-H. 1990, Combinatorial and synergistic properties of CaMV 35S enhancer subdomains, *EMBO Journal*, **9** (6): 1685-1696.

Binet, M. N., Weil, J. H., and Tessier, L. H. 1991, Structure and expression of sunflower ubiquitin genes, *Plant Molecular Biology*, **17** (3): 395-407.

Birch, R. G. 1997, Plant transformation: problems and strategies for practical application, *Annual Review of Plant Physiology and Plant Molecular Biology*, **48**: 297-326.

Birch, R. G., Bower, R., Elliot, A., Potier, B., Franks, T., and Coredeiro, G. 1996, Expression of foreign genes in sugarcane, *Proceedings of the International Society of Sugarcane Technologists XXII Congress*, p. 368.

Birch, R. G. and Franks, T. 1991, Development and optimisation of microprojectile systems for plant genetic transformation, *Australian Journal of Plant Physiology*, **18**: 453-469.

Bolle, C., Sopory, S., Lübberstedt, T., Herrmann, R. G., and Oelmüller, R. 1994, Segments encoding 5'-untranslated leaders of genes for thylakoid proteins contain *cis*-elements essential for transcription, *Plant Journal*, **6** (4): 513-523.

Bower, R. and Birch, R. G. 1992, Transgenic sugarcane plants via microprojectile bombardment, *Plant Journal*, **2**: 409-416.

Callis, J., Fromm, M., and Walbot, V. 1987, Introns increase gene expression in cultured maize cells, *Genes and Development*, **1**: 1183-1200.

Chopra, S., Del-favero, J., Dolferus, R., and Jacobs, M. 1992, Sucrose synthase of *Arabidopsis*: Genomic cloning and sequence characterization, *Plant Molecular Biology*, **18**: 131-134.

Christensen, A. H., Sharrock, R. A., and Quail, P. H. 1992, Maize polyubiquitin genes: structure, thermal perturbation of expression and transcript splicing, and promoter activity following transfer to protoplasts by electroporation, *Plant Molecular Biology*, **18** (4): 675-689.

Christensen, A. H. and Quail, P. H. 1996, Ubiquitin promoter-based vectors for high-level expression of selectable and/or screenable marker genes in monocotyledonous plants, *Transgenic Research*, **5**: 213-218.

Dellaporta, S. L., Wood, J., and Hicks, J. B. 1983, A plant DNA minipreparation, *Plant Molecular Biology Reporter*, **1**: 19-21.

Donath, M., Mendel, R., Cerff, R., and Martin, W. 1995, Intron-dependent transient expression of the maize *GapA1* gene, *Plant Molecular Biology*, **28**: 667-676.

Fu, H., Kim, S. Y., and Park, W. D. 1995a, A potato *Sus3* sucrose synthase gene contains a context-dependent 3' leader element and a leader intron with both positive and negative tissue-specific effects, *Plant Cell*, **7**: 1395-1403.

Fu, H., Kim, S. Y., and Park, W. D. 1995b, High-level tuber expression and sucrose inducibility of a potato *Sus4* sucrose synthase gene require 5' and 3' flanking sequences and the leader intron, *Plant Cell*, **7**: 1387-1394.

Gallo-Meagher, M. and Irvine, J. E. 1993, Effects of tissue type and promoter strength on transient GUS expression in sugarcane following particle bombardment, *Plant Cell Reports*, **12**: 666-670.

Garbarino, J. E., Oosumi, T., and Belknap, W. R. 1995, Isolation of a polyubiquitin promoter and its expression in transgenic potato plants, *Plant Physiology*, **109** (4): 1371-1378.

Gibeaut, D. M. 2000, Nucleotide sugars and glycosyltransferases for synthesis of cell wall matrix polysaccharides, *Plant Physiology and Biochemistry*, **38**: 69-80.

Hansom, S., Bower, R., Zhang, L., Potier, B., Elliot, A., Basnayake, S., Cordeiro, G., Hogarth, D. M., Cox, M., Berding, N., and Birch, R. G. 1999, Regulation of transgene expression in

sugarcane, *Proceedings of the International Society of Sugarcane Technologists XXIII Congress*, *New Delhi. STAI, New Dehli (Ed. V. Singh)* pp 278-289.

Higo, K., Ugawa, Y., Imamoto, M., and Korenaga, T. 1999, Plant cis-acting regulatory DNA elements (PLACE) database, *Nucleic Acids Research*, **27** (1): 297-300.

Hoffman, N. E., Ko, K., Milkowski, D., and Pichersky, E. 1991, Isolation and characterization of tomato cDNA and genomic clones encoding the ubiquitin gene ubi3, *Plant Molecular Biology*, **17** (6): 1189-1201.

Huang, S., An, Y. Q., McDowell, J. M., McKinney, E. C., and Meagher, R. B. 1997, The Arabidopsis ACTII actin gene is strogly expressed in tissues of the emerging inflorescence, pollen, and developing ovules, *Plant Molecular Biology*, **33** (1): 125-139.

Jefferson, R. A., Kavanagh, T. A., and Bevan, M. W. 1987, GUS fusions: B-glucuronidase as a sensitive and versatile gene fusion marker in higher plants, *EMBO Journal*, **6** (13): 3901-3907.

Jeon, J.-S., Lee, S., Jung, K.-H., Jun, S.-H., Kim, C., and An, G. 2000, Tissue-preferential expression of a rice a-tubulin gene, *OsTubA1*, mediated by the first intron, *Plant Physiology*, **123**: 1005-1014.

Komatsu, A., Moriguchi, T., Koyama, K., Omura, M., and Akihama, T. 2002, Analysis of sucrose synthase genes in citrus suggests different roles and phylogenetic relationships, *Journal of Experimental Botany*, **53** (366): 61-71.

Last, D. I., Brettell, R. I. S., Chamberlain, D. A., Chaudhury, A., Larkin, P. J., Marsh, E. L., Peacock, W. J., and Dennis, E. S. 1991, pEmu: an improved vector for gene expression in cereal cells, *Theoretical and Applied Genetics*, **81**: 581-588.

Li, Y., Ma, H., Zhang, J., Wang, Z., and Hong, M. 1995, Effects of the first intron of rice Waxy gene on the expression of the foreign genes in rice and tobacco protoplasts, *Plant Science*, **108**: 181-190.

Luehrsen, K. R. and Walbot, V. 1991, Intron enhancement of gene expression and the splicing efficiency of introns in maize cells, *Molecular and General Genetics*, **225** (1): 81-93.

Maas, C., Laufs, J., Grant, S., Korfhage, C., and Werr, W. 1991, The combination of a novel stimulatory element in the first exon of the maize Shrunken-1 gene with the following intron 1 enhances reporter gene expression up to 1000-fold, *Plant Molecular Biology*, **16** (2): 199-207.

Martin, T., Frommer, W. B., Salanoubat, M., and Willmitzer, L. 1993, Expression of an *Arabidopsis* sucrose synthase gene indicates a role in metabolization of sucrose both during phloem loading and in sink organs, *The Plant Journal*, **4** (2): 367-377.

Martinez-Trujillo, M., Limones-Briones, V., Chavez-Barcenas, T., and Herrera-Estrella, L. 2003, Functional analysis of the 5' untranslated region of the sucrose phosphate synthase rice gene (*sps1*), *Plant Science*, **165**: 9-20.

McElroy, D., Blowers, A. D., Jenes, B., and Wu, R. 1991, Construction of expression vectors based on the rice actin 1 (Act1) 5' region for use in monocot transformation, *Molecular and General Genetics*, **231** (1): 150-160.

McElroy, D., Zhang, W., Cao, J., and Wu, R. 1990, Isolation of an efficient actin promoter for use in rice transformation, *Plant Cell*, **2**: 163-171.

Mena, M., Cejudo, F. J., Isabel-Lamoneda, I., and Carbonero, P. 2002, A role for the DOF transcription factor BPBF in the regulation of gibberellin-responsive genes in barley aleurone, *Plant Physiology*, **130**: 111-119.

Morello, L., Bardini, M., Sala, F., and Breviario, D. 2002, A long leader intron of the *Ostub 16* rice B-tubulin gene is required for high-level gene expression and can autonomously promote transcription both *in vivo* and *in vitro*, *Plant Journal*, **29** (1): 33-44.

Mun, J. H., Lee, S. Y., Yu, H. J., Jeong, Y. M., Shin, M. Y., Kim, H., Lee, I., and Kim, S. G. 2002, Petunia actin-depolymerizing factor is mainly accumulated in vascular tissue and its gene expression is enhanced by the first intron, *Gene*, **292**: 233-243.

Nairn, C. J., Winesett, L., and Ferl, R. J. 1988, Nucleotide sequence of an actin gene from *Arabidopsis thaliana*, *Gene*, **65**: 247-257.

Nelsestuen, G. L. and Kirkwood, S. 1971, The mechanism of action of uridine diphosphoglucose dehydrogenase, *Journal of Biological Chemistry*, **246** (12): 3828-3834.

Norris, S. R., Meyer, S. E., and Callis, J. 1993, The intron of *Arabidopsis thaliana* polyubiquitin genes is conserved in location and is a quantitative determinant of chimeric gene expression, *Plant Molecular Biology*, **21**: 895-906.

Pearson, L. and Meagher, R. B. 1990, Diverse soybean actin transcripts contain a large intron in the 5' untranslated leader: structural similarity to vertebrate muscle actin genes, *Plant Molecular Biology* , **14** (4): 513-526.

Plesch, G., Ehrhardt, T., and Mueller-Roeber, B. 2001, Involvement of TAAAG elements suggests a role for Dof transcription factors in guard cell-specific gene expression, *The Plant Journal*, **28** (4): 455-464.

Plesch, G., Kamann, E., and Mueller-Roeber, B. 2000, Cloning of regulatory sequences mediating guard-cell specific gene expression, *Gene*, **249**: 83-89.

Plesse, B., Criqui, M. C., Parmentier, Y., Fleck, J., and Genschik, P. 2001, Effects of the polyubiquitin gene Ubi. U4 leader intron and first ubiquitin monomer on reporter gene expression in Nicotiana tabacum, *Plant Molecular Biology*, **45** (6): 655-667.

Rethmeier, N., Seurinck, J., VanMontagu, M., and Cornelissen, M. 1997, Intron-mediated enhancement of transgene expression in maize is a nuclear, gene-dependent process, *The Plant Journal* **12** (4): 895-899.

Schenk, P. M., Remans, T., Sagi, L., Elliott, A. R., Dietzgen, R. G., Swennen, R., Ebert, P. R., Grof, C. P., and Manners, J. M. 2001, Promoters for pregenomic RNA of banana streak badnavirus are active for transgene expression in monocot and dicot plants, *Plant Molecular Biology*, **47** (3): 399-412.

Shaw, J. R., Ferl, R. J., Baier, J., StClair, D., Carson, C., McCarty, D. R., and Hannah, L. C. 1994, Structural features of the maize *sus1* gene and protein, *Plant Physiology* **106**: 1659-1665.

Snyman, S., Meyer, G. M., Carson, D. L., and Botha, F. C. 1996, Establishment of embryogenic callus and transient gene expression in selected sugarcane varieties, *South African Journal of Botany*, **62**: 151-154.

Terada, R. and Shimamoto, K. 1990, Expression of CaMV 35S-GUS gene in transgenic rice plants, *Molecular and General Genetics*, **220**: 389-392.

Thangavelu, M., Belostotsky, D., Bevan, M. W., Flavell, R. B., Rogers, H. J., and Lonsdale, D. M. 1993, Partial characterization of the *Nicotiana tabacum* actin gene family: Evidence for pollen specific expression of one of the gene family members, *Molecular and General Genetics*, **240**: 290-295.

Tzafrir, I., Torbert, K. A., Lockhart, B. E., Somers, D. A., and Olszewski, N. E. 1998, The sugarcane bacilliform badnavirus promoter is active in both monocots and dicots, *Plant Molecular Biology*, **38**: 347-356.

Vasil, V., Clancy, M., Ferl, R. J., Vasil, I. K., and Hannah, L. C. 1989, Increased gene expression by the first intron of maize *shrunken-1* locus in grass species, *Plant Physiology*, **91**: 1575-1579.

Wang, J., Jiang, J., and Oard, J. H. 2000, Structure, expression and promoter activity of two polyubiquitin genes from rice (*Oryza sativa* L.), *Plant Science*, **156**: 201-211.

Wang, Y., Zhang, W., Cao, J., McElroy, D., and Wu, R. 1992, Characterization of *cis*-acting elements regulating transcription from the promoter of a constitutively active rice actin gene, *Molecular and Cellular Biology*, **12** (8): 3399-3406.

Wei, H., Albert, H. H., and Moore, P. H. 1999, Differential expression of sugarcane polyubiquitin genes and isolation of promoters from two highly-expressed members of the gene family, *Journal of Plant Physiology*, **155**: 513-519.

Xu, Y., Yu, H., and Hall, T. C. 1994, Rice triosephosphate isomerase gene 5' sequence directs β-glucuronidase activity in transgenic tobacco but requires an intron for expression in rice, *Plant Physiology*, **106**: 459-467.

Yamamoto, Y., Tsuji, H., and Obokata, J. 1995, 5'-leader of a photosystem I gene in *Nicotiana sylvestris*, *psaDb*, contains a translational enhancer, *Journal of Biological Chemistry*, **270**: 12466-12470.

Yanagisawa, S. 1996, Dof DNA binding proteins contain a novel zinc finger motif, *Trends in Plant Science*, **1**: 213-214.

Yanagisawa, S. 2000, Dof1 and Dof2 transcription factors are associated with expression of multiple genes involved in carbon metabolism in maize, *The Plant Journal*, **21** (3): 281-288.

Yanagisawa, S. and Schmidt, R. J. 1999, Diversity and similarity among recognition sequences of Dof transcription factors, *The Plant Journal*, **17** (2): 209-214.

Yanagisawa, S. and Sheen, J. 1998, Involvement of maize Dof zinc finger proteins in tissue-specific and light-regulated gene expression, *Plant Cell*, **10**: 75-89.

Zhang, W., McElroy, D., and Wu, R. 1991, Analysis of rice *Act1* 5' region activity in transgenic rice plants, *Plant Cell*, **3**: 1155-1165.

Zhu, Q., Dabi, T., and Lamb, C. 1995, TATA box and initiator functions in the accurate transcription of a plant minimal promoter in vitro, *Plant Cell*, **7**: 1681-1689.

CHAPTER 6

CONCLUSIONS

The aim of this study was to isolate a developmentally regulated promoter, specific for young tissues that could be used for sugarcane transformation. For this purpose an endogenous gene that is expressed in the desired pattern in sugarcane, was selected. As sugarcane is a complex polyploid where a single gene copy can be represented by up to ten alleles, it was necessary to first establish whether promoter isolation from sugarcane is a viable option. At the start of this study, it was not known whether all of these alleles are expressed or if some of the gene copies have accumulated sequence changes inhibiting their expression, which would result in many sequences that represent silent copies of a specific gene, adjacent to non-functional promoters. If this were the case, it would greatly complicate promoter isolation from the sugarcane genome, as many sequences would have to be evaluated for promoter activity in order to find a single active promoter, making it a costly and time consuming exercise. This study provides the first evidence for the simultaneous expression of distinct haplotypes corresponding to a single gene in one sugarcane plant. Whether this is a common feature of sugarcane genes is not known at present. An interesting finding was that certain haplotypes seem to be represented at a greater frequency than others. It could be possible that these alleles occur more frequently than others, or that stronger promoters regulate the expression of these alleles. Further studies of allelic variation and expression of these alleles in sugarcane are required to fully explain this phenomenon. However, the possibility that multiple alleles are expressed to provide the required levels of a specific enzyme, rather than the expression of one dominant allele is encouraging for sugarcane gene and promoter isolation.

The promoter that was targeted for isolation in this study regulates the expression of UDP-glucose dehydrogenase. Based on evidence from other species, it was expected that expression of this gene would be strongly correlated with a demand for structural polysaccharides in the developing tissues of sugarcane. To ensure that this gene represented a suitable target for the isolation of a developmentally regulated promoter, the distribution of this enzyme in different sugarcane tissues and cell-types was investigated. A promoter that could be used to alter

metabolism in young developing tissues to increase the sucrose yield, for example, would have to be active in sucrose storing tissues i.e. the storage parenchyma. Results showed that the UDP-glucose dehydrogenase transcript is present in virtually all cell types present in sugarcane internodes and that abundance of the protein decreased in all cell types with increasing tissue maturity. It is therefore likely that UDP-glucose dehydrogenase dominates over the *myo*-inositol pathway in the provision of UDP-glucuronic acid as precursor for structural polysaccharides in the most actively growing sugarcane tissues, making its promoter an ideal candidate for the purposes of this study. This work represents the first reported isolation and analysis of a gene encoding UDP-glucose dehydrogenase from a monocotyledonous species. The importance of UDP-glucose dehydrogenase in the formation of cell walls is currently being investigated by expressing the isolated gene in an anti-sense orientation in transgenic sugarcane. It is not yet understood why cell wall polysaccharides from grasses are so different from those of dicotyledonous plants. Transgenic plants with modified hemicellulose contents resulting from altered levels of UDP-glucose dehydrogenase could provide some novel insights into this phenomenon.

The sequence upstream of the UDP-glucose dehydrogenase coding region was isolated and evaluated for its ability to drive transgene expression in both a transient system and stably transformed sugarcane. This sequence was found to contain a large intron preceded by a non-coding first exon. Introns in this position have often been shown to play an important role in the regulation of expression of the adjacent coding region. Results from this study demonstrated that an active promoter was indeed isolated, and that the presence of the intron was essential for strong reporter gene expression. Tissue specific expression was observed in transgenic sugarcane where expression of the reporter gene was regulated by the UDP-glucose dehydrogenase promoter and first intron. Deletion of the intron resulted in reduced expression of the reporter gene to barely detectable levels. Interestingly, an intron in this position in the gene was also demonstrated for several other plant species, including maize, sorghum, rice and *Arabidopsis*. Further characterisation of the intron sequence, for example by placing it in the context of different promoters, could clarify the role of this intron in the expression of UDP-glucose dehydrogenase in plants. If the intron is found to contribute to the regulation of the gene expression pattern, the sequences it contains could possibly be used to obtain tissue specific

expression in transgenic sugarcane, by combining these sequences with promoters that are known to be active in sugarcane. Given the fact that many promoters are silenced when introduced into sugarcane, redirecting the expression by promoters that are not could provide an attractive alternative.

The sequence of the promoter was also investigated using in silico analysis for possible clues relating to the regulation of UDP-glucose dehydrogenase gene expression. Deletion analysis to confirm the role of identified *cis*-acting sequence elements required for optimal activity can now be done. Also, as more promoter sequences become available, a comparison of putative transcription factor binding sites located in the promoters of genes with similar expression profiles could confirm the involvement of the identified motifs in the regulation of genes involved in cell wall synthesis. In this study the first developmentally regulated promoter was isolated from sugarcane. Further analysis of this promoter could provide valuable knowledge about the regulation of gene expression in sugarcane.

APPENDIX 1

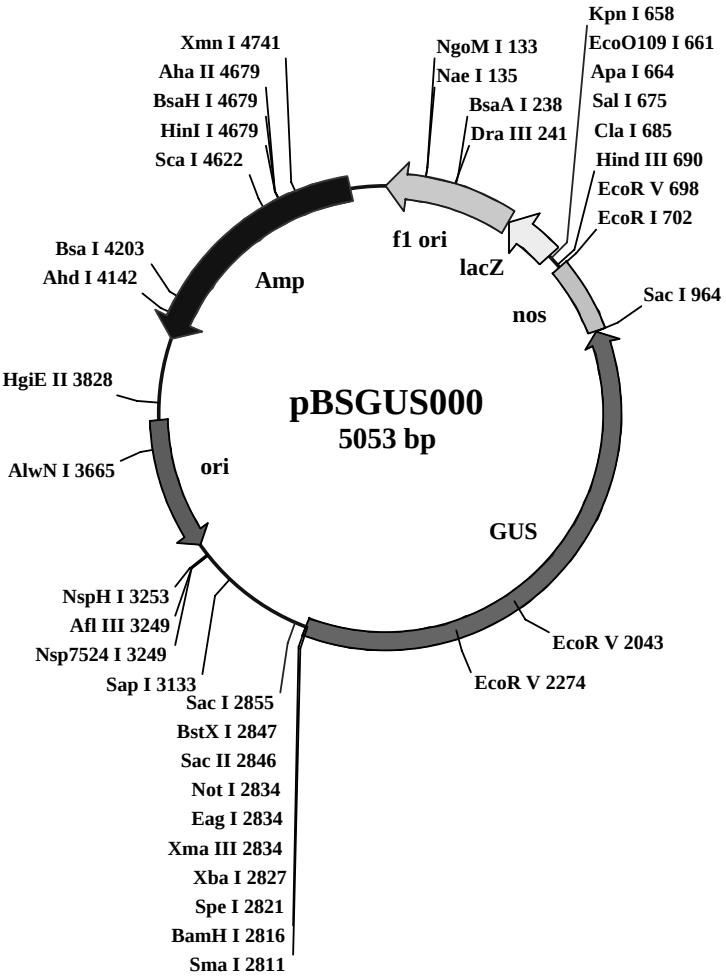
Nucleotide sequence of the UDP-glucose dehydrogenase promoter, intron and 5'-UTR.
The intron sequence is shaded in grey and the 5'-UTR is underlined.

TCTAGAATTTAAGTACGCATCAGTTAAGCGAGAAAACAGCGGCATTTGTTAACTTGCAGCATGT
TCGTTTCGCTGGTTTATTGTGGCTAAAAGTACTGTTCGCTAGTTTGTTGTGAGAGAAAAACACTG
TTGGATGGCTGCTGATTCCGCTGAATAGCTCAACCCAACAGAGACTTGGTAGCTTCAACATGTG
CCGATCCATGACTTAGAACTGTGGAAATTATTGAACTTTCTCACTAGAGTCTTTTCGTAAACGGT
AGCTTCAGTCTTTGTCAGTCTATTGCTAGCATTAATCAAATGACAGTGACTGCAGGCTAAGAAA
GTTCTCCCTTTTCCTGCTTACACTCCGTCTCTGCATAAAGAACGTCTCTCCAATATTTTTTTCCT
TTTTAATTTTTCTTTTAACTTTTTAGTAGATGGCATCAACGCTGGAGTTTCCTTTCCTTTCAGTAG
ATCCGGCCTCTGTTTCTTAATCAGCAAACAAAAACACACGTTTTGTTTGGTGGAGAGTAATCAC
TGGTAGAAGCATCAGCATCAGCTTCCTGGATCTGACTGTGTCCAGGCTATGAAAATTTCTACTG
GCTGGGCTCCGTTGATCCGGCCTGCATTTCTCACGTCTCTGTCACCGATGCTGTAACTTTTGCGT
CCTCGATTGTTACTCGCTCCATATTACTCCCTTCGTCTTGTAAAAAGAGTCATTCTCGCTTCCCA
AAAAGTCAACAGCTTTTAACTTTAACCAAATATATTTAACAAAATATTAATATTTATGGTACAT
AATTAGTATCTTAGGTAGATCTTTGAATATACTTTCATAATAAACTTATTTAAAGATATAAATGT
GCACGTATTTTCTACAAACGTAGTTAAAGTTGATAAAGTTTAACCAGCACGCTTTTCACGGTGA
CTCTTTTTATGGGACGGAGTGGTATATGGTTCAGAGCCCAATTATTCGTTTGCTGCCTGTAGTTG
ACTGTGGATTAAATTACATCAACTCTTTGGTTATTCTTCAGACCATGTTCCTTGACTTGCTTCCA
GTTCATGGGCTACATGTACGCACCCATCTTGCTGCCTGACCTTACTACTCTCTCTGTGCACTGAA
CAGAGAGAAGAGCCAGATCTGATCAGAGAGTAATCAGACTGCTCAGTTACACCAAAACTGCTC
AATCAGCAACTGATGAAAAAAAAAAGAAAAGGGAGAAGCTTTATCAGCACCTTTCGTTTATGC
GCTTGATGGTGCTTGAGAATAATTTTCTTTTTCTGTGAGGAGGAGGTTGAGGGCATAGAAGGCA
ATGGTTAGGTGGAGTGTCATGGACTCATGTGAGAGCCTTTTTGATTAATGCAAGTTTGGTGGCA
TTGGCATGCATGAGCAGCACTCGTCGTCTCCTCAGTTGCTTCTGGTTTCTGACCAGCCTTGACGC
GGTTACTATTTTTTTTTTCTACGGGACACGGCACGGTTATTTTTTGAAGCCGAGGGGTAGATCTG
GTTCGGTGGGTTTGCGGGCCGGGGTATTTAAGGACTCGTGGCCTCGCTTCTCCTCTGCACACAC
ACAAAGCGAGCTCCTCTTCTCCCACGCATCGCGCGAAGGAAGAGAGAGATCGCCGCTCCTCTTC
TCTTCGTCGCCTCTACTGGTGTGGGTGCCCGAGATCTTGCTGCAGTTCCTCGATATCTGGTGAGA
CTCGCGCCCTGATCCCACGCTGTTACTTCAATTTTTTTTCCTTTCTGATGAGTCCCGCGCCTGATC
ATGCTAGTTGAGGATCGTGATTTCATTCTGGGGCTAGTTTGTTCCTCGGCTTCGTTCTGCTGCAA
GCAGTTGTCTCCTCGAGTTCGTTCTCTGAGGACGACCCTGTCTTCAGATCTGGAAGTTTCGTGGT
TTGGATCTGATAGGAGTTACGGTTACGCGTTGTAGTGGCTGAAGAGGATTTGCGCCCCCCTAAT
GTTGCCTTTGCACCCCTACTGGATCAGAGTTTCAGATCTTTCTGTCTGGTCGCAGTTTTAAACGC
CTTGACGGTGCTCATATACGAGTAAATTTCATTTTATAATTCGGTGGTCATGGATCTGTATCCTG
CGATTCCTGCCCTTCACTCTGTTTAGGTGGTGTTTGTGTCAGGTAACGGCAGTGCACGTAATGGG
AATGGGATCACAGCTGCCTGGTTGTGTAATAGGTGGTTGCTTAATTTGTCTGGTTTTTCTTAGAA
AAGGTTTTAGAAAGCATTAATTTCACATGTATTCGGTGGAGAAGAGGGGCCATCCCAATCTTGA
TCAGCAGAAAACCTGAGATGTGCCGTGCCTGCTTTAACAAATATGAGTTCTGATGAATGTGTCC
TAACACTTCCACCCAAAACAAACCTTTTTAGATGTGGTTTGACCTGGTCGGTAGATGTTGCCAAT
TGCTATACCTAGTATATCCGAACTCCTATTTATTCAGTTCACTTAAGTAACCTTCTCGGAACATG
CTTAGCTATGCTCTGTTATTTCAGTCGTGCACTAGTAGTCTAGTGTTCAGGATCTGGTGCTTTTG
CTTGTGAATCCTGATTGCTTCTGTTAACTTGGGAAACTTGGTTGACATGTCTCATCTGTCAGATA
TTCTCAGGGTATGTGCTCTATTTTACAATCCACTAATGAAGCACCTTTTGCAATTACAGGTCACA
GATCTATCTGACAAGATG

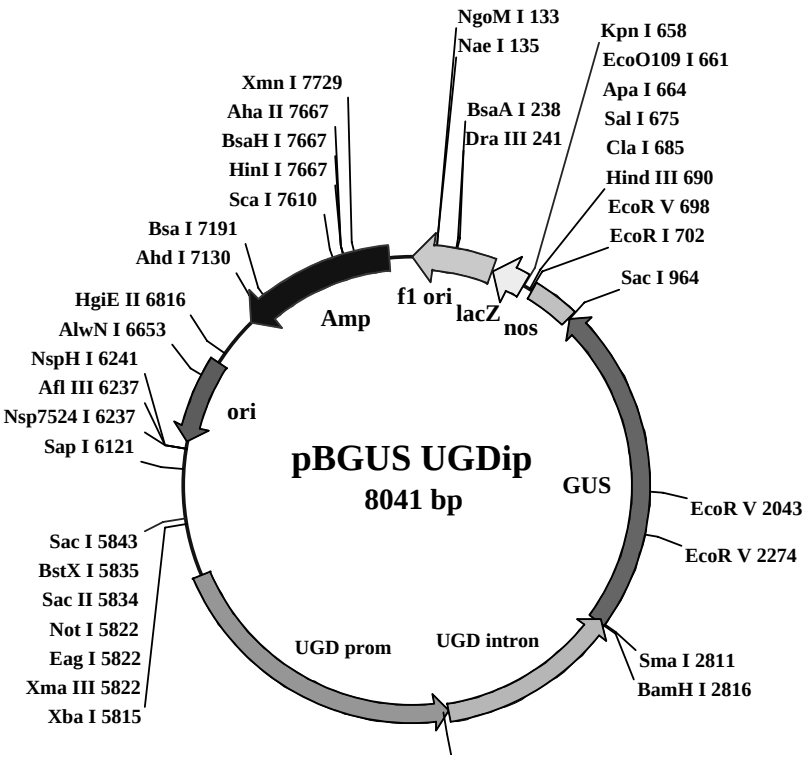
APPENDIX 2

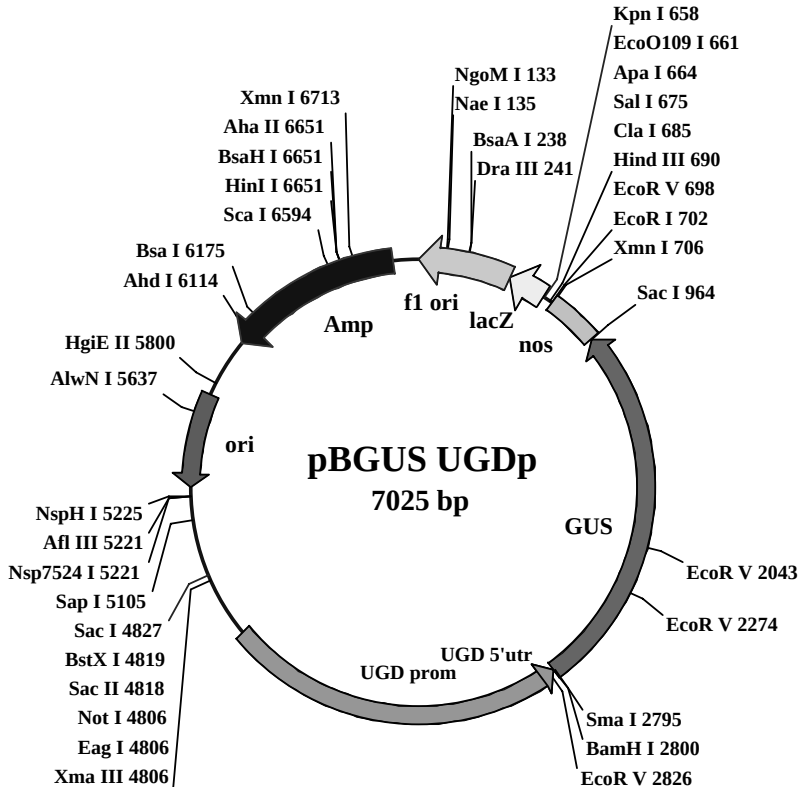
Schematic representation of reporter gene constructs. **A** pBGUS000**:** A *BamH* I-*Eco* RI fragment containing the GUS gene and adjacent *nos* terminator, was removed from the construct, pAHC27 (Christensen *et al*., 1992) and cloned into a pBluescript II KS + cloning vector (Stratagene). **B** pBGUS UGDip: The *Xba* I - *BamH* I treated 2700 bp fragment containing the putative UDP-glucose dehydrogenase promoter and 5'-UTR intron was cloned into the promoterless pBGUS 000 vector, prepared using the same restriction enzymes. **C** pBGUS UGDp: The *Xba* I - *BamH* I treated 1700 bp fragment containing the putative UDP-glucose dehydrogenase promoter without the 5'-UTR intron was cloned into the promoterless pBGUS 000 vector, prepared using the same restriction enzymes

A



B


NgoM I 133
Nae I 135
Kpn I 658
EcoO109 I 661
Apa I 664
Sal I 675
Cla I 685
Hind III 690
EcoR V 698
EcoR I 702
Sac I 964
Xmn I 7729
Aha II 7667
BsaH I 7667
HinI I 7667
Sca I 7610
Bsa I 7191
Ahd I 7130
BsaA I 238
Dra III 241
f1 ori
lacZ
nos
Amp
HgiE II 6816
AlwN I 6653
NspH I 6241
Afl III 6237
Nsp7524 I 6237
Sap I 6121
ori
pBGUS UGDip
8041 bp
GUS
EcoR V 2043
EcoR V 2274
Sac I 5843
BstX I 5835
Sac II 5834
Not I 5822
Eag I 5822
Xma III 5822
Xba I 5815
UGD prom
UGD intron
Sma I 2811
BamH I 2816


Kpn I 658
EcoO109 I 661
Apa I 664
Sal I 675
Cla I 685
Hind III 690
EcoR V 698
EcoR I 702
Xmn I 706
Sac I 964
NgoM I 133
Nae I 135
Xmn I 6713
Aha II 6651
BsaH I 6651
HinI I 6651
Sca I 6594
Bsa I 6175
Ahd I 6114
BsaA I 238
Dra III 241
f1 ori
lacZ
nos
Amp
HgiE II 5800
AlwN I 5637
ori
pBGUS UGDp
7025 bp
GUS
EcoR V 2043
EcoR V 2274
NspH I 5225
Afl III 5221
Nsp7524 I 5221
Sap I 5105
Sac I 4827
BstX I 4819
Sac II 4818
Not I 4806
Eag I 4806
Xma III 4806
UGD 5'utr
UGD prom
Sma I 2795
BamH I 2800
EcoR V 2826