Molecular Cloning and Characterization of *Leucaena leucocephala* β-Glucosidase, a Family 1 Glycosyl hydrolase

A THESIS
SUBMITTED TO THE
UNIVERSITY OF PUNE

FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

IN
BIOTECHNOLOGY

By
Noor Mahammad Shaik

Under the Guidance of

Dr. B. M. Khan
HEAD
PLANT TISSUE CULTURE DIVISION
NATIONAL CHEMICAL LABORATORY
PUNE – 411 008
INDIA

March 2009

Dedicated to my family and friends

Contents Page No.

CERTIFICATE

This is to certify that the work incorporated in the thesis entitled **"Molecular Cloning and Characterization of *Leucaena leucocephala* β-Glucosidase, a Family 1 Glycosyl hydrolase"** submitted by **Noor Mahammad Shaik** was carried out under my supervision at the Plant Tissue Culture Division, National Chemical Laboratory, Pune. Materials obtained from other sources have been duly acknowledged in the thesis.

Dr. B. M. KHAN

(Research Guide)

DECLARATION

I hereby declare that the work reported in this thesis titled **"Molecular Cloning and Characterization of *Leucaena leucocephala* β-Glucosidase, a Family 1 Glycosyl hydrolase"** is entirely original and submitted for the Degree of **Doctor of Philosophy** to the University of Pune. It was carried out by me at Plant Tissue Culture Division, National Chemical Laboratory, Pune - 411 008, Maharashtra, India, under the supervision of **Dr. B. M. Khan**. I further declare that it has not formed the basis for the award of any degree or diploma of any other University or Institution.

March, 2009

Noor Mahammad Shaik
PTC Division,
NCL, Pune-411008.

Acknowledgements

I am deeply grateful to my Research Guide Dr. Bashir Mohammad Khan, who has been a constant source of inspiration, motivation and a mentor. His calm and balanced nature, systematic approach to scientific problem resolution and careful interpretation of experimental results combined with his limitless virtue of patience, support and affection have helped me immensely in completing this thesis work and I am deeply indebted to him.

Special thanks are due to Dr. S. K. Rawal for being an unrelenting source of encouragement and support. His expert advice always gave new strength to face the challenges.

I also wish to acknowledge Dr. S. Sivaram, Director, NCL for providing necessary infrastructure to carry out my research.

I greatly aknowledge the contributions of Prof.S.Ramakumar, chairman, Bioinformatics centre IISc Bangalore for their cooperation and encouragement.

My heartfelt gratitude goes to Dr.Sucheta.N.Karande for making me feel homely in pune during the course of this work.

Thanks are due to all members of the scientific and supporting staff of the Plant Tissue Culture Division, Biochemical Sciences Division and NCL Library.

I thank my labmates and friends RC, Rohini, Neelima, Nookaraju, Manish, Sushim, Sameer, Arun , Pallavi, Abhilash, Sumita, Ruby, Santosh Jhadav, Sameena didi, Ektabhabhi, Gayatriji, Rajeshri, Kannan and Santosh Gupta for their encouragement and help during the course of this work. Azfar, Malini, Ulfat and Jay Kumar deserve special thanks, for the cordial atmosphere in the lab.

I thank trainees Karishma, Sobia, Priya, Ramakrishna Reddy, Amol and prashant babu, Sulakshnana, Mayaji, Jayesh, Anurag and Kiran for their help during my experimental work.

I thank my friend Anurag at IISC,Bangalore for his contribution in computational analysis for this work.
I thank shashidhara, biochemical sciences division, for his contribution in enzyme kinetic studies of this work.

I take this opportunity to thank my friends Bhushan sir, Manish Pathak Sreekanth, Nagraj, JP, Sagar, Manje, Elan, Santosh, Chetan, Atul F., Satya, Ambrish, Anish, Aarohi, Nitin, SuryaPrasad, Sachin, Shabab, Shadab, Poonam, Ansary, Asad, Kishore bhaiya, Asutosh, Sunil, Bhuban, Kiran, Shweta, Swapna, Shyla, Sujata and Venkatesh for their help, support, best wishes and lighter moments shared together throughout my work period.

I thank my friends at NCCS Pune, Vikku, Pankaj, Ranveer, Bhaskar, Amita for their kind cooperation in this work. I thank my friend at NIV, styender for his kind coopration in this work.

Special mention is a must for Ashish, Rishi, Somesh , Mohsin, Anubhav,Rodu, tonima and sameer for being my extended family during the course of my work. They have given me a perfect family atmosphere and a home away from home.

I would like to gratefully acknowledge the fellowship awarded by the CSIR, India.

I owe my deepest gratitude to my grandparents, parents, brothers and sister who have constantly encouraged me to excel. I would like to thank my fiancé Shabetaj for her motivation and support during this work.

Finally I thank Almighty for blessing me with good health to carry out the work and guiding me to realize a higher consciousness of love for all forever.

Noor Mahammad Shaik

Abbreviations

AA	Amino acid
AldOMT	5-Hydroxyconiferaldehyde O- methyltransferase
BAP	6- Benzylaminopurine
bp	Base pairs
BSA	Bovine serum albumin
C3H	Coumarate 3- hydroxylase
C4H	Cinnamate 4-hydroxylase
CAD	Cinnamyl alcohol dehydrogenase
CAld5H/ F5H	Coniferaldehyde 5- hydroxylase / Ferulate 5- hydroxylase
CBG	Coniferin β glucosidase
CCoAOMT	Caffeoyl coenzyme A 3-O- methyltransferase
CCR	Cinnamoyl coenzyme A reductase
cDNA	Complementary DNA
Ci/ mmol	Curie per milli mole
CIAP	Calf Intestinal Alkaline Phosphatase
COMT	Caffeate O-methyltransferase
Cps	Counts per second
Da	Dalton
DEPC	Diethylpyrocarbonate
DNA	Deoxyribose nucleic acid
DTT	Dithiothritol
EDTA	Ethylene Diamine Tetra Acetic acid
EDTA	Ethylene diamine tetra acetic acid disodium salt
ELISA	Enzyme linked immuno sorbent assay
F5H/CAld5H	Ferulate 5-hydroxylase/ coniferaldehyde 5-hydroxylase
FTIR	Fourier Transform Infra Red
g /L	grams per litre
g	gram
G	Guaiacyl
gDNA	GenomicDNA

GSP	Gene Specific Primers
h	Hour(s)
IPTG	Isopropyl β-D-thiogalactoside
Kb	Kilobase pairs
KDa	Kilo Daltons
Kg	Kilogram
Km	Michaelis-Menton constant
L	Litre
LD_{50}	Lethal dose 50%
MCS	Multiple cloning sites
mg	milligram
min	Minute(s)
mL	millilitre
mM	millimolar
mRNA	messenger RNA
NAA	1-Napthyl aceticacid
nM	nano molar
nm	nanometer
NUP	Nested Universal Primers
O/N	Overnight
PAL	Phenylalanine ammonia lyase
pg	picogram
pmol	picomole
PMSF	Phenyl methyl sulphonyl fluoride
ppm	Parts per million
RNA	Ribose nucleic acid
rpm	Rotations per minute
RT	Room temperature
s	second(s)
S	Syringyl
SAD	Sinapyl alcohol dehydrogenase
SDS	Sodium dodecyl sulphate (sodium lauryl sulphate)
SMQ	Sterile Milli Q

Soln	Solution
sp.	Species
TDZ	Thiadiazuron
U	Units
UDP-GT	UDP-glycosyltransferase
UPM	Universal Primer Mix
UTR	Untranslated Region
UV	Ultraviolet
V	Volt
v/v	volume / volume
Vmax	Maximum velocity
w/v	weight / volume
X-gal	5-bromo-4-chloro-3-indolyl-β-D-galactoside
α	Alpha
β	Beta
λ	Lamda
%	Percentage
°C	degree Celsius
µg	microgram
µg/L	Micrograms per liter
µL	microlitre
µm	micrometer
µM	micromolar
4CL	4-Coumarate coenzyme A ligase

1. General Introduction

1.1. Glycosyl hydrolases and families

Glycoside hydrolases (GHs) are widely distributed enzymes that hydrolyse the glycosidic bond between two or more carbohydrates or between a carbohydrate and a non-carbohydrate moiety (Xu *et al*., 2004). These enzymes have been classified into families based on amino acid sequence similarities (Henrissat, 1991; Henrissat and Davies, 1997). It is reported that there are more than 100 Glycosyl Hydrolase families (Opassiri *et al*., 2006). The distinguishing features and representative members of these enzymes are described on the CAZy (ModO) server (Coutinho and Henrissat, 1999). Carbohydrates and their glycoconjugates are one of the most diverse groups of organic molecules in the biosphere. The selective cleavage of glycosidic bonds is crucial in a variety of fundamental biological processes for all living organisms. The large (and growing) number of glycoside hydrolase families reflects the diversity of substrates and the need for selective cleavage of the glycosidic bonds (Verdoucq *et al*., 2004). Glycoside hydrolases have been reported in Eukarya, Archaea, and Bacteria. The β-glucosidases that have been characterized to date fall predominantly in glycosyl hydrolase families 1 and 3 (Henrissat *et al*., 1991), with family 1 enzymes being more numerous in plants (Opassiri *et al*., 2006).

1.1.1. Family 1 glycosyl hydrolases

The glycosyl hydrolase (GH) family 1 in the carbohydrate-active enzyme database (http://afmb.cnrs-mrs.fr/CAZY) is grouped together with 16 other families in clan GH-A, which is the biggest of the 13 clans of glycosidases. Although a number of three-dimensional structures already exist in family 1, only eight are from eukaryotic sources, six of which are of plant origin. These GHs are well known to be involved in plant defense mechanisms (Barleben *et al*., 2007).

1.1.1.1 Functional genomic analysis of Family 1 glycosyl hydrolases in *Arabidopsis*

Multigene families are believed to have arisen from a single ancestral gene by gene duplication via either unequal crossing-over or intra/ inter-chromosomal transposition after chromosome breaks. In addition to these processes, duplication events involving

large segments of chromosomes are thought to have further increased the number of genes in what were initially small multigene gene families (Vision *et al*., 2000; Simillion *et al*., 2002; Blanc *et al*., 2003).

The recent completion of the *Arabidopsis* genome sequencing project, together with its automated annotation process, enabled for the first time to gauge the number of GH Family 1 hydrolases in a typical angiosperm (Initiative, 2000). Reannotation resulted in the recognition of 48 members within *Arabidopsis* GH Family 1. The 47 members share a common evolutionary origin and the forty-eighth member of the family is a β-glucosidase-like gene (At3g06510). It was earlier designated as sfr2 (Thorlby and Warren, 2002), which belongs to a distinct lineage in phylogenetic tree. Six β-glucosidase (BGLU) genes appear to encode S-glycosidases (myrosinases). In these enzymes, the TI/LNQL motif replaces the TFNEP (acid/base catalyst) peptide motif typical of Family 1 hydrolases. The remaining 41 BGLU genes are believed to encode O-glycosidases, but some family members may hydrolyze non-carbohydrate substrates like hydroxyisourate also (Raychaudhuri and Tipton, 2002).

At least six β -O-glycosidase genes (BGLU1, BGLU2, BGLU5, BGLU6, BGLU14, and BGLU43) and two myrosinase genes (BGLU36 and BGLU39) appear to be pseudogenes (as defined by Snyder and Gerstein, 2003), because, if expressed, their predicted polypeptide products would lack key motifs or residues essential for β-glycosidase activity due to reading frame shifts and/or premature translation termination. Examination of a multisequence alignment of the deduced Arabidopsis BGLU polypeptides reveals that all except those encoded by the eight pseudogenes exhibit the highly conserved GH Family 1 motifs (e.g. S/AAY/FQV/WEG, RFSIWSRIFP, TF/LNEP, APGRCS, I/VT/ SENG, GYFAWSLXDNFEW, and variants thereof). Over 40 members including four putative β-S-glucosidases were recognized, but their biological roles remain unknown.

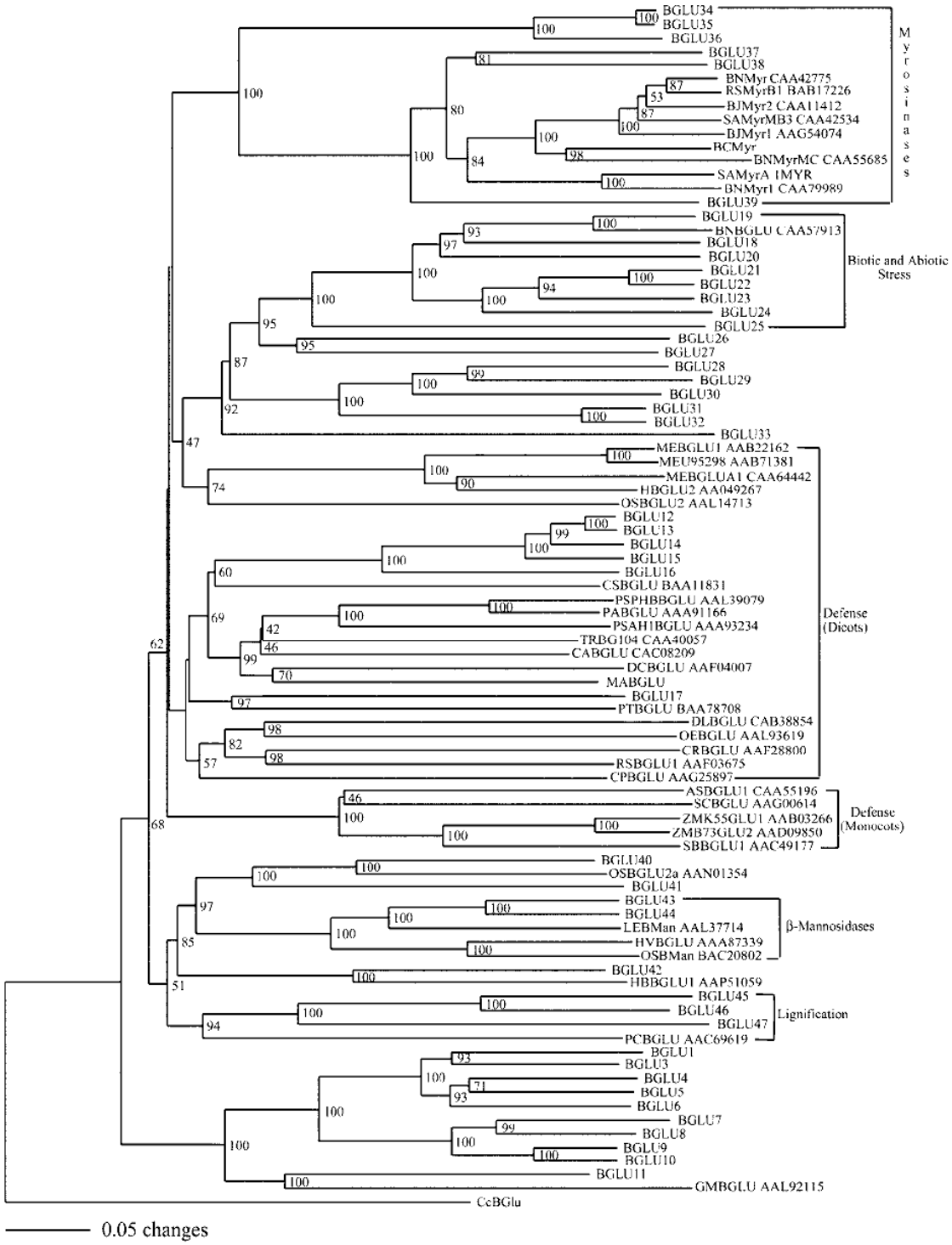


Fig1.1. A phylogenetic tree depecting the relationship between GH Family 1 members from different species.

The neighbor-joining tree (Fig 1.1) was constructed using the program PAUP 4.0b2 after alignment of aminoacid sequences by ClustalX followed by manual editing.The numbers shown at each node represent support values (percent) obtained by bootstrap analysis (1000 replicats). The *C.elegance* β-glucosidase (NP_497558) was used to root the tree. All branches are drawn to scale as indicated by the scale bar (=0.05 substitutions/site). Putative biological roles were ascribed here to several functions to all family members using biochemical, molecular, and genetic approaches (Source: Xu *et al*., 2004).

The biochemical properties, including substrate specificities, of several GH Family 1 members from other plant species have been characterized during the past two decades. To find the phylogenetic and functional relationship among Arabidopsis BGLU genes a phylogenetic tree was constructed that includes members from a wide variety of taxa including *Arabidopsis*. Based on their (*Arabidopsis* BGLU) clustering with genes of known biochemical function from other species they were assigned putative biological roles. Several interesting observations can be made from Figure 1.1: BGLU45, BGLU46, and BGLU47 cluster with the *Pinus contorta* β-glucosidase that exhibits pronounced specificity toward coniferin and syringin (Dharmawardhana *et al*., 1995, 1999). BGLU43 and BGLU44 cluster with mannosidases. When these *Arabidopsis* BGLU homologs were expressed in heterologous system and their substrate specificity tested, it is in accordance with the phylogenetic tree (Xu *et al*., 2004; Escamilla-Travino *et al*., 2006).

1.1.1.2. Functional genomic analysis of Family 1 Glycosyl Hydrolases in *Oryza*

The completion of the *Oryza sativa* L. spp. japonica rice genome project and the complementary indica rice (*O.sativa* L. spp. indica) genome project by the Beijing Genomic Institute (BGI) has allowed genome-wide analysis of gene families in this important monocot grain crop (IRGSP 2005,Yu *et al*., 2002). The sequence and mapping information provided to the public databases by these projects enabled to identify the genes for glycosyl hydrolase family 1 members (putative β-glucosidases) in rice, determine their gene structures and genomic organization, and model their protein products and phylogenetic relationships (Opassiri *et al*., 2006). Forty β-glucosidase

genes, including 34 full-length genes, 2 pseudogenes, 2 gene fragments, and 2 intronless genes, were identified. Thirty-six out of 40 genes are found in both japonica and indica rice with 98–100% sequence identity. In order to better characterize the function of the GH1 multi-enzyme family in rice, recombinant expression of these genes or their cDNAs to produce the enzymes is necessary (Opassiri *et al*., 2006).

1.2. Functions of Family 1 glycosyl hydrolases

Glycoside Hydrolase (GH) Family 1 members typically hydrolyze substrates of the type G-O-X (or G-SX), where G represents a β-linked glucosyl, galactosyl, mannosyl, fucosyl, 6-phosphoglucosyl, 6-phosphogalactosyl, or glucuronide residue, whereas X represents another glycosyl residue or a non carbohydrate group (i.e. aglycone). The nature of the aglycone moiety of substrates is believed to be critical for the physiological functions of these enzymes (Xu *et al*., 2004). GH Family 1 hydrolases play many diverse and important roles in living organisms. For example, in cellulolytic bacteria and fungi, they are key components of the cellulase complexes secreted by these organisms to degrade cellulose to glucose (Be´ guin, 1990; Fowler, 1993). Whereas in mammals the liver cytosolic β-glucosidase hCBG is a xenobiotic metabolizing enzyme that hydrolyses flavonoid glucosides (Berrin *et al*., 2003). In higher plants, these hydrolases have been implicated in such fundamental processes as chemical defense against herbivory (Niemeyer, 1988; Poulton, 1990; Rask *et al*., 2000), lignifications (Dharmawardhana *et al*., 1995), hydrolysis of cell wall-derived oligosaccharides during germination (Leah *et al*., 1995), and regulation of the biological activity of phytohormones by hydrolysis of their inactive hormone-glucoside conjugates (Falk and Rask, 1995).

Plant family 1 glycosyl hydrolases tend to show high specificity for their aglycones, though many hydrolyze synthetic, nonphysiological substrates, like p-nitrophenol (pNP)-β-glycosides (Esen, 1993). The aglycones span a wide range of structures, including sugars (Leah *et al*., 1995; Akiyama *et al*., 1998; Opassiri *et al*., 2004), hydroxaminic acids (Babcock *et al*.,1994), isoflavonoids (Chuankhayan *et al*., 2005), rotenoids (Svasti *et al*., 1999), alkaloids (Geerlings *et al*., 2000; Warzecha *et al*., 2000), hydroxyquinones (Duroux *et al*.,1998) and cyanogenic nitriles (Poulton *et al*.,1990). It is

the specificity for these aglycones which is thought to specify the function of most of these enzymes (Esen, 1993). Since many β-glucosidases function in plants, it is important that these enzymes specifically hydrolyze their own substrates and not other substrates with which they may come into contact. It seems evident that the substrate specificity, localization of the enzymes with respect to potential substrates, and the activities of the substrates and hydrolysis products will determine the roles of these enzymes (Opassiri *et al*., 2006).

1.3. Mechanism of action of Family 1 Glycosyl Hydrolases

Glycosyl hydrolases are retaining enzymes. They operates via a mechanism involving two steps: the formation of the glycosyl–enzyme with concomitant aglycon departure (glycosylation step); and hydrolysis of the glycosyl–enzyme by a water molecule (deglycosylation step) (Sinnott *et al*., 1990; McCarter and Withers, 1994; Burmeister *et al*., 1997).

β-Glucosidases belonging to the family 1 glycoside hydrolases catalyze the hydrolysis of the glucosidic bond between the anomeric carbon (C1 of the glucose) and the glucosidic oxygen by a mechanism in which the anomeric configuration of the glucose is retained (Davies and Henrissat, 1995). The catalytic mechanism is illustrated in Fig. 1.2 Two conserved glutamic acid residues serve as a catalytic nucleophile and a general acid/base catalyst, respectively. In retaining β- glucosidases, the catalytic glutamic acid residues are situated on opposite sides of the β-glucosidic bond of the docked substrate at a distance of 5.5 Å (Davies and Henrissat, 1995). As the initial step in catalysis, the nucleophile (B) performs a nucleophilic attack at the anomeric carbon, which results in formation of a glucose–enzyme intermediate. In this process, aglucone (R_1) departure is facilitated by protonation of the glucosidic oxygen by the acid catalyst (AH). During the second catalytic step (deglucosylation), a water molecule is activated by the catalytic base to serve as a nucleophile for hydrolysis of the glucosidic bond and release of the glucose (Davies and Henrissat, 1995; Morant *et al*., 2008).

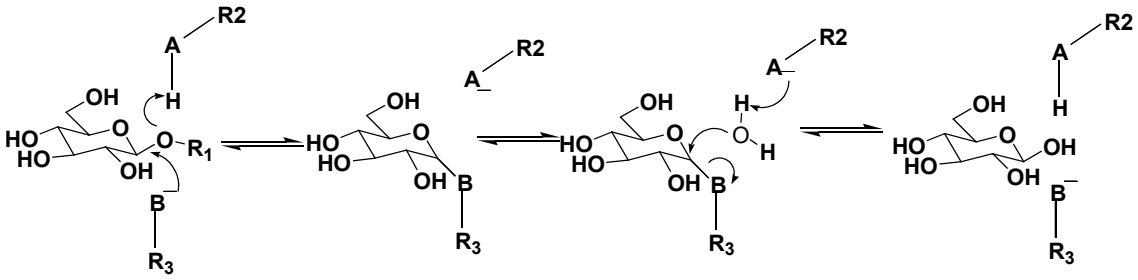


Fig.1.2. Reaction mechanism of retaining β-glucosidases. A glutamic acid residue in the conserved TFNEP motif (Davies and Henrissat, 1995) serves as a general acid/basecatalyst (AH) while a glutamic acid residue in the I/VTENG motif (Davies and Henrissat, 1995) serves as a nucleophile (B). The reaction cycle is specified in the text.

1.4. β-glucosidases for plant defense strategies

Plants have evolved various strategies to defend themselves against herbivores and pathogens. Although some of these strategies are constitutive, i.e. present at all times, others are induced only in response to herbivore feeding or pathogen infection (Frost *et al*,. 2008). There are wide variations in the composition and concentration of constitutive defenses and these ranges from mechanical defenses to digestibility reducers and toxins. Most external mechanical defenses and large quantitative defenses are constitutive, as they require large amounts of resources to produce and difficult to mobilize (Traw *et al*., 2002). Induced defenses include secondary metabolic products, as well as morphological and physiological changes (Karban *et al*., 1997).

Plants synthesize a plethora of secondary metabolites to defend themselves against herbivores and microorganisms and adapt to different types of abiotic environmental stresses. Traditionally, plant defense compounds are grouped into phytoanticipins and phytoalexins. Phytoanticipins are preformed defense compounds and are constitutive defenses whose production may increase under stress, whereas Phytoalexins are antimicrobial defense metabolites produced de novo in response to biotic and abiotic stresses (Vanetten *et al*., 1994). Plant secondary metabolites involved in plant defense include alkaloids, cyanogenic glycosides, glucosinolates, terpenoids, and phenolics (Wink *et al*., 2004).

An overwhelming number of these plant secondary products is reported to occur in glucosylated form (Warzecha *et al*., 2000). Glycosylation serves to protect the plant

against the toxic effects of its own chemical defense system, to increase solubility and facilitate storage (Jones and Vogt, 2001). Plant defense compounds such as hydrogen cyanide, saponins, coumarins, and naphthoquinones are released by deglycosylation of their substrates by specific β- glucosidases (Cairns *et al*., 2000).

1.5. β-Glucosidases as a bio-activating component in plant chemical defense

In intact plant tissue, the β-glucosidases and their defense compounds are compartmentalized separately. Upon cell disruption, caused for example by a chewing insect, the defense compounds are bioactivated via hydrolysis of the glucosidic linkage catalyzed by β-glucosidases. This two-component system i.e Glycoside/glycosidase (β-glucosidases), of which each of the individual components is chemically inert, provides plants with an immediate chemical defense against attacking herbivores and pathogens (Morant *et al*., 2008). The four most well characterized preformed two-component defense systems: cyanogenic glucosides, benzoxazinoid glucosides, avenacosides and glucosinolates, with emphasis on the β-glucosidases responsible for their bio-activation are reviewed (Morant *et al*., 2008).

1.5.1. Benzoxazinoid glucosides and their β-glucosidases

The benzoxazinoids (also referred to as hydroxamic acids) DIBOA (2,4-dihydroxy-1,4-benzoxazin-3-one) and DIMBOA (2,4- dihydroxy-7-methoxy-1,4-benzoxazin-3-one (figure 1.3) are defense chemicals widely distributed within Poaceae (Niemeyer, 1988) and have also been found in a number of eudicotyledenous plant species (Baumeler *et al*., 2000). Benzoxazinoids have been implicated in resistance of maize against insects, fungi and bacteria (Niemeyer, 1988) and maize mutants deficient in DIMBOA biosynthesis are compromized with respect to their disease resistance (Frey *et al*., 1997). DIMBOA is highly effective in conferring resistance to phloem feeding aphids (Givovich *et al*., 1994; Givovich and Niemeyer, 1995; Cambier *et al*., 2001) as evidenced by the failure to establish a population of *Metopolophium dirhodum* aphids on young maize seedlings (Cambier *et al*., 2001). Besides their importance as natural pesticides, benzoxazinoids and their degradation products show significant allelopathic

effects (Burgos and Talbert, 2000). Hence from an agricultural point of view, benzoxazinoid synthesizing crops are interesting not only because of their strong resistance to herbivores and pathogens, but also for their potential as cover crops to control weeds and in crop rotation schemes to reduce nematode infestation (Morant *et al*., 2008).

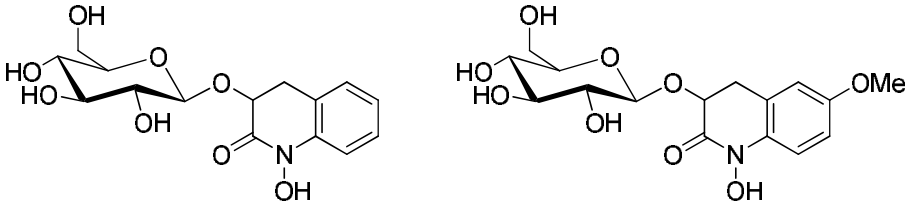


Fig. 1.3. Benzoxazinoid glucosides.

1.5.2. The oat avenacoside/avenacosidase two-component system

Oat (*Avena sativa*) is one cereal that does not accumulate Benzoxazinoid glucosides (BxGlcs). Instead, oat accumulates saponins which are wide spread defense compounds found in many plant species, although oat represents the only saponin accumulating cereal (Osbourn, 2003). Saponins possess detergent-like properties (Osbourn, 1996) and have been proposed to exert their antibiotic effect by insertion and complex formation with sterols in cellular membranes leading to formation of pores and membrane disruption (Morrissey and Osbourn, 1999). Saponins are mainly characterized as antimicrobial although insecticidal effects have also been reported (Weissenberg *et al*., 1998). How plants are able to store active saponins and prevent them from exerting their membrane disrupting effects in planta is unknown. Oat accumulates two forms of saponins. Roots contain avenasides, which are triterpenoids stored directly in their active, monodesmosidic form. Avenaside deficient *sad* mutants show significantly increased disease susceptibility (Papadopoulou *et al*., 1999), which substantiates the classification of plant saponins as defense compounds. Oat leaves accumulate non-active bisdesmosidic saponins, avenacosides (Fig 1.4), that need to be bio-activated by a specific β-glucosidase in order to exert their biocidal effects (Nisius, 1988). Only little is

known about the genes and enzymes involved in biosynthesis of the structurally complex saponins.

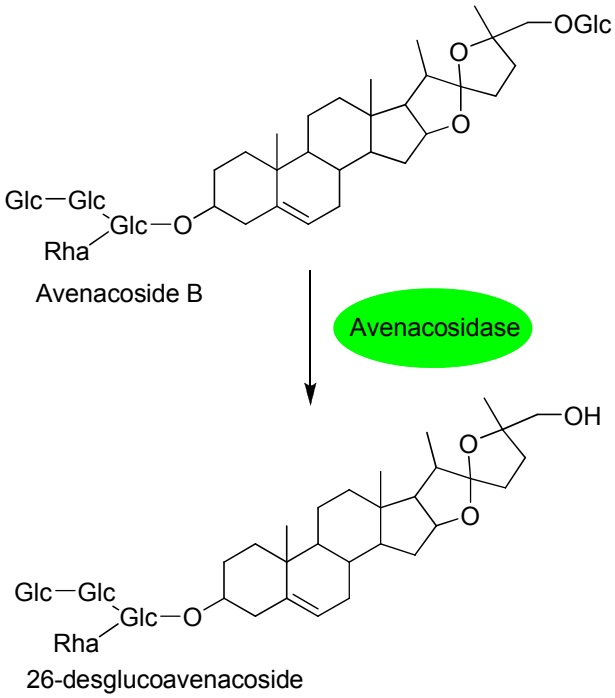


Fig. 1.4. Bio-activation of avenacoside B by hydrolysis of the C26 b-glucosidic bond. Upon tissue disruption in oat (shown in the picture) leaves, avenacosides A and B (molecules shown) are deglucosylated at the C26 position by avenacosidase to yield biologically active 26-desglucoavenacosides.

1.5.3. The glucosinolate–myrosinase system

Glucosinolates are a category of secondary products present primarily in the species of the order Capparales. When tissue is damaged, for example by herbivory, glucosinolates are degraded in a reaction catalyzed by thioglucosidases, denoted myrosinases, also present in these species. Thereby, toxic compounds such as nitriles, isothiocyanates, epithionitriles and thiocyanates are released (Fig 1.5). The glucosinolate/myrosinase system is generally believed to be part of the plant's defense against insects, and possibly also against pathogens (Rask *et al*., 2000).

Glucosinolates have been studied most extensively in the Brassicaceae family which includes many oilseed and vegetable crops. The distinctive flavor of members of this

family, such as *Brassica nigra* (black mustard), *Armoracia lapathifolia* (horseradish), and *Brassica oleracea* (cauliflower, broccoli, cabbage, brussels sprouts), is dependent on the hydrolysis of glucosinolates and on the nature and amounts of the products thus released (Fenwick *et al*., 1983).

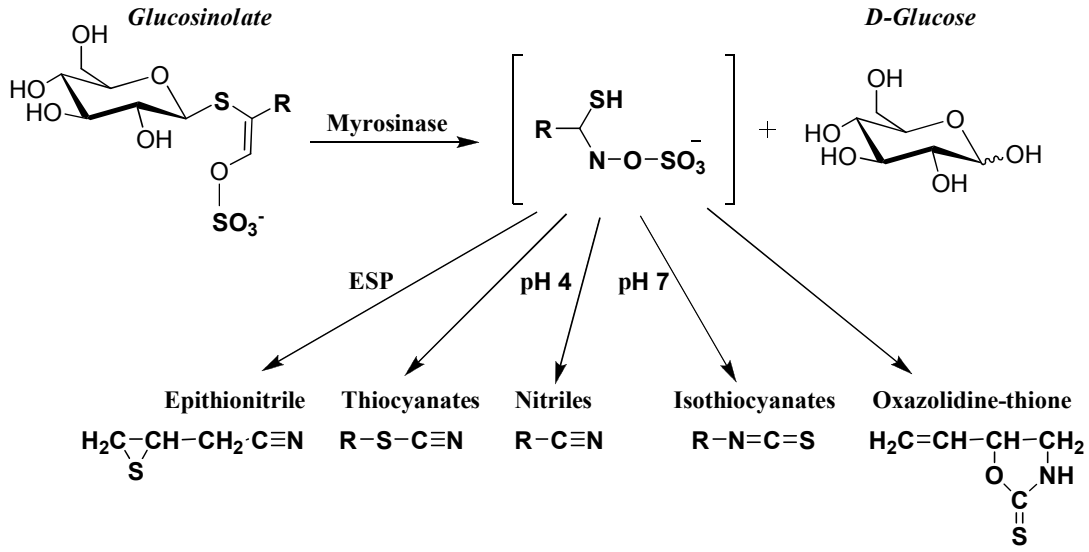


Fig.1.5. The general structure of glucosinolates, their intermediate and final degradation products. Conditions favoring the formation of certain degradation products are indicated. ESP (epithio specifier protein).

The glucosinolates play important roles in agriculture owing to the toxic effects of their breakdown products on most insects and herbivores (Lambrix *et al*., 2001; Kliebenstein *et al*., 2002; Agrawal and Kurashige, 2003; Lazzeri *et al*.., 2004) and due to their use as natural biofumigants (Brown and Morra, 1995; Lazzeri *et al*., 2004; Zasada and Ferris, 2004). Glucosinolates also serve as attractants and feeding stimuli for some, often highly specialized herbivores (Gabrys and Tjallingii, 2002; Mewis *et al*., 2002). As observed for specialist herbivores feeding on cyanogenic plants, some specialized insects are able to sequester glucosinolates present in their food plants for use in their own defense against predators (Müller *et al*., 2001; Müller *et al*., 2002).

1.5.4. Cyanogenic glucosides and cyanogenic β-glucosidases

Cyanogenic glucosides (CNGs) are amino acid derived phytoanticipins found in more than 2650 different plant species from ferns and gymnosperms to monocotyledenous and eudicotyledenous angiosperms (Bak *et al*., 2006). This suggests that cyanogenesis was an early evolutionary event (Saupe, 1981). Upon tissue disruption, the cyanogenic glucosides are hydrolyzed to yield an unstable aglucone, which either spontaneously or enzymatically degrades into a ketone or an aldehyde and toxic HCN (Poulton, 1990; Møller and Seigler, 1999; Morant *et al*., 2003). HCN blocks cellular respiration via inactivation of the mitochondrial cytochrome oxidase (Nelson, 2006).

Cyanogenic glucosides are present in a disproportionately large number of crops (Jones, 1998) suggesting a role of cyanogenic glucosides as important natural pesticides, i.e. a trait that has been favored during domestication of modern crop plants. The natural pesticide properties of cyanogenic glucosides are exploited by using cyanogenic plants for biofumigation purposes, e.g. in order to suppress nematode activities (Widmer, 2000; Widmer and Abawi, 2002).

1.5.4.1. The biosynthesis and bio-activation of cyanogenic glucosides

The biosynthesis and bio-activation of cyanogenic glucosides has been the focus of much research. Cyanogenic glucosides are synthesized from valine, isoleucine, leucine, phenylalanine, and tyrosine and from the non-protein amino acid 2-cyclopentenyl-glycine. The genes and enzymes of the entire cyanogenic glucoside biosynthetic pathway are known from sorghum, which produces the tyrosine-derived dhurrin at up to 30% dry weight in etiolated seedling tips (Halkier and Møller, 1989).

The Tyrosine is converted into the cyanogenic glucoside via the concerted action of two cytochromes P450, CYP79A1 and CYP71E1, and a UDPG-glucosyltransferase, UGT85B1 (Bak *et al*., 1998a; Jones *et al*. 1999; Hansen *et al*.., 2003; Thorsøe *et al*., 2005). Biosynthesis, catabolism and detoxification of CNGs in plants, insects and higher animals is shown in Fig 1.6. Due to its unique genetic simplicity, the cyanogenic glucoside pathway has a pioneering status in plant metabolic engineering (Bak *et al*., 1999; Kristensen *et al*., 2005; Morant *et al*., 2007), and the transfer of the entire dhurrin pathway to the non-cyanogenic model plant *Arabidopsis thaliana* unequivocally demonstrated the ability of cyanogenic glucosides to deter herbivores (Tattersall *et al*.,

2001). Cyanogenic β-glucosidases have been characterized from a wide variety of cyanogenic plants including sorghum, cassava, white clover, rubber tree (*Hevea brasiliensis*), black cherry (*Prunus serotina*), flax (*Linum ussitatissimum*) and *Lotus japonicus* (Fan and Conn, 1985; Kuroki and Poulton, 1986; Hosel *et al*., 1987; Selmar *et al*., 1987; Kuroki and Poulton, 1987; Pocsi *et al*., 1989; Mkpong *et al*., 1990; Morant *et al*., 2008)

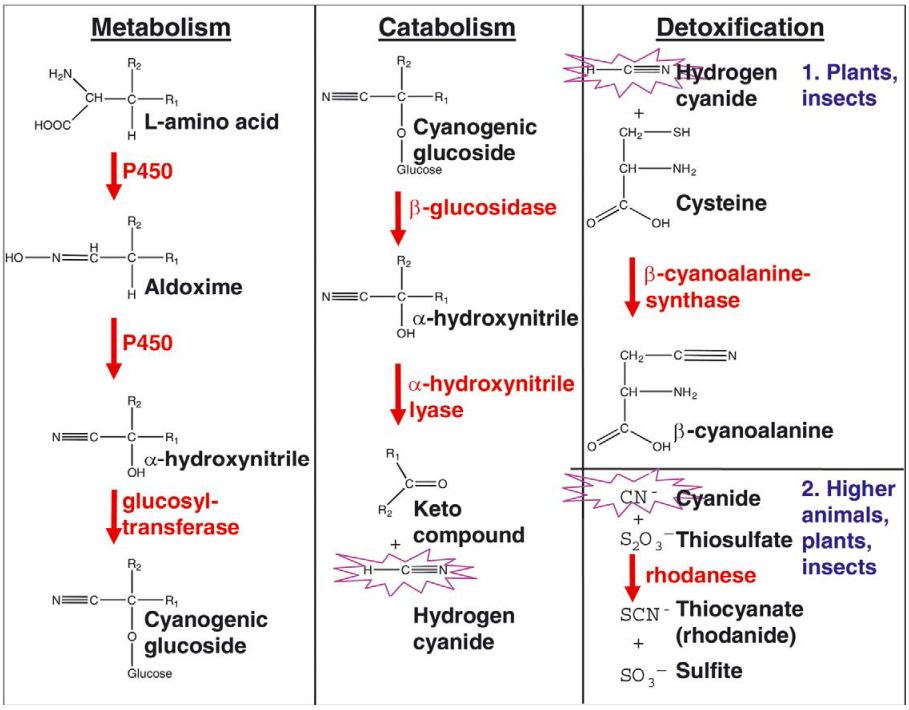


Fig.1.6. Biosynthesis, catabolism and detoxification of CNGs in plants, insects and higher animals. Enzymes involved are shown in red. HCN is highlighted in purple

1.5.4.2. Chemical structure of some cyanogenic glycosides

The structure of some cyanogenic glycosides and some examples of their occurrence are given in Table 1 (Source: Vetter *et al*., 2000). The CGs are glycosides of α-hydroxynitriles, all known compounds are β-linked, mostly with D-glucose.

Table 1
The chemical structure of some cyanogenic glycosides (according to Tapper and Reay, 1973)

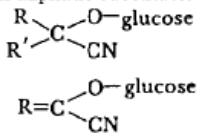
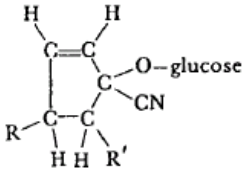
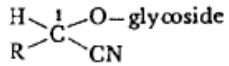
	Substituent	Glycoside	Sugar	Configuration at C1	Occurrence
(A) Glycosides with aliphatic substitutes					
	R═R′═CH_3-	linamarin	D-glucose		*Linum* spp. *Trifolium* spp.
	R═CH_3-, R′═CH_3CH_2-	lotaustralin	D-glucose		*Lotus* spp. *Maniholt* spp.
	R═$(CH_3)_2CH$	aciapetalin	D-Glucose		*Acacia* spp.
	R═$HCO_2CH═CH(CO_2HCH_2)C$	Triglochinin	D-glucose		*Triglochin* spp.
	R═H, R′═H	deidaclin	D-glucose		*Deidamia* spp.
		tetraphyllin A	D-glucose		*Tetrapathæa* spp.
	R═OH, R′═H	tetraphyllin B	D-Glucose		*Tetrapathæa* spp.
	R═OH, R′═H	Gynocardin	D-glucose		*Gynocardia* spp. *Pangium* spp.

Table 1 (*continued*)

	Substituent	Glycoside	Sugar	Configuration at C1	Occurrence
(B) Glycosides with aromatic substituents					
	phenyl	prunasin	D-glucose	D	*Prunus* spp.
	phenyl	amygdalin	gentiobiose	D	*Prunus* spp.
	phenyl	lucumin	primeverose	D	*Lucuma* spp.
	phenyl	vicianin	vicianose	D	*Vicia* spp.
	phenyl	sambunigrin	D-glucose	L	*Sambucus* spp.
	p-hydroxyphenyl	dhurrin	D-glucose	L	*Sorghum* spp.
	p-hydroxyphenyl	taxiphyllin	D-glucose	D	*Taxus* spp.
	p-Hydroxyphenyl	Zierin	D-Glucose	–	*Zieria* spp.
	p-Glucosyloxyphenyl	Proteacin	D-Glucose	L	*Macadamia* spp.
C. Glycosides with a free α-hydroxynitrile		*p*-Glucosyloxymandelonitrile		–	*Nandina* spp.

The enzymic hydrolysis produces the aglycone and the sugar moiety. The CGs can be grouped according to chemical nature of substituents, namely aliphatic, aromatic groups and into the glycosides with a free α- hydroxynitrile. Some of these CGs are better known than the others because the carrying plant species (group) has a greater practical importance. Several economically important plants are highly cyanogenic for example linamarin in Manihot esculenta, *Linum usitatissimum*, *Trifolium repens*, dhurrin in Sorghum species, amgydalin in rosaceous plants, lotaustralin in *Lotus corniculatus* (Vetter, 2000).

1.6. β- and γ- hydroxynitrile glucosides (Non-cyanogenic β-glucosides)

Among the plants harboring cyanogenic glucosides several also produce β- and γ-hydroxynitrile glucosides (Fig.1.7), which do not release HCN upon hydrolysis. Hence they are also known as non-cyanogenic glucosides. Because of the striking structural similarities of α-, β- and γ hydroxynitrile glucosides and a high frequency of co-occurrence it has been proposed that these compounds are biosynthetically related (Lechtenberg and Nahrstedt, 1999). Recent research has indeed established a biosynthetic connection between α-, β- and γ hydroxynitrile glucosides but the details remain unclear (Bjarnholt *et al*., 2008; Forslund *et al*., 2004; Morant *et al*., 2007).

1.6.1. Relationship between α, β and γ-hydroxynitrile glucoside biosynthesis

An investigation of the co-occurrence of isoleucine-derived α-, β- and γ -hydroxynitrile glucosides in Lotus, Ribes and Rhodiola species demonstrated that the presence of β- and γ hydroxynitrile glucosides was accompanied by presence of lotaustralin (α-hydroxynitrile) in all 13 species analyzed (Bjarnholt *et al*., 2008). A parallel result was obtained on the co-occurrence of the leucine-derived a-hydroxynitrile glucoside epi-heterodendrin and the β- and γ -hydroxynitrile glucosides C7–C10 (Fig. 1.7) in all eight Rosaceae species studied (Lechtenberg *et al*., 1996). Furthermore, *L. japonicas* transformed with an RNAi construct targeted against CYP79D3 and CYP79D4 displayed a transient decrease in the level of all hydroxynitrile glucosides (Morant *et al*., 2007). This means that the ile-ox (Isoleucine -oximes) produced by CYP79D3 and

CYP79D4 is most likely a shared intermediate in the biosynthesis of α-, β- and γ -hydroxynitrile Glucosides.

1.6.2. The proposed route for biosynthesis of β- and γ-hydroxynitrile glucosides

Nielsen *et al*. (2002) suggested that the biosynthesis of the non-cyanogenic β- and γ-hydroxynitrile glucosides reflects evolutionary diversification of the second P450 enzyme in the general cyanogenic glucosides biosynthetic pathway, such as CYP71E1 in *S. bicolor*, CYP71E7 in *M. esculenta* and their putative orthologs in other plant species. In the suggested pathway for synthesis of the β- and γ-hydroxynitrile glucosides the putative CYP71 α paralog, a multifunctional CYP71αβγ, is thought to be able to hydroxylate any of the carbon atoms in the nitrile intermediate (Nielsen *et al*., 2002). The putative α-, β- and γ-hydroxynitriles formed are subsequently either glucosylated to produce the saturated hydroxynitrile glucosides or undergo a C–C dehydration reaction followed by another C-hydroxylation to afford the unsaturated β- and γ-hydroxynitriles which can then be glucosylated. All known CYP79A1 orthologs have been shown to be highly substrate specific for the amino acid corresponding to the cyanogenic glucosides found in the plant in which the enzyme was identified (Andersen *et al*., 2000; Forslund *et al*., 2004; Kahn *et al*., 1999).

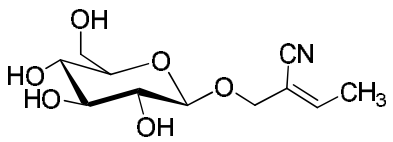
1.6.3. Evolution of cyanogenic glucoside biosynthetic enzymes

All known and putative CYP71a orthologs are more or less promiscuous enzymes able to produce α-hydroxynitriles from oximes and/or nitriles derived from several of the six known amino acid precursors and even some artificial nitriles. The *M. esculenta* CYP71E7 recombinant protein and microsomal preparations converted val-ox, ile-ox, tyr-ox and phe-ox into α-hydroxynitriles (Koch *et al*., 1992; Morant *et al*., 2008). This supports the notion that the aliphatic cyanogenic glucosides are evolved from the aromatic pathway.

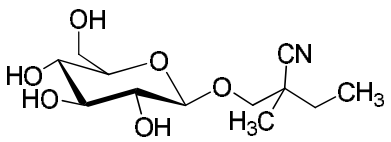
The qualitative and quantitative analyses of the hydroxynitrile glucoside distribution within different species of *Ribes*, *Rhodiola* and *Lotus* further support the proposed

connection between the biosynthetic pathways of cyanogenic and non-cyanogenic hydroxynitrile glucosides. Bjornjholt *et al*., (2008) found isoleucine-derived β- and γ-hydroxynitrile glucosides co-occurring with α-hydroxynitrile glucosides in all investigated *Ribes* and *Rhodiola* species and in two species of *Lotus* (Bjornjholt 2008). Recently, it has also been demonstrated that cyanogenic β-glucosidases can hydrolyze β- and γ-hydroxynitrile glucosides *in vivo* (Morant *et al*., 2008; Nielsen *et al*., 2006). β-Glucosides of β- and γ-hydroxynitriles derived from isoleucine are known as rhodiocyanosides (Fig.1.7) and always co-occur with the isoleucine-derived cyanogenic glucoside, lotaustralin (Forslund *et al*., 2004). The co-occurrence strongly suggests that rhodiocyanosides are synthesized via the same biosynthetic pathway as lotaustralin. This is supported by the fact that CYP79D3 and CYP79D4 both catalyze the first and rate limiting step in cyanogenic glucoside and rhodiocyanoside biosynthesis in *L. japonicus* (Forslund *et al*., 2004; Morant *et al*., 2007). In leaves of *L. japonicus*, the rhodiocyanosides are efficiently hydrolyzed upon tissue disruption by the cyanogenic β-glucosidases likewise responsible for cyanogenic glucoside hydrolysis (Morant *et al*., 2008). The co-localization of rhodiocyanosides and cyanogenic glucosides in *L. japonicus* (Forslund *et al*., 2004) and their parallel hydrolysis upon tissue disruption suggests that rhodiocyanosides like cyanogenic glucosides are defense compounds, although this remains to be shown experimentally. A parallel scenario is observed in barley (*Hordeum vulgare*) which accumulates the cyanogenic glucoside epiheterodendrin and four non-cyanogenic hydroxynitrilealk(en)yl glucosides, all derived from leucine (Nielsen *et al*., 2002). It is in agreement with the β-glucosides of α-, β- and γ-hydroxynitriles, have previously been proposed to be synthesized via the same biosynthetic pathway (Moller and Seigler, 1999; Nielsen *et al*., 2002).

<u>β-hydroxynitrile glucosides</u>



9d Rhodiocyanoside D



10d Epidermin

γ-hydroxynitrile glucosides

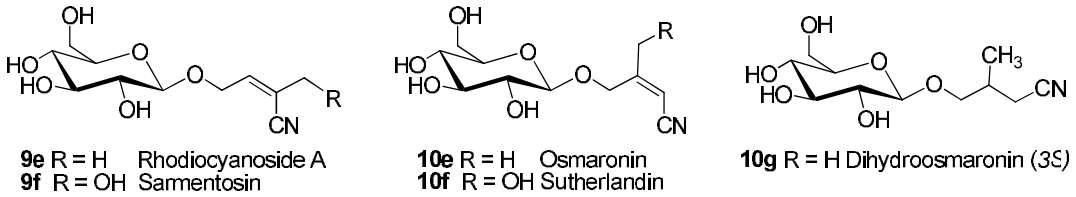


Fig.1.7. Selected structures of aliphatic hydroxynitrile glucosides. (Source: Nakamura *et al*., 2007).

1.7. Importance of *Leucaena leucocephala* in pulp and paper industry in India

Paper industry in India mainly uses bamboos, *Eucalyptus* sp., *Casuarina* sp. and *Leucaena* sp. as a source for paper pulp. Selection of the species depends upon availability, price and acceptability by any one given industrial unit. In bamboo growing countries, like India, the proportionate use of bamboos and hardwood species is in the ratio of 15:85. Although all these plant species are of importance to the paper industry, *Leucaena* sp. is exclusively used in India and about 25% of raw material for pulp and paper industry is contributed by this hard wood tree. To meet the increasing demand of high quality wood for paper industry it is essential to provide designer plant species. However, as a safeguard for the future no plant should be harvested from areas that may challenge sustainability. It will thus be crucial to raise plantations of the plant species with elite materials and or genetically modified plants that meet the demands of the pulp and the paper industry in economical and sustainable manner.

1.7.1. *Leucaena* classification

Leucaena is a native of Central America and it has been naturalized pan-tropically. Members of the genera are vigorous, drought tolerant, highly palatable, high yielding, rich in protein and grow in a wide range of soils (Hughes, 1998). However, these attributes are limited by the occurrence of anti-nutritive factors in the fodder, such as tannins and mimosine (Hammond *et al*., 1989 a, b). *Leucaena* occupies 2 to 5 million

hectares of land worldwide (Brewbaker and Sorenson, 1990). They are recognized as some of the fastest growing and most useful trees in the tropics (NRC, 1984). *Leucaena* is represented by 22 species. Of these, 6 are intraspecific taxa and 2 are widespread spontaneous hybrids. Most of the species are diploid 2n=52 or 56. However, 4 species are tetraploid 2n = 4X=104 or 112 (Hughes, 1998). *L. leucocephala* is a member of the genus related to the other species within the Mimosoideae sub-family, its subspecies and other related genera.

Classification of *Leucaena leucocephala*

Kingdom	Plantae
Super division	Tracheobionta
Division	Spermatophyta
Subdivision	Magnoliophyta
Class	Magnoliopsida
Subclass	Rosidae
Order	Fabales
Family	Fabaceae
Genus	Leucaena
Common name	Lead tree, white popinac
Sub species	Glabrata (Rose; S. Zarate); Ixtahuacana (Hughes) and *Leucocephala* (Benth) Var. Peru and Cunningham
Related genera	Desmanthus; Schleinitzia; Calliandropsis; Neptunia; Alantsilodendron; Gagnebina; Dichrostachys; and Kanaloa

1.7.2. Fodder production and feed value of *Leucaena leucocephala*

Leucaena leucocephala produces 6 to 18 tones of forage dry matter per hectare in the wet season and 2 to 3 tones of dry matter per hectare in the dry season. All parts of *Leucaena* are edible to livestock (Shelton, 1995). *Leucaena* leaves contain about 19 to 26 percent crude protein. They are rich in source of carotene and vitamins. The provitamin-A in *Leucaena* is among the highest ever recorded in a plant specimens (Jones, 1994). Leucaena is rich in calcium and phosphorus, but deficient in sodium and iodine. The in- vivo digestibility is 50 to 70 percent (Norton and Poppi, 1995). The

presence of mimosine, a tyrosine analogue, a toxic non protein amino acid limits the use of this plant as forage (Allison *et al*., 1990; Gupta and Atreja, 1999; Tangendjaja and Willis, 1980).

Fig. 1.8. *Leucaena leucocephala* plants at different stages. (a) *Leucaena leucocephala* plants in its full bloom, (b) *Leucaena* plants with immature pods, (c) *Leucaena* plants with mature pods and flowers.

1.7.3. *Leucaena* Wood

Leucaena wood has a thin bark which is about 8% dry matter at the age of 5 years. The sapwood is yellow-white, while the heartwood is yellow to reddish brown. This is similar to the density, tensile, compression, bending and shear strength of oak, ash, birch and sugar maple. It is fine textured and workable. It absorbs preservatives, and can be treated against termites (Pottinger and Huges, 1995). Leucaena wood is among the best hardwoods for the paper and rayon making. It produces pulp that is high in holocellulose, low in silica, ash, lignin, alcohol-benzene soluble and hot water soluble. Pulp yield is 50 to 52%. Its short fibre is suitable for rayan production (Pottinger and Hughes, 1995). Wood from giant Leucaena has a heating value of 4640 Kcal.kg-1 at the age of 2 to 4 years, and 7000 Kcal.kg-1 at the age of 8 years which is equivalent to 70% of the heating value of fossil fuel. In power generation, wood replaces fossil fuel in

generating electricity and for the production of charcoal for producer gas generators that power vehicles, boats and irrigation pumps.

1.8. Aphid infestation on *Leucaena leucocephala*

The rapid spread of the leucaena psyllid *Heteropsylla cubana* westward during the mid-1980s from tropical America to Asia, Australia and finally to Africa has restricted the utilisation of the important multipurpose tree legume *Leucaena leucocephala* (Geiger *et al*., 1995). Control options have included the use of insecticides considered uneconomic, and biological control using natural predators and genetic resistance (Napompeth *et al*., 1994). Some progress in biological control of psyllids using natural predators has been made and today psyllids rarely devastate stands of *L. leucocephala* to the extent that occurred in the mid 1980s (Napompeth *et al*., 1994). However, the impact of the psyllid remains important in most *L. leucocephala* based production systems. The effect on production appears to reduce as the plants mature, so that *Leucaena* managed for wood production is less affected than Leucaena managed for forage production (Austin *et al*., 1997). This is probably due to the higher proportion of mature green leaves not attacked by psyllids that maintain photosynthate to the plant even when young growth has been severely defoliated (Wheeler and Brewbaker, 1990). However, production losses can still be substantial. Genetic resistance is the most cost-effective method to overcome the psyllid problem (Mullen *et al*., 2003). Attempts to breed psyllid resistant Leucaena by crossbreeding resistant species with highly nutritive and productive species (such as Leucaena leucocephala) have been hindered because breeders do not know which particular plant traits are the primary influences of resistance to psyllids (Finlay-Doney and Walter, 2005).

1.9. Potential of biotechnology in improving plant with resistance to insect and pathogen attack

Insect-resistant crops have been one of the major successes of applying plant genetic engineering technology to agriculture; cotton (*Gossypium hirsutum*) resistant to lepidopteran larvae (caterpillars) and maize (*Zea mays*) resistant to both lepidopteran and coleopteran larvae (rootworms) have become widely used in global agriculture and

have led to reductions in pesticide usage and lower production costs (Toenniessen *et al*., 2003; Brookes and Barfoot, 2005; Gatehouse, 2008). These biotechnological approaches use successful expression of either the Bt toxins, lectins and protease inhibitors in plants. However the ability to activate plant natural defense pathways and to engineer inducible defense molecules may enhance the resistance durability of transgenic plants. Novel approaches for insect resistant plant includes engineering secondary metabolism plant defensive compounds and of volatile communication (Gatehouse, 2008).

1.10 Engineering Secondary Metabolism of Plant Defensive Compounds

The availability of genes encoding the biosynthetic enzymes of secondary metabolism has made transfer of biosynthetic pathways between plants feasible. Genes encoding two Cyt P450 oxidases and a UDP-glycosyltransferase from *Sorghum bicolor* have been transferred to *Arabidopsis* (Tattersall *et al*., 2001), resulting in the production of the cyanogenic glycoside dhurrin from Tyrosine (Kristensen *et al*., 2005). The resulting plants produced hydrogen cyanide on tissue damage and showed enhanced resistance to attack by the flea beetle, *Phyllotreta nemorum*, a specialist feeder on crucifers. Other secondary metabolites that have been produced in transgenic plants include the alkaloid, caffeine in tobacco by the introduction of three genes encoding N-methyl transferases (Kim *et al*., 2006).

1.11. The scope of the present study

As we have discussed earlier that different plant based defense systems exists in several plant species. It is necessary to understand plant based insect resistance system in *Leucaena* for development of insect resistance varieties. Such secondary metabolism can be biotechnologically manipulated to enhance the resistance against insects and pathogens.

1.12. Objectives of the present study

- To isolate, clone and characterize a β-glucosidase gene from *Leucaena leucocephala* probably involved in defense.
- To express the cDNA in a heterologous system and purification of the recombinant protein
- Biochemical and bioinformatics characterization of the recombinant enzyme
- To study the transcriptional expression profile of the β-glucosidase gene in different tissues of the plant

2. Materials and methods

This chapter deals with the general laboratory techniques routinely followed during the course of work. Other important specific methodologies followed will be discussed separately in the respective chapters.

2.1. Plant Material

2.1.1. *Leucaena leucocephala*

Studies on in vitro plant regeneration were carried out using seeds obtained from field grown L. *leuco Cephala* cultivar K-636. Seeds of *L. leuco Cephala* (K-636) were treated with conc. H_2SO_4 for 2 – 3 min and then washed extensively with tap water. The scarified seeds were surface sterilized with 0.1% (w/v) $HgCl_2$ for 10 min followed by five rinses with sterile deionized water. The seeds were soaked in sterile water for two days and then transferred to ½ MS basal medium (Murashige and Skoog, 1962) supplemented with 2% sucrose and 1.5% glucose for germination. The medium was solidified with 0.8% agar. The pH of the medium was adjusted to 5.8 prior to auto Claving. The culture bottles were incubated at 25± 2 ^{0}C and 60% relative humidity under 16 h photoperiod, light intensity 24.4 µmol / m^2/s. One month old axenic cultured pants were the raw material for further experiments.

2.2. Glassware

Glassware used in all the experiments were pro Cured from "Borosil", India. Test tubes (25 mm x 150 mm), glass bottles (70 mm x 125 mm), petridishes (85 mm x 15 mm; 55 mm x 15 mm), conical flasks (100, 250 & 500 mL; 1, 2 & 5 L capacity) and pipettes (1, 2, 5, 10 and 25 mL capacity) were used during the course of study.

2.2.1. Preparation of Glassware

Glassware used for all the experiments were cleaned by boiling in a saturated solution of Sodium bicarbonate for 1h followed by repeated washing in tap water. Thereafter, it was immersed in 30% HNO_3 solution for 30 min followed by repeated washing in tap water and rinsed with distilled water. Washed glassware was thereafter dried at room temperature. Test tubes and flasks were plugged with absorbent cotton (Mamta Surgical

Cotton Industries Ltd., Rajasthan, India). Auto Claving of the glassware and above items was done at 121 ^{0}C and 15 psi for 1 h.

2.3. Plastic ware

Sterile disposable filter sterilization units (0.22 μm) and petridishes (55 mm and 85 mm diameter) were pro Cured from "Laxbro", India. Eppendorf tubes (1.5 mL and 2 mL capacity), microtips (10, 200 and 1000 μL capacity) and PCR tubes (0.2 mL and 0.5 mL capacity) were obtained from "Tarsons" and "Axygen", India.

2.4. Chemicals

Tris, IPTG, X-gal, SDS, BSA, EDTA, PMSF, PIPES, DTT, PVPP, APS, Imidazole, TES buffer, Urea and Ethidium bromide were purchased from Sigma-Aldrich (USA), Bio-world (USA). Agarose, restriction enzymes, T4 DNA ligase, RNase A and lysozyme were obtained from NEB (USA), Promega (USA), Bioenzymes (USA) and Amersham (UK). Different kits were purchased from BD CLONETECH (JAPAN). Invitrogen (USA), Promega (USA), and Sigma-Aldrich (USA). Taq DNA polymerase was obtained from Sigma-Aldrich (USA) and Bangalore Genei (India). Plasmid vectors, pGEM-T Easy Vector and pET30b (+) were purchased from Invitrogen (USA), Promega (USA) and Novagen (USA) respectively. Megaprime labeling kit and Hybond-N+ membrane were obtained from Amersham (UK). [α- ^{32}P]-dATP and [α- ^{32}P] –dCTP were obtained from Bhabha Atomic Research Centre (BARC), India. X-ray films were obtained from Konica (Japan) or Kodak (USA). Agarose A (Sigma), Affi-gel matrices, gold particles were purchased from BioRad (USA). All other chemicals and solvents of analytical grade were purchased from HiMedia, Qualigens Fine Chemicals and E-Merck Laboratories, India. All chemicals used in the tissue culture study were of analytical grade (AR) and were obtained from "Qualigens", "S.D. Fine Chemicals" or "HiMedia", India. The Sucrose, glucose and agar-agar were obtained from "Hi- Media". Bacto-Agar for microbial work was obtained from "DIFCO" laboratories, USA. Substrates for enzyme assays i.e Glycosides both natural and synthetic were purchased from Sigma-Aldrich (USA). Almond β-glucosidase enzyme also obtained from Sigma-Aldrich (USA). Chromogenicsynthetic glucoside.VRA-Glc was purchasesd from PPR Diognostics, UK.

2.5. Equipment

Tab 2.1. Equipment

S.No.	Equipment	Make
1	Balances	Contech/ Sartorious
2	Water bath	Julabo/
3	Dry Bath	Eppendorf/BGenei
4	Incubator	New Brunswick
5	Centrifuge	Sorvall/Haereus/eppendorf/Sigma
6	Gel Do Cumentation system	Bio-Rad
7	Thermo Cycler PCR machine	MJResearch/Stratagene
8	Spectrophotometer	Applied Biosystem
9	Power pack	Bio-Rad
10	Agarose Gel Electrophoresis Units	Bangalore Genei/ Bio-Rad
11	protein Gel Electrophoresis Units	Hoeffer Scientific/ BioRad
12	Speed Vac concentrater	Savant
13	pH-Meter	Global
14	Water purification system	Millipore Unit (Milli RO/ Milli Q)
15	Microwave oven	Bilbol
16	Fridge/ Deep freezer	Vestfrost/Leonard/Godrej
17	Magnetic rotator	REMI
18	Laminar Air Flow	Microfilt India

2.6. Buffers And Solutions

2.6.1. Buffers and Solutions for DNA Electrophoresis

Tab. 2.2. Buffers And Solutions for DNA Electrophoresis

Name	Ingredients	Preparation and Storage
50x TAE	2 M Tris 0.05 M EDTA	pH was adjusted to 8.0 using glacial acetic acid and stored at RT
TBE buffer	90 mM Tris 90 mM Boric acid 2 mM EDTA	RT
DNA loading buffer	0.25 g Xylencyanol 0.25 g Bromophenol blue 0.25 g Ficoll 400 1.46 g EDTA make up the volume to 100ml with H_2O	The solutions were filtered sterilized using 0.22 µ filter and stored at RT

2.6.2. Buffers and Solutions for Genomic DNA Isolation, Southern and Slot Blot

Tab. 2.3. Buffers and Solutions for g-DNA isolation, Southern and Slot Blot

Name	Ingredients	Preparation and Storage
Extraction buffer	100 mM Tris-HCl (pH 8.0) 20 mM Na EDTA (pH 8.0) 1.4 M NaCl 2.0% (w/v) CTAB Add β-mercaptoethanol to 0.2 % before use.	RT
Depurination buffer	0.25 N HCL	Freshly prepared
Denaturation buffer	1.5 M NaCl 1M NaOH	RT
Neutralization buffer	1.5 M NaCl 1.0 M Tris HCl (pH 7.5)	RT
20 X SSC	3 M NaCl 0.3 M Sodium citrate (pH 7.0)	RT
Hybridization buffer	1% BSA 1.0 mM EDTA pH 8.0 0.5 M Sodium phosphate pH 7.2 7% SDS	RT
Low stringency wash buffer	6 X SSC 0.1% SDS	RT
Moderate stringency wash buffer	2 X SSC 0.1% SDS	RT
High stringency wash buffer	0.2 X SSC 1% SDS	RT

2.6.3. Solutions for the transformation and selection of *E. coli*

Tab. 2.4. Solutions for the transformation and selection of *E. coli*

Name	Ingredients	Preparation and Storage
IPTG solution	200 mg/mL in SMQ	Sterile filtration and storage at -20°C
X-Gal	20 mg/mL 5-bromo-4-chloro-3-indolyl-β-D-galactoside In N,N'-Dimethyl formamide (DMF)	Wrap in aluminum foil and store at -20°C
Ampicillin	100 mg/mL in SMQ	Sterile filtration and storage at -20°C
Tetracycline	12.5 mg/mL in 70% ethanol	Sterile filtration and storage at -20°C
Kanamycin	50 mg/mL in SMQ	Sterile filtration and storage at -20°C
Riffamycin	50 mg/mL in DMSO	Sterile filtration and storage at -20°C
Hygromycin	25 mg/mL in SMQ	Sterile filtration and storage at -20°C

2.6.4. Buffers and Solutions for Plasmid isolation (Alkaline lysis)

Tab. 2.5. Buffers and Solutions for Plasmid isolation (Alkaline lysis)

Name	Ingredients	Preparation and Storage
Solution I (TEG)	25 mM Tris-HCl (pH 8.0), 10 mM EDTA (pH 8.0), 50 mM Glucose.	Stored at 4°C
Solution II	0.2 N NaOH, 1% SDS	Freshly prepared
Solution III	3 M Potassium acetate (pH 4.8)	Stored at 4°C
RNase A	10 mg/mL	Stored at -20°C
Other solutions	Chloroform, Absolute ethanol, 3.0 M Sodium acetate 70% ethanol Deionized water	

2.6.5. Buffers and Solutions for Gel Electrophoresis (PAGE)

Tab. 2.6: Buffers and Solutions for Gel Electrophoresis (native/ SDS-PAGE)

Name	Ingredients	Preparation and Storage
Monomer solution	29.2% acrylamide 0.8% bis-acrylamide in water	Stored at 4°C in the dark
Stacking gel	Distilled water 3.4 mL 1 M Tris-HCl (pH 6.8) 0.63 mL Acrylamide/bis 30%, 0.83 mL 10% (w/v) SDS, 0.05 mL (SDS-PAGE) 10% (w/v) APS, 0.05 mL TEMED 5 µL	
Separating gel (12%)	Distilled Water 3.3 mL 1.5 M Tris-HCl (pH 8.8), 2.5 mL Acrylamide/Bis 30%, 4.0 mL 10% (w/v) SDS, 0.1 mL (SDS-PAGE) 10% (w/v) APS, 0.1 mL TEMED, 4 µL	
2x Protein loading buffer	Distilled Water, 2.7 mL 0.5 M Tris-HCl (pH 6.8), 1.0 mL Glycerin, 2.0 mL 10% (w/v) SDS, 3.3 mL (SDS-PAGE) β-Mercaptoethanol, 0.5 mL 0.5% (w/v) Bromophenol blue, 0.5 mL	Store at 4°C
10x SDS-electrode buffer	Tris base 15 g Glycine 72 g SDS 5 g Water up to 500 mL	Store at 4°C, dilute 1:10 before use
Staining solution	Coomassie-blue R 250, 25 mL Methanol, 100 mL Acetic acid, 20 mL Water up to 200 mL	
Coomassie blue	Coomassie-blue R 250, 0.5 g Water up to 50 mL	Dissolve the dye in water and filtre
Destaining solution	Methanol, 30 mL Acetic acid, 20 mL	

	Distilled water up to 200 mL	
Silver staining Fixing solution	30% ethanol, (150 mL) 10% acetic acid, (50 mL) Distilled water up to 500 mL	Store at room temperature
Sensitizing solution	30% ethanol, (150 mL) 0.5 M sodium acetate 0.2% $Na_2S_2O_3$ Distilled water up to 500 mL	Store at room temperature
Silver solution	0.2% silver nitrate (0.6 g) 0.01 % formaldehyde (30 µL) Distilled water up to (300 mL)	Prepare fresh
Developing solution	6% Na_2CO_3 (18 g) 0.02% formaldehyde (60 µL) Distilled water up to 300 mL	Prepare fresh
Stop solution	1.5% Na_2EDTA (4.5 g) Distilled water up to 300 ml	Store at room temperature

2.6.6. Buffers and Solutions for Protein Purification under Native Conditions

Tab 2.7. Buffers and Solutions for Protein Purification (based on imidazole concentration)

Name	Ingredients	Preparation and Storage
Binding buffer	50 mM Tris 150 mM NaCl 10 mM imidazole 5% glycerol (pH 8.0)	Adjust pH by adding concentrated HCl and Store at 4^oC
Wash buffer	50 mM Tris 300 mM NaCl 25 mM imidazole 10% glycerol (pH 8.0)	Adjust pH by adding concentrated HCl and Store at 4^oC
Elution buffer	50 mM Tris 300 mM NaCl 200 mM imidazole 10% glycerol (pH 8.0)	Adjust pH by adding concentrated HCl and Store at 4^oC

2.6.7. Buffers and Solutions for Protein Extraction under Denaturating Conditions

Tab 2.8. Buffers and Solutions for Protein Extraction under Denaturating conditions

Name	Ingredient	Preparation and Storage
Lysis buffer	50 mM Tris-HCl (pH 8.0) 5 mM EDTA 100 mM NaCl 0.5% Triton-X100 0.7 mM DTT	Stored at 4^oC

	0.1 mM PMSF (Freshly added) 10 mM Mg SO_4 Lysozyme100µg/mL (Added freshly)	
Sonication buffer	100 mM Tris HCl (pH8.0) 50 mM Glycine	Stored at 4°C
Dispersion buffer:	100 mM Tris-HCl (pH8.0) 50 mM Glycine 8 M Urea /6M GuHCl	Stored at 4°C

Tab 2.9: Buffers and Solutions for Protein Purification

Name	Ingredients	Preparation and Storage
Binding buffer	50 mM Tris 300 mM NaCl 5% Glycerol (pH 8.0)	Adjust pH by adding concentrated HCl and Store at 4°C
Wash buffer	100 mM-200 mM Citrate-Phosphate buffer 150 mM NaCl (pH 6.3)	Adjust pH by adding concentrated Citric acid or Na_2HPO_4 Store at 4°C
Elution buffer	100 mM-200 mM Citrate-Phosphate buffer 150 mM NaCl (pH 4.5)	Adjust pH by adding concentrated Citric acid or Na_2HPO_4 Store at 4°C

2.6.8. Buffers and Solution for the β-glucosidase Enzyme Assay

Tab 2.10. Buffers and Solutions for the β-glucosidase Enzyme Assay

Name	Ingredients	Preparation and Storage
Citrate-Phosphate buffer	100 mM Cirtic acid 200 mM Na_2HPO_4	Adjust pH by adding either concentrated citric acid or Na_2HPO_4 and Store at 4°C
Sodium Chloride	150 mM NaCl	RT

2.6.9. Different Media Used for Studies

Tab. 2.11. Different Media Used for Studies

Name	Ingredients	Preparation and storage
Luria Bertani Broth (LB)	1% Bactotryptone 0.5%Yeast extract 1% NaCl	pH adjusted to 7.0 with NaOH, store at room temperature or at +4°C
SOB media	2% Bactotryptone 0.5% Yeast extract 10 mM NaCl 10 mM $MgCl_2.6H_2O$ 2 mM KCl	pH adjusted to 6.8 with NaOH, store at room temperature or at +4°C
TB buffer	10 mM PIPES 15 mM CaCl2	pH was adjusted 6.8 with KOH. $MnCl_2$ was added to

	250 mM KCl	final concentration of 55 mM and filter sterilized
YEP	1% Bactotrypton 1% Yeast extract 0.1% Glucose	pH adjusted to 7.0 with NaOH, store at room temperature or at +4°C

2.7. Host Cells

Tab 2.12. Host Cells

E.coli	Genotype
DH 5α	F' _80_lacZ_M15 end A1 hsdR17 (rk-mk+) supE44 thi-1_-gyrA96 relA1 _(lacZYA-argFV169) deoR
JM 109	e14-(McrA-) recA1 endA gyrA96 th-1 hsdR17(rk-mk+) supE44 relA1 _(lac-proAB) [F' traD36 proAB lacqZ _M15
XL1 Blue	RecA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac [F´ proAB lacIqZΔM15 Tn10 (Tetr)
TOPO 10	F- mcrA Δ(mrr-hsdRMS-mcrBC) φ80lacZΔM15 ΔlacX74 recA1 araD139 Δ(araleu) 7697 galU galK rpsL (StrR) endA1 nupG
BL 21	F-, ompT hsdSB (rB – mB -) gal dcm (DE3) pLysS (CamR)

2.8. Methods

2.8.1. Bacterial culture conditions

E. coli was grown at 37°C with shaking at 200 rpm in Luria Bertani (LB) broth/ SOB broth and maintained on LB/ SOB plates with 1.5% agar (Sambrook *et al*., 1989). For plasmid DNA preparation recombinant E. coli was grown in LB media supplemented with appropriate antibiotic.

2.8.2. Bacterial transformation

2.8.2.1. Preparation of competent cells using $CaCl_2$

A single colony of *E.coli* XL1 Blue was ino Culated in 5 ml of LB medium containing tetracycline (12.5 mg / L). LB medium (50 mL) was ino Culated with 1% of the overnight grown E. coli culture and allowed to grow till 0.5 O.D. at 600 nm. The cells were harvested by centrifugation at 5,000 g for 10 min at 4°C, suspended in 100 mM ice-cold $CaCl_2$ and kept on ice for 30 min. Cells were centrifuged, the pellet suspended in 1mL of 100 mM ice-cold $CaCl_2$ and stored as aliquots of 200 µL at 4°C.

2.8.2.2. *E. coli* transformation

The competent *E. coli* cells were transformed according to Sambrook *et al* (1989). Briefly, DNA (~50 ng in 10 µL or less) was added to the (200 µL) competent E. coli cells, mixed and kept on ice for 30 min. The cells were then incubated at 42^{o}C for 2 min. To each tube 800 µL of LB broth was added and further incubated at 37^{o}C for 1 h. About 100 µL of the transformed competent cells were spread onto LB plates containing appropriate antibiotics, IPTG and X-gal as per need (Sambrook *et al*., 1989).

2.8.3. PCR Colony Screening

This method bypasses DNA purification, and relies on the selectivity of PCR amplification to determine whether a bacterial colony of interest does indeed contain the desired DNA. Simply adding a small portion of a bacterial colony to a PCR master mix will introduce enough templates DNA for amplification. A single bacterial colony was added to 1.5 mL eppendorf PCR tube containing 25 µL of water. The micro tip was agitated in the water to remove the colony. Subsequently, the 2 –3 µL of this suspension was put into 15 µL of PCR reaction. The remaining components were added to the PCR reaction and subjected to normal cycling parameters for the particular primers. If insert orientation, as well as presence, needs to be determined, utilization of a forward vector-specific primer and a reverse insert-specific primer, or vice versa, allows such determination. If only the presence of the insert needs to be determined, then two insert-specific primers can be used. An additional 5 min denaturation step at 95^{o}C before the amplification cycles will aid in lysing the bacteria to enhance PCR product amplification success. The resulting PCR products were checked on an agarose gel for the presence of the predicted band.

2.8.4. Isolation of Nucleic Acids And Polymerization Chain Reaction (PCR)

2.8.4.1. Isolation of plasmid DNA from *E. coli* cells

The alkaline lysis method of Sambrook *et al*., (1989) was improvised upon so that 12-24 samples could be pro Cessed conveniently for plasmid DNA extraction within 3 h, with yields of 5-30 µg per 1.5 mL culture depending on the host strain and the plasmid

vector. An important feature of this proto Col was the use of PEG for purification, which resulted in precipitation of high quality super-coiled plasmid DNA free of contamination. The bacterial cultures were grown overnight (O/N) with shaking (200 rpm) at 37^{o}C in LB broth, with appropriate antibiotic(s). About 1.5 to 3 mL culture was centrifuged for 1 min at 4000 g to pellet the bacterial cells. The pellet was resuspended in 100 µL of TEG buffer (Tab: 2.5) by vigorous pipetting, 200 µL of Soln. II (Tab: 2.5) was added, mixed by inversion till the solution becomes clear and incubated on ice for 5 min. The cell lysate was neutralized by addition of 150 µL of Soln. III (Tab: 2.5), mixed well and incubated on ice for 5 min. The cell debris was removed by centrifuging for 5 min at 12000 g at room temperature. The supernatant was transferred to a clean tube, RNase A was added to a final concentration of 20 µg/ mL (Sambrook *et al*., 1989) and incubated at 37^{o}C for 20 min. To the above solution 400 µL of chloroform was added, mixed for 30 s and centrifuged for 5 min at 12000 g at 4^{o}C. The upper aqueous layer was transferred to a clean tube, 1/10th volume sodium acetate and one volume absolute ethanol was added with mixing and kept at -20^{o}C for 1-2 h. The sample was centrifuged at 12000 g for 10 min at room temperature. The pellet was washed thrice with 70% ethanol and dried under vacuum. The dried pellet was dissolved in 40 µL of deionized water and 40 µL of PEG/NaCl solution (20% PEG 8000 in 2.5 M NaCl) was added. The mixture was incubated on ice for 20 min and the plasmid DNA pelleted out by centrifugation at 12000 g for 15 min at 4^{o}C. The supernatant was aspirated carefully, the pellet washed with 70% ethanol and air-dried. The dried pellet was resuspended in 20 µL deionized water and stored at -20^{o}C.

2.8.4.2. Isolation of plant Genomic DNA

Genomic DNA was isolated by using the proto Col of Lodhi *et al*. (1994). Fresh young leaves were collected, frozen in liquid nitrogen and crushed to a fine powder. About 1g of ground tissue was extracted with 10 mL extraction buffer. The slurry was poured into a clean auto Claved 50 mL centrifuge tube and 100 mg insoluble polyvinylpolypyrrolidone (PVPP) added. The tube was inverted several times to thoroughly mix the slurry, incubated at 60^{o}C for 30 min and then allowed to cool down to room temperature. 12 mL of chloroform: isoamylalcohol mix was added and the contents mixed by inverting the tube gently till an emulsion formed. The mix was then

centrifuged at 6,000 g for 15 min at room temperature. Supernatant was carefully collected in a fresh tube and chloroform: isoamylalcohol (24:1) extraction step repeated. To the clear supernatant 0.5 volume of 5 M NaCl was added and mixed gently. Next two volumes of cold (-20^{o}C) 95% ethanol was added and the sample kept at 4^{o}C until DNA strands appeared. The tube was centrifuged at 3,000 g for 3 min and then at 5,000 g for next 3 min. The supernatant was poured off, the DNA pellet washed with cold (4^{o}C) 70% ethanol and air-dried. DNA was dissolved in 400 µL of TE buffer. The DNA solution was treated with 1 µL RNase A (10 mg/ mL) per 100 µL DNA and incubated at 37^{o}C for 30 min. The sample was extracted with chloroform: isoamyl alcohol to remove RNAase A. DNA was re-precipitated and dissolved in 40 µL TE buffer. Purity of DNA was checked spectrophotometrically by measuring the absorbance ratio at 260/280 nm. DNA was stored at 4^{o}C.

2.8.4.3. Restriction digestion of DNA

Plasmid and genomic DNA restriction digestion was set up as per restriction enzymes manufacturer recommendations.

2.8.4.4. Extraction and purification of DNA from agarose gels

The restriction digested DNA or PCR amplified products were run on an agarose gel in 1X TAE buffer. The gel was stained with ethidium bromide (0.5 µg/ mL) and viewed using a hand held long wavelength UV illuminator. The fragment of interest was excised from the gel and weighed. A 100 µg gel slice was transferred to a 1.5 mL micro Centrifuge tube and 300 µL Buffer DE-A (Axygen^{TM} GEL elution kit, Biosciences, USA) added. The tube was incubated at 70^{o}C for 5 to 10 min with intermittent mixing until the gel slice was completely dissolved. The gel mixture was cooled down to room temperature and 150 µL of Buffer DE-B was added. The above molten agarose was put into Axyprep column and placed into 2 mL microfuge tube. The assembly was centrifuge at 12,000 g for 1 min and filtrate was discarded. 500 µL of wash buffer 1 (provided by Axygen) was added and centrifuged at 12,000 g for 30s, filtrate was discarded. 700 µL of wash buffer 2 was added and spin at 12,000 g for 30 s, filtrate was discarded. It was repeated again with buffer 2. 1 min empty spin was given to ensure the complete removal of salt. Axyprep column was transferred into a fresh 1.5 mL

microfuge tube and 25-30 µL of elution buffer was added to the centre of the membrane and kept it as such for 1 min at room temperature. Centrifuge at 12000 g for 1 min. The eluted DNA was stored at 4°C. This eluted PCR product was further used for further steps.

2.8.4.5. Total RNA Isolation

RNase free environment was created and maintained as described by Blumberg (1987). All glass and plastic ware was DEPC (0.1% in water) treated overnight and auto Claved. The pestle and mortar were also DEPC treated and then baked at 300 ºC for 6 h. All materials were dried in a vacuum oven. Total RNA from different plant tissues was isolated using TRIzol reagent. The plant tissue was collected, washed with DEPC treated water, frozen in liquid nitrogen and crushed to a fine powder. To 100 mg of the fine powder 1 mL TRIzol reagent was added and mixed thoroughly using a vortimix. Chloroform: isoamyl alcohol (300 µL) was added and mixed thoroughly using vortimix. The tubes were centrifuged at 4°C at 13,000 g for 15 min. The supernatant was transferred to 1.5 mL tubes and the chloroform: isoamyl alcohol step repeated. The aqueous phase was transferred to 1.5 mL tubes and half volume isopropanol added. It was mixed thoroughly and kept for RNA precipitation for 1 h at room temperature. Total RNA was pelleted out by centrifugation at 13,000 g for 15 min at 4°C. The RNA pellet was washed with 70% ethanol twice and dried in a SpeedVac centrifugal concentrator. RNA pellet was dissolved in 40 µL of DEPC treated water and stored at -80°C in aliquots. Purity of RNA was confirmed by measuring OD at 260/280 nm and also by visualization on 1.5% TAE Agarose gel.

2.8.4.6. mRNA Purification

Total RNA was quantified spectrophotometrically. The amount of RNA was in the range of 1 mg to 3 mg. Appropriate amount of OBB Buffer and Oligotex Suspension were added according to manual instruction. The sample was incubated for 3 min at 70°C in a heating blo Ck. Sample was removed from the heating blo Ck, and place at 20°C to 30°C for 10 min. Oligotex: mRNA complex was pelleted down by centrifugation for 2 min at maximum speed (14,000–18,000 g), and carefully removed the supernatant by pipetting. Oligotex: mRNA pellet was resuspended in the appropriate

amount of OW2 buffer by vortexing and pipetted onto a small spin column placed in a 1.5 mL micro Centrifuge tube and centrifuged for 1 min at maximum speed. Spin column was transferred to a new RNase-free 1.5 mL micro Centrifuge tube, appropriate amount of OW2 buffer added to the column and again centrifuged for 1 min at maximum speed, flow-through was discarded. Spin column was transferred to a new RNase free 1.5 mL micro Centrifuge tube. Appropriate amount of hot (70^{o}C) OEB buffer was pipetted onto the column and resuspended by pipetting up and down; spin it down for 1 min at maximum speed. The Eluted mRNA sample was quantified spectrophotometrically and used for further downstream pro Cesses.

2.8.4.7. Spectrophotometric determination of nucleic acid concentration

DNA concentration was determined by measurement of the absorption at 260 nm. A Lambda 25 Perkin Elmer Photometer was used to determine the concentration of 1:50 diluted RNA or DNA samples in a volume of 1ml in a 10 mm light path quartz cuvette. Absorbance readings (A_{260}) should fall between 0.1 and 1.0 to be accurate. Sample dilution was adjusted accordingly. An absorbance of 1.0 at 260 nm corresponds to 50 µg DNA/ mL.

2.8.4.8. cDNA first strand synthesis by Reverse Transcription

Complementary DNA (cDNA) is synthesized from a mature mRNA template in a reaction catalyzed by the enzyme reverse transcriptase. The resulting molecule is a DNA-RNA hybrid and the pro Cess is called as cDNA 1st strand synthesis. For DNA double strand synthesis this hybrid molecule is digested with RNase H (specific for degrading RNA strand in a DNA- RNA hybrid), DNA second strand is synthesized using DNA polymerase I (Kimmel and Berger, 1987). In the present study cDNA first strand was synthesized using ImPromII™ Reverse Transcription System (Promega, USA). The reactions were set up as per the manufacturer's guidelines. Briefly, reverse transcription reactions of up to 1 µg of total RNA performed in 20 µL reactions comprised of components of the ImPromII Reverse Transcription System. Experimental RNA was combined with the oligo (dT) $_{15}$ primer. The primer/template mix was isothermally denatured at 70^{o}C for 5 min and snap chilled on ice. A reverse transcription reaction mix was assembled on ice to contain nuclease-free water, reaction buffer,

reverse transcriptase, magnesium chloride, dNTPs and ribonuclease inhibitor RNasin®. As a final step, the template-primer combination was added to the reaction mix on ice. Following an initial annealing at 25°C for 5 min, the reaction was incubated at 42°C for up to 1 h. The cDNA synthesized was directly added to amplification reactions. The first strand reaction was set up as follows:

Reagent	Volume
Experimental RNA (1µg)	1 µL
Primer [Oligo(dT)$_{15}$ or Random (10 pmol)	1 µL
DEPC treated Water	3 µL
Final volume	**5 µL**

The tubes were incubated at 70°C for 5 min and then chilled in ice-water for 5 min. Tubes were briefly spun in a micro Centrifuge to collect the condensate and maintain the original volume. The tubes were kept closed and on ice until addition of the reverse transcription reaction mix. The reverse transcription reaction mix was prepared by combining the following components of the ImProm-II Reverse Transcription System in a sterile 1.5 mL micro Centrifuge tube on ice.

Reagent	Volume
ImProm-II. 5X Reaction Buffer	4.0 µL
MgCl2 (15 mM)	2.0 µL
dNTP Mix (7.5 mM)	1.0 µL
RNasin® Ribonuclease Inhibitor (40U/ µl)	0.5 µL
ImProm-II Reverse Transcriptase	1.0 µL
Nuclease-free water	6.5 µL
Final volume	**15.0 µL**

An aliquot of 1.0 µg total RNA and oligo (dT)$_{15}$ or Random hexamer primer (10 pmol) mix total vol 5 µl was added to the above reaction for a final reaction volume of 20 µL per tube. The tube was incubated at 25°C for 5 min for primer annealing and then at

42°C for 1 h for cDNA first strand synthesis. Reverse transcriptase was thermally inactivated by incubation at 70°C for 15 min prior to pro Ceeding with PCR amplification (Chumakov, 1994).

2.8.4.9. Polymerase Chain Reaction (PCR)

PCR is a powerful technique to amplify a desired nucleotide sequence using sequence specific primers. This amplification may be either of and from a single template or of a template from a mixture of templates (Saiki *et al*., 1985; Mullis and Faloona, 1987; Saiki *et al*., 1988; Mullis, 1990; Arnheim and Erlich, 1992). This technique has been successfully used for various purposes like fishing out of gene(s) from genomic DNA or from cDNA population (Todd *et al*., 1987), introducing restriction sites of interest in the amplified product for directional cloning (Scharf *et al*., 1986), creating sequence mismatch/ deletion/ addition resulting in mutant version of a gene or nucleotide sequence (Goodenow *et al*., 1989), differentiating between two alleles (Kwok *et al*., 1990) etc. In the present study applications of PCR were exploited for a few of the above specified applications. The PCR reaction mixture and cycling conditions used were as follows:

Reaction mixture

Reagent	**Volume**
Sterile deionized water	6.2 µL
Template (50 ng/ µL)	1.0 µL
Forward primer (6 pmol)	1.0 µL
Reverse primer (6 pmol)	1.0 µL
dNTPs (0.2 mM)	4.0 µL
10 x Buffer (Mg^{+2} 1.5 mM)	1.5 µL
Taq DNA Polymerase (1U/µL)	0.3 µL
Total volume	**15.0** µL

PCR cycle conditions

1 cycle 95°C 5 min

35 cycles 95°C 1 min / 45-65°C 30-45 s (annealing temperature was dependent on

primer Tm) / 72^{o}C 1 min 30 s

1 cycle 72^{o}C 10 min

1 cycle 4^{o}C hold.

2.8.4.10. Rapid amplification of cDNA ends (RACE)

Generally, using reverse transcription PCR, either partial cDNA fragments (both 5' and 3' ends missing) or cDNA with full 5'end missing are amplified from total cDNA. If a partial cDNA sequence is known, unknown sequences to the 5' and 3' of the known sequence can be reverse transcribed from RNA, amplified by PCR using RACE (Frohmman *et al*., 1988). RACE Ready cDNA Kit (Invitrogen, USA) was used for RACE. The reactions were set up as per the manufacturer's guidelines. Briefly, the RACE technique is based on oligo-capping and RNA ligase-mediated (RLM) RACE methods (Maruyama and Sugano, 1994; Vollo Ch *et al*., 1994). The GeneRacer method involves selectively ligating an RNA oligonucleotide (GeneRacer RNA Oligo) to the full-length 5' ends of decapped mRNA using T4 RNA ligase. Application of GeneRacer race ready cDNA Kit is as follows:

- Identifying the 5' and 3' untranslated regions of genes
- Studying heterogeneous transcriptional start sites
- Characterizing promoter regions
- Obtaining the complete cDNA sequence of a gene
- Amplifying the full-length cDNA for downstream cloning and expression

The prerequisite to begin the RACE was to have the good quality RNA or mRNA. 1-5 µg total RNA or 50- 250 ng purified mRNA in DEPC water was used for the RACE reaction**.** Dephosphorylation reaction of 10 µL was set in 1.5 mL sterile micro centrifuge tube. The reaction mix is as follows.

Reagent	**Volume**
RNA	2 µL
10X CIAP Buffer	1 µL
RNaseOut™ (40 U/µL)	1 µL
CIP (10 U/µL)	1 µL
DEPC water	5 µL
Total Volume	**10 µL**

The reaction mixture was mixed gently by pipetting, followed by brief spin to collect the liquid at the bottom and incubated at 50^{o}C for 1 h. The reaction volume was mad eup to 100 µL by adding DEPC water and precipitated by adding 2 µL of 10 mg/mL glycogen, 10 µL of 3 M sodium acetate (pH 5.2) and 220 µL 95% ethanol. The precipitated sample was washed with 70% ethanol and pellet was air-dried. Air-dried pellet was dissolved in 7 µL of DEPC water. 5' mRNA cap structure from full- length mRNA was removed by following reaction:

Reagent	**Volume**
Dephosphorylated RNA	7 µL
10X TAP Buffer	1 µL
RNaseOut™ (40 U/µL)	1 µL
TAP (0.5 U/µL)	1 µL
Total Volume	**10 µL**

The reaction mixture was mixed gently by pipetting, followed by brief spin to collect the liquid at the bottom and incubated at 37^{o}C for 1 h. The reaction volume was made up to 100 µL by adding DEPC water and precipitated by adding 2 µL of 10 mg/mL glycogen, 10 µL of 3 M sodium acetate (pH 5.2) and 220 µL 95% ethanol. The precipitated sample was washed with 70% ethanol and pellet was air-dried. Air-dried pellet was dissolved in 7 µL of DEPC water. After decapping of full-length mRNA the RNA Oligo was ligated to the 5' end of mRNA. The reaction of ligating the RNA Oligo to decapped mRNA is as follows:-

7 µL of dephosphorylated, decapped RNA was added to the tube containing the pre-aliquoted, lyophilized GeneRacer™ RNA Oligo (0.25 µg). RNA Oligo was resuspended by pipetting up and down several times. Mixture was incubated at 65 °C for 5 minutes and kept on ice

Reagent	**Volume**
10 X Ligase Buffer	1 µL
10 mM ATP	1 µL
RNaseOut™ (40 U/ µL)	1 µL
T4 RNA ligase (5U/µL)	1 µL
Total Volume	**11 µL**

The reaction mixture was mixed gently by pipetting, followed by brief spin to collect the liquid at the bottom and incubated at 37°C for 1 h. The reaction volume was made up to 100 µL by adding DEPC water and precipitated by adding 2 µL of 10 mg /mL glycogen, 10 µL of 3 M sodium acetate (pH 5.2) and 220 µL 95% ethanol. The precipitated sample was washed with 70% ethanol and pellet was air-dried. Air-dried pellet was dissolved in 10 µL of DEPC water. After ligation of GeneRacer™ RNA Oligo to decapped, full-length mRNA, the mixture was ready to reverse transcribe the mRNA into cDNA. 1 µL of the desired primer and 1 µL of dNTP Mix (25 mM each) were added to the ligated RNA and incubated at 65°C for 5 min. Following reagents were added to the 11 µl ligated RNA and primer

Reagent	Volume
5 X RT Buffer	4 µL
SuperScript™ III RT (200 U/µL)	1 µL
0.1 M DTT	1 µL
RNaseOut™ (40 U/µL)	1 µL
Sterile water	2 µL
Total Volume	**20 µL**

The reaction mixture was mixed gently by pipetting, followed by brief spin to collect the liquid at the bottom and incubated at 45°C for 1 h. The reaction was inactivated by incubating the reaction mixture at 70°C for 15 min. 1 µL of RNase H (2 U) was added to reaction mixture to chew up the DNA / RNA hybrid by incubation at 37°C for 20 min. Centrifuged the sample and kept at -20°C for further use in PCR.

PCR Setup

Reactions were setup to amplify either the 5′ end or the 3′ end of gene of interest. The reaction was set up as given in Table below:

Reagent	5′ RACE	3′ RACE
GeneRacer 5′ Primer10 µM	3 µL	-
Reverse GSP 10 µM	1 µL	-

GeneRacer™ 3′ Primer 10 µM	-	3 µL
Forward GSP 10 µM	-	1.0 µL
RT Template	1.0 µL	1.0 µL
10X High Fidelity PCR Buffer	5.0 µL	5.0 µL
dNTP Solution (10 mM each)	1.0 µL	1.0 µL
Platinum® Taq DNA Polymerase High Fidelity, 5U/ µL	0.5 µL	0.5 µL
$MgSO_4$, 50 mM	2.0 µL	2.0 µL
Sterile Water	36.5 µL	36.5 µL
Total Volume	**50.0 µL**	**50.0 µL**

Cycling

Temperature	Time	Cycles
94°C	2 min	1
94°C	30 S	5
72°C 1 min/ 1 kb	2 min	
94°C	30 S	5
70°C 1 min/ 1 kb	2 min	
94°C	30 S	25
66°C	30 S	
72°C	2 min	
72°C	10 min	1

Nested PCR:

Nested PCR was done to increase the specificity and sensitivity of RACE products for the 5′ or 3′ ends of gene. 1 µL of the original amplification reaction was used as template for nested PCR. Reactions were set up as described below:

Reagent	5′ RACE	3′ RACE
Gene Racer 5′Nested Primer10 µM	1.0 µL	-
Reverse Nested GSP 10 µM	1.0 µL	-
Gene Racer 3′ Nested Primer 10 µM	-	1.0 µL
Forward Nested GSP 10 µM	-	1.0 µL
Primary PCR Product	1.0 µL	1.0 µL

10 X High Fidelity PCR Buffer	5.0 µL	5.0 µL
dNTP Solution (10 mM each)	1.0 µL	1.0 µL
Platinum® Taq DNA Polymerase High Fidelity, 5U/µL	0.5 µL	0.5 µL
$MgSO_4$, 50 mM	2.0 µL	2.0 µL
Sterile Water	38.5 µL	38.5 µL
Total Volume	**50.0 µL**	**50.0 µL**

Following program was used for the nested PCR reactions.

Cycling

Temperature	Time	Cycles
94°C	2 min	1
94°C	30 S	25
66°C	2 min	
72°C	2 min	
72°C	10 min	1

10-20 µL of nested PCR product was analysed on a 1% agarose/ethidium bromide gel. Nested PCR products were shorter by the number of bases between the original primers and the nested primers. The largest product was representing the most full-length message. The band(s) were excised, cloned and sequenced.

2.8.4.11. Quantitative real time PCR (QRT PCR)

Quantitative Real Time PCR is a powerful tool for gene expression analysis and was first demonstrated by Higuchi *et al*., (1992, 1993). QPCR quantitates the initial amount of the template most specifically, sensitively and reproducibly, and is a preferable alternative to other forms of quantitative reverse transcriptase PCR that detect the amount of final amplified product at the end-point (Freeman *et al*., 1999; Raeymaekers, 2000). QPCR monitors the fluorescence emitted during the reaction as an indicator of amplicon production during each PCR cycle (i.e., in real time) as opposed to the endpoint detection (Higuchi *et al*., 1992, 1993). The QPCR system is based on the detection and quantitation of a fluorescent reporter (Lee *et al*., 1993; Livak *et al*., 1995). There are three main fluorescence-monitoring systems for DNA amplification (Wittwer

et al., 1997a): (1) hydrolysis probes, (2) hybridizing probes and (3) DNA-binding agents (Wittwer *et al*., 1997b, Vander Velden *et al*., 2003). Most commonly used are the hydrolysis probes, which include TaqMan probes (Heid *et al*., 1996), molecular beacons (Tan *et al*., 2004; Vet and Marras, 2005) and scorpions (Saha *et al*., 2001; Terry *et al*., 2002). They use the 5' exonuclease activity of Taq polymerase to measure the amount of target sequences in cDNA samples. The use of fluorescent probe technologies reduces the risk of sample contamination while maintaining convenience, speed and high throughput screening capabilities (Reischl *et al*., 2002).

In the present study the Brilliant® II QPCR Master Mix (Stratagene, USA) was used. This kit supports quantitative amplification and detection with multiplex capability, and shows consistent high performance with various fluorescent detection systems, including molecular beacons and TaqMan® probes. The kit supports PCR amplifications and detection of a variety of DNA targets, including genomic DNA, plasmid DNA, and cDNA. The Brilliant II QPCR master mix includes SureStart® Taq DNA polymerase, a modified version of Taq2000™ DNA polymerase with hot start capability. A passive reference dye (an optional reaction component) is provided so as to control the final dye concentration, increasing the flexibility of the reagents for use with different platforms.

2.8.4.11.1. Pre-proto Col considerations

Magnesium Chloride Concentration

Magnesium chloride concentration in the PCR reaction mix affects the specificity of the PCR primers and probe hybridization (Eckert and Kunkel, 1990). The Brilliant® II QPCR Master Mix kit contains $MgCl_2$ at a concentration of 5.5 mM (in the 1 x solution), which is suitable for most targets.

Probe Design

Probes are designed in a fashion so as to have a melting temperature 7–10^{o}C higher than the annealing temperature of the primers. Lyophilized custom molecular beacon or TaqMan probes are constituted in 5 mM Tris-HCl (pH 8.0) and 0.1 mM EDTA (low TE buffer).

Fluorescence Detection

Fluorescence may be detected either in real-time or at the endpoint of cycling using a realtime spectrofluorometric thermal cycler. For endpoint analysis, PCR reactions can be run on any thermal cycler and then analyzed with a fluorescence plate reader that has been designed to accommodate PCR tubes and is optimized for the detection of fluorescent PCR reactions. Readings may be acquired before and after PCR for comparison.

Optimal Concentrations for Experimental Probes and Primers

The optimal concentration of the experimental probe is determined empirically. The optimal concentration is the lowest concentration that results in the lowest Ct and an adequate fluorescence for a given target concentration. In present study TaqMan® probes were used.

TaqMan® Probes

The TaqMan probe concentration can be optimized by varying the final concentration from 100 to 500 nM in increments of 100 nM.

PCR Primers

The optimal concentration of the upstream and downstream PCR primers is determined empirically. The optimal concentration is the lowest concentration that results in the lowest Ct and an adequate fluorescence for a given target concentration. The primer concentration for use with TaqMan probes can be optimized by varying the primer concentration from 50 nM to 600 nM. The best concentrations of the upstream and downstream primers are not always of equal molarity.

Reference Dye

A passive reference dye may be added to compensate for non-PCR related variations in fluorescence. Fluorescence from the passive reference dye does not change during the course of the PCR reaction but provides a stable baseline to which samples are normalized. In this way, the reference dye compensates for changes in fluorescence

between wells caused by slight volume differences in reaction tubes. The excitation and emission wavelengths of the reference dye are 584 nm and 612 nm respectively.

Reference Dye Dilution

Reference dye is prepared fresh prior to setting up the reactions. All reaction tubes containing reference dye are protected from light. Initial dilutions of the reference dye are made in nuclease-free PCR-grade H_2O and used at a final concentration of 30 – 300 nM depending upon the instrument configuration.

Data Acquisition With a Spectrofluorometric Thermal Cycler

Acquisition of real-time data generated by fluorogenic probes was as recommended by the instrument manufacturer. Data collection was either at the annealing step of each cycle (3- step cycling proto Col) or the annealing/extension step (2-step cycling proto Col).

2.8.4.11.2. Preparing the Reactions

The reference dye was diluted 1:500 (recommended in kit) using nuclease-free PCR-grade H_2O resulting in a final reference dye concentration of 30 nM in the reaction mixture. Real time PCR model Stratagene Mx 3000 Pi was used in present study. The experimental reactions were prepared by adding the following components in order:

Reagent Mixture

Component	Volume
2 X master mix	12.5µL
Experimental probe (optimized concentration)	x µL
Upstream primer (optimized concentration)	x µL
Downstream primer (optimized concentration)	x µL
Diluted reference dye	0.375 µL
Experimental gDNA, cDNA or plasmid	x µL
Final volume	**25 µL**

2.8.4.11.3. PCR Cycling Programs

The reactions were placed in the instrument and a 2 step program was run (recommended for TaqMan reactions).

Two-step Cycling Proto Col

Temperature	Time	No. of Cycles
95°C	10 min	1
95°C	30 S	40
60°C	30 S	
72°C	1 min	

Initial 10 min incubation was to fully activate the DNA polymerase. The temperature cycler was set to detect and report fluorescence during the annealing/extension step of each cycle.

2.8.5. Nucleic Acids Blotting/Hybridization

2.8.5.1. Slot Blot Hybridization

For slot blot hybridization DNA or RNA samples were diluted according to experimental requirements. The DNA samples were denatured by adding 1/10th volume 3 M NaOH and incubation at 65°C for 10 min. To the denatured sample an equal volume of 6 X SSC was added. Two layers of Whatman 3 MM filter paper wetted with sterile deionized water and 6 X SSC were placed in the Slot Blot apparatus followed by Hybond-N+ membrane (Amersham, UK) treated in the manner as above. The Slot Blot unit was assembled and wells washed with 500 µL of 6 X SSC by applying vacuum. After washing, samples prepared earlier were applied in the wells and vacuum applied till whole sample volume passed through the well slit and wells appear dry. The unit was carefully disassembled and the membrane taken out. The membrane was air dried and then baked for 2 h at 80°C to immobilize DNA. Hybridization and autoradiography were carried out as described in the following section 2.12.3 & 2.13.

2.8.5.2. Random Primer Labeling

Random primer labeling of the DNA probes (Feinbeng and Vogelstein, 1983, 1984) was done using the Megaprime DNA labeling kit (Amersham, UK). Reaction (50 µL) was set up as follows:

Component	**Volume**
25 ng DNA (used as probe)	5.0 µL
Primer solution (Random hexanucleotides) (3.5 A_{260} U)	5.0 µL
Final Volume	**10.0 µL**

Above mixture was heated in a boiling water bath for 10 min and cooled to room temperature facilitating primer annealing to the DNA. This reaction mixture was added to the following reaction mixture.

Component	**Volume**
10 X reaction buffer (500 mM Tris-HCl, pH 8.0; 100 mM $MgCl_2$; 10 mM DTT; 0.5 mg/ mL acetylated BSA)	5.0 µL
dATP (0.5 mM)	4.0 µL
dGTP (0.5 mM)	4.0 µL
dTTP (0.5 mM)	4.0 µL
(α-^{32}P)-dCTP (Sp. activity 3000 Ci m/mol)	5.0 µL
Sterile deionized water	16.0 µL
Exonuclease free Klenow fragment (2 U/ µL)	2.0 µL
Total volume	**50.0 µL**

The reaction was carried out at 37^{o}C for 45 min and stopped by incubation in a boiling water bath for 10 min and snap chilled on ice.

2.8.6. Hybridization

The blots made as in sections 2.8.5.1 and 2.8.5.2 above were prehybridized at 45^{o}C in 30 mL of hybridization buffer for 6-8 h in a hybridization incubator (Robin Scientific, USA). The buffer was decanted and fresh buffer added with the denatured radiolabelled probe. Hybridization was carried out at 50^{o}C for 14-18 h. The solution was decanted and the membrane washed with low stringency buffer at 55-65^{o}C for 15 min followed by a high stringency wash at 55^{o}C for 15 min. The moist blot was wrapped in Saran wrap and exposed to X-ray film at –70^{o}C in a cassette with intensifying screen.

2.8.7. Expression And Purification of Recombinant Protein

2.8.7.1. Expression of Recombinant Protein

A single isolated bacterial colony from freshly streaked plates (grown on LB agar medium containing 50 μg / mL kanamycin) was used to ino Culate 5 mL liquid LB medium containing the same concentrations of the antibiotic. Culture was grown overnight with shaking at 225 rpm at 37 °C. One 1 ml aliquot of each culture was used to ino Culate 100 ml liquid cultures containing 50 μg/mL kanamycin. Once the cultures reached OD_{600} 0.4 - 0.5, recombinant protein expression was induced by the addition of isopropyl β-D-thiogalactopyranoside (IPTG), and the culture was grown either for 4–6 h at 37 °C (inclusion bodies) or for 9-12 h at 20°C (for soluble protein) with shaking at 150 rpm. Cells were harvested by centrifugation at 6000 rpm for 10 min at 4 °C. Pellets were resuspended in 6.25 mL lysis buffer. Cells were disrupted by sonication for 5 mins at 70 amplitude on a Sonifer Cell Disruptor. $MgSO_4$ was added to a final concentration of 10 mM and lysozyme to a final concentration 100 μg/mL to the disrupted cells and kept at 37^{o}C for ½ h. It was centrifuged at 10000 rpm for 10 minute and supernatant was saved as lysate and purified using Ni^{+} NTA beads either under native conditions (described in Chapter 2: section 2.6.6) or denaturation conditions (described in Chapter 2: section 2.6.7).The purified protein was tested for β-glucosidase activity (described in Chapter 2: section 2.8.7.3). Pellet was resuspended in 2 mL sonication buffer. Suspension was again sonicated for 1 min at 70 amplitude to disrupt the inclusion bodies. It was dissolved in 3 mL of dispersion buffer and an aliquot of 20 μL checked on SDS PAGE to check the expression (Chapter 2: section 2.6.5 and table: 2.6).

2.8.7.2. Affinity Purification of Recombinant Protein Using Ni^+ NTA Beads

The recombinant protein, among several other bacterial proteins is loaded on affinity matrix column such as Ni-agarose. This affinity matrix contains bound metal ion nickel, to which the poly histidine-tag binds with micro molar affinity. The matrix is then washed with buffer to remove unbound proteins. This can be achieved either by increasing the imidazole concentration in buffer (chapter 2: section 2.6.6, Table 2.7 or lowering the pH of the buffer (chapter 2: section 2.6.6, Table 2.8)

The column(s) and buffers were equilibrated to room temperature. The Ni^+ beads column was equilibrated with binding buffer for 30 minutes and then three bed volume of binding buffer was passed through the column followed by addition of soluble protein in dispersion buffer. The column was kept at 4°C for 1 h for binding of recombinant protein to Ni^+ beads. Flow through was collected in different tube afrter 1 h and column was washed with two bed volume of washing buffer (chapter 2: section 2.6.7). The washing efficiency may be improved by the addition of 20 mM imidazole and histidine-tagged proteins are then usually eluted with 150-300 mM imidazole. (Chapter 2: section 2.6.6, Tab 2.7) Unbound proteins can also be washed by lowering the pH of wash buffer to 5.9. The 6xHis-tagged protein was eluted with elution buffer with pH 4.9 (chapter 2: section 2.6.6, Tab 2.8)

The 6xHis-tagged bound protein was eluted in 4 aliquots of elution buffer, 0.5 mL each. Protein elution was monitored by measuring the absorbance at 280 nm of collected fractions. The eluted protein was analyzed for β-glucosidase activity (Chapter 2, section 2.8.7.3) and also by SDS-PAGE.

2.8.7.3. β-glucosidase Enzyme Assay.

Generally β-glucosidases have broad substrate specificity and they hydrolyze a range of artificial substrates. p-nitrophenyl-β-D-glycopyranoside (pNPG) is the most commonly used synthetic chromogenic substrate used to monitor β-glucosidase activity. The enzyme hydrolyses the glycosidic bond and the released aglycone after basification gives yellow colour. The concentration of the parantrophenol (aglucone) can be determined by measuring the optical density at 420 nm. In the present study the enzyme was incubated with 100 µM pNPG in 500 µL of 100-200mM Citrate –Phosphate buffer for 20 min and reaction was stopped by adding 500 µL 1 M Na_2CO_3. The O.D of the

reaction mixture was measured at 420nm. The reaction mixture without enzyme served as control.

2.8.7.4. Polyacrylamide gel electrophoresis (PAGE)

PAGE system is the widely used electrophoresis system for protein separations (Laemmeli, 1970). The resolution in a Laemmli gel is excellent because the treated peptides are stacked in a stacking gel before entering the separating gel.

2.8.7.5. Preparation of the Separating Gel

A vertical slab gel (Hoeffer Scientific, U.S.A.) was assembled using 1.0 mm spacers. In a side armed vacuum flask, 10% separating gel solution was made according to Table. 2.6 except for the addition of ammonium persulfate and TEMED. The solution was degassed by applying vacuum while stirring on a magnetic stirrer. TEMED and ammonium persulfate were added, and mixed gently without generating bubbles. The solution was pipetted into the gel cassette leaving 1.5 cm from the top unfilled. The gel solution was overlaid with water saturated n-butanol to remove trapped air bubbles and to layer evenly across the entire surface. When a sharp liquid-gel inter-surface was observed after the gel polymerization, the slab was tilted to pour off the overlay.

2.8.7.6. Preparation of the Stacking Gel

Stacking gel solution was prepared according to Table. 2.6, excluding ammonium per sulfate and TEMED. As in the separating gel, this solution was degassed. TEMED and ammonium per sulfate were added, mixed and overlaid on the separating gel. A comb was inserted taking care not to trap air bubbles beneath the comb teeth. The gel was left to polymerize.

2.8.7.7. Preparation of the sample

Equal parts of the protein sample and the loading buffer were mixed in a micro Centrifuge tube and kept in a boiling water bath for 90 s for SDS-PAGE. Then the samples were centrifuged at 10,000g for 10 min, supernatant collected and used for electrophoresis. Gel was run at room temperature at 100 V. The protein samples for native PAGE were mixed with native dye and gel was run at 80 V at 10ºC.

2.8.7.8. Loading and Running the Gel

The comb was removed from the gel, the wells were flushed with distilled water and drained off completely. The wells were filled with tank buffer and the samples under laid using a syringe. Lower and upper buffer chambers were filled with tank buffer. Voltage was set between 70 and 80 volts. The run was stopped when the dye reached the bottom of the gel.

3.1. Introduction

β- Glycosidases that belong to the family 1 glycoside hydrolases catalyze hydrolysis of the β-glycosidic bond in β-glycosides consisting of two carbohydrate moieties or a carbohydrate moiety linked to an aryl or alkyl aglycone. In plants, β-glycosidases serve a number of diverse and important functions, including bioactivation of defense compounds (Nisius, 1988; Poulton, 1990; Morant *et al*., 2003; Halkier and Gershenzon, 2006; Suzuki *et al*., 2006), cell wall degradation in endosperm during germination (Leah *et al*., 1995), activation of phytohormones (Kristoffersen *et al*., 2000; Lee *et al*., 2006), and lignification (Dharmawardhana *et al*., 1995; Escamilla-Trevino *et al*., 2006). In addition, β-glycosidases play key roles in aroma formation in tea, wine, and fruit juice (Mizutani *et al*., 2002; Fia *et al*., 2005; Maicas and Mateo, 2005). Plants produce myriad of secondary metabolites involved in defense against pathogens and herbivores. These defense compounds are often stored as β-glycosides and bioactivated by specific β-glycosidases (Morant *et al.,* 2008). β-glycosidases which act as the bioactivator include the α-hydroxynitrile glycosides (cyanogenic glycosides) that are found in numerous different plant species (Poulton, 1990; Hughes, 1993; Bak *et al*., 2006; Morant *et al*., 2007; Bjarnholt and Møller, 2008; Bjarnholt *et al*., 2008), benzoxazinoid glycosides in *Zea mays*, *Triticum aestivum*, and *Secale cereale* (Niemeyer, 1988; Sue *et al*., 2000a, 2000b), avenacosides in *Avena sativa* (Nisius, 1988; Kim *et al*., 2000), isoflavonoid glycosides in legumes (Cairns *et al*., 2000; Chuankhayan *et al.*, 2005, 2007a, 2007b; Suzuki *et al*., 2006), and glucosinolates found mainly in Brassicales (Halkier and Gershenzon, 2006). Most of these β-glucosidases belong to the Family 1 glycosyl hydrolases.

Complete genome sequence allowed to gauge the total number of members of the Family 1 Glycosyl hydrolases (GH) in model dicot *Arabidopsis* and monocot Rice, and found to be 47 and 40 in number respectively. However, functions of all the GH Family 1 members are not yet known. Functional expression of these GH family 1 members from other plants also helps to better understand the functions of all these GH Family 1members *in vivo*. Though there are many family 1 GH members characterized from

different plants, till date there are no studies on any of GH Family 1 member from *Leucaena leucocephala*

In the present study, using Polymerization Chain Reaction (PCR) based approach, a partial genomic DNA clone and a full-length cDNA clone of a Family 1 Glycosyl hydrolases has been isolated from *L. leucocephala*. This leguminous tree is used extensively as source of raw material in Indian pulp and paper industry.

3.2. Materials and Methods

3.2.1. Isolation of the family 1 glycosyl hydrolase gene from *Leucaena leucocephala*

There are different approaches to isolate genes like gDNA library screening; cDNA library screening and direct PCR based approach (RACE). An approach has been used here to isolate family 1 glycosyl hydrolases gene using PCR based method *i.e.* through RACE. The prerequisite for RACE reaction is to have a partial known sequence of the gene to be isolated. So the first thing is to isolate a partial cDNA of the gene.

3.2.2. Designing primers

The selected family 1 glycosyl hydrolases sequences publically accessible from the National Center for Biotechnology Information (NCBI) were aligned to generate a family 1 glycosyl hydrolases nucleotide consensus sequence. The alignment of these amino acids and their corresponding nucleotide sequences indicated that large stretches of the amino acids and nucleotides are highly conserved and thus could be used as primer targets.

3.2.2.1. Multiple sequence alignment of amino acid sequences of Family 1 Glycosyl hydrolases

The enzymes which are involved in plant defense by cleaving defensive glycosides, belonging to the Family 1 glycosyl hydrolases were aligned using multiple sequence alignment program Clustal W 1.8. The amino acid sequences of five enzymes, two from *Prunus* and one each from *Trifolium*, *Sorghum* and *Dalbergia* were selected for the alignment: Amygdalin hydrolase PsAH (AAA93234) and Prunasin Hydrolase PsPH

(AAL39079) from *Prunus serotina*, Cyanogenic β-glucosidase TrCBG (CAA40057) from *Trifolium*, Dalcochinin β-glucosidase DcBglu1 (AAF04007) from *Dalbergia* and Dhurrinase SbDhr1 (AAC49177) from *Sorghum bicolor*. (GenBank accession numbers are given in bracket)

3.2.2.2. Multiple sequence alignment of nucleotide sequence of conserved regions of the Family 1 Glycosyl hydrolases

The corresponding nucleotide sequences of the consensus regions of the family1 glycosyl hydrolases were aligned using ClustalW 1.8 multiple sequence alignment program. The homologous regions were manually edited and primers were designed from them.

3.2.3. Genomic DNA extraction and PCR

Genomic DNA was extracted from young and disease free leaves of *L. leucocephala* (Chapter 2; section 2.8.4.2). PCR reactions were set as described earlier (Chapter 2; section 2.8.4.9). PCR reactions were set up with *L. leucocephala* genomic DNA using all possible combinations of the forward and the reverse primers.

3.2.4. cDNA synthesis and amplification of partial cDNA fragments

3.2.4.1. cDNA synthesis

Total RNA was extracted from *in vitro* grown seedlings of *L. leucocephala* (Chapter 2; section 2.8.4.5). mRNA was purified using Oligotex (Chapter 2; section 2.8.4.6). cDNA first strand was synthesized using ImPromII™ Reverse Transcription System (Promega, USA) as described previously (Chapter 2; section 2.8.4.8).

3.2.4.2. Amplification of partial cDNA fragments

PCR reactions were setup with cDNA as template. Primers designed from conserved regions of Family 1 Glycosyl hydrolases were used.

3.2.5 Rapid Amplification of cDNA Ends (5'RACE and 3' RACE)

RACE ready cDNA was synthesized using RACE Ready cDNA Kit (Invitrogen, USA) as described previously (Chapter 2; section 2.8.4.10).

3.2.5.1. 3'RACE

Primary PCR was done using RACE Ready cDNA as template. Gene specific primer LRF1 was used as forward primer and Generacer 3' as reverse primer. The resultant PCR product was diluted and used as template for secondary PCR. LRF1 used as forward primer and generacer 3'nested primer as reverse primer. Details of PCR setup and cycling were described earlier (Chapter 2; section 2.8.4.10)

3.2.5.2. 5'RACE

Primary PCR was done using RACE Ready cDNA as template. Gene specific primer CBGR2 was used as forward primer and Generacer 5' as reverse primer. The resultant PCR product was diluted and used as template for secondary PCR. CBGR1 used as forward primer and generacer 5'nested primer as reverse primer. Details of PCR setup and cycling were described earlier (Chapter 2; section 2.8.4.10)

3.2.5.3. PCR amplification of full-length β-glucosidase

Gene specific forward and reverse primers were designed from the start and stop codon of 5'and 3'RACE sequences respectively and designated as AtgN(5'GGT ACC ATG ATG AAG AAG GTG ATG GTA GTA3') and NpTR3(**5'CTC GAG TTA ATA TTT TTG AAG GAA GTT CCT G 3'**). PCR was done using cDNA as template. An expected size of approximately 1.5 kb amplicon was amplified. The amplicon was cloned into pGEMTEasy vector and sequenced for confirmation.

3.2.6. Slot Blot Hybridization

For slot blot hybridization genomic DNA samples of *Leucaena leucocephala* and plasmid DNA of the clone (pLlbglu1) containing the full-length *Leucaena* β-glucosidase were used. DNA samples were diluted according to experimental requirements. The DNA samples were blotted on to Hybond-N+ membrane (Amersham, USA) using Slot blot apparatus (Hoefer Scientific, USA). A 5'RACE clone (cCBGR1) of *Leucaena* β-

glucosidase was used to amplify ~1.0 kb amplicon and used as probe. Random primer labeling of the DNA probe was done using the Megaprime DNA labeling kit (Amersham, USA). Pre-hybridization and hybridization of the blots were carried out in hybridization incubator (Robin Scientific, USA) as per detailed procedure described in Chapter 2, Section 2.8.5. The blot was checked for 20-50 counts per seconds by Geiger Muller counter, wraped in Saran wrap and exposed to X-ray film at -70^{o}C in a cassette with intensifying screen. After four days films were developed using developer and fixer solutions described in Appendix.

3.2.7. Characterization of the *L. leucocephala* family1 glycosyl hydrolase using bioinformatics tools

Nucleotide and amino acid sequence analysis of the *L. leucocephala* family1 glycosyl hydrolase was done using software pDRAW 32 and online bioinformatics analysis facility available at www.justbio.com, www.expasy.org and www.ncbi.nlm.nih.gov. Phylogenetic trees are constructed after alignment of amino acid sequences in Clustal X and Neighbour-Joining trees in Mega 4.0 software.

3.3 Results and Discussion

3.3.1. Multiple sequence alignment of amino acid sequences of Family1 Glycosyl hydrolases

Five Amino acid sequences of Family 1 Glycosyl hydrolases from *Prunus, Trifolium*, *Dalbergia* and *Sorghum*, involved in defense by cleaving defensive glycosides are selected from NCBI database. These sequences were aligned using Clustal X 1.8 multiple sequence alignment program. These sequences were having maximum homology at the 5' region. The conserved regions from these Family 1 Glycosyl hydrolases, chosen for designing primers are shown in red (Fig 3.1)

```
PsAH            -MATKLGSLL--------LCA-LLLAGFALTNSKAAKTDPPIHCASLNR-----------
PsPH            -MALQFRSLL--------LCVVLLLLGFALATTNAAGTDPPGVCTTLNR-----------
TrCBG           ------------------------LLSITTTHIHAFKP-LPISFDDFSD-----------
DcBglu1         MLAMTSKAIL--------LLG--LLALVSTSASIDFAKEVRETITEVPP-----------
SbDhr1          -MALLLASAINHTAHPAGLRSHPNNESFSRHHLCSSPQNISKRRSNLSFRPRAQTISSES
                                        .:                .

PsAH            -----------SSFDALEPGFIFGTASAAYQFEGAAKEDGRGPSIWDTYTHNHSERIKDG
PsPH            -----------TNFDTLFPGFTFGAATAAYQLEGAANIDGRGPSVWDNFTHEHPEKITDG
TrCBG           -----------LNRSCFAPGFVFGTASSAFQYEGAAFEDGKGPSIWDTFTHKYPEKIKDR
DcBglu1         -----------FNRSCFPSDFIFGTASSSYQYEG----EGRVPSIWDNFTHQYPEKIADR
```

```
SbDhr1          AGIHRLSPWEIPRRDWFPPSFLFGAATSAYQIEGAWNEDGKGPSTWDHFCHNFPEWIVDR
                              . : ..* **:*::::* **    :*: ** ** : *:..* * *

PsAH            SNGDVAVDQYHRYKEDVRIMKKMGFDAYRFSISWSRVLPNGKISGGVNEDGIKFYNNLIN
PsPH            SNGDVAIDQYHRYKEDVAIMKDMGFDAYRFSISWSRILPNGTLSGGINKKGIEYYNNLTN
TrCBG           TNGDVAIDEYHRYKEDIGIMKDMNLDAYRFSISWPRVLPKGKLSGGVNREGINYYNNLIN
DcBglu1         SNGDVAVDQFHRYKKDIAIMKDMNLDAYRMSISWPRILPTGRVSGGINQTGVDYYNRLIN
SbDhr1          SNGDVAADSYHMYAEDVRLLKEMGMDAYRFSISWPRILPKGTLAGGINEKGVEYYNKLID
                :***** *.:* * :*: ::*.*.:****:****.*:**.* ::**:*. *:.:**.* :

PsAH            EILRNGLKPFVTIYHWDLPQALEDEYGGFLSPNIVDHFRDYANLCFKKFGDRVKHWITLN
PsPH            ELLSNGIEPLVTLFHWDVPQALVDEYGGLLSPRIVDDFEAYANVCYNEFGDRVKRWTTLN
TrCBG           EVLANGMQPYVTLFHWDVPQALEDEYRGFLGRNIVDDFRDYAELCFKEFGDRVKHWITLN
DcBglu1         ESLANGITPFVTIFHWDLPQALEDEYGGFLNHSVVNDFQDYADLCFQLFGDRVKHWITLN
SbDhr1          LLLENGIEPYITIFHWDTPQALVDAYGGFLDEEDYKDYTDFAKVCFEKFGKTVKNWLTFN
                  * **: * :*::*** **** * * *:*.    ..:  :*.:*:: **. **.* *:*

PsAH            EPYTFSSSGYAYGVHAPGRCSAWQKLNCTGGNSATEPYLVTHHQLLAHAAAVKLYKDEYQ
PsPH            EPYTVSHHGYTIGIHAPGRCSSWYDPTCLGGDSSTEPYLVTHHLLLAHAAAVKLYKENYQ
TrCBG           EPWGVSMNAYAYGTFAPGRCSDWLKLNCTGGDSGREPYLAAHYQLLAHAAAARLYKTKYQ
DcBglu1         EPSIFTANGYAYGMFAPGRCSPSYNPTCTGGDAGTETYLVAHNLILSHAATVQVYKRKYQ
SbDhr1          EPETFCSVSYGTGVLAPGRCSPGVSCAVPTGNSLSEPYIVAHNLLRAHAETVDIYN-KYH
                **  .   .*  *  ******   .     *::  *.*:.:*  : :** :. :*: :*:

PsAH            ASQNGLIGITLVSPWFEPASEAEEDINAAFRSLDFIFGWFMDPLTNGNYPHLMRSIVGER
PsPH            ASQNGVIGITTVSHWFEPFSESQEDKDATSRALDFMYGWFMDPLTRGDYPQTMRSIVGSR
TrCBG           ASQNGIIGITLVSHWFEPASKEKADVDAAKRGLDFMLGWFMHPLTKGRYPESMRYLVRKR
DcBglu1         EHQKGTIGISLHVVWVIPLSNSTSDQNATQRYLDFTCGWFMDPLTAGRYPDSMQYLVGDR
SbDhr1          KGADGRIGLALNVFGRVPYTNTFLDQQAQERSMDKCLGWFLEPVVRGDYPFSMRVSARDR
                   .* **::       * ::   * :*  * :*   ***:.*:. * **  *:  . .*

PsAH            LPNFTEEQSKLLKGSFDFIGLNYYTTRYASNAPKITSVHA--SYITDPQVNAT-AELKGV
PsPH            LPNFTEEQSKSLTGSYDYIGVNYYSARYASAYTNNYSVPTPPSYATDAYVNVTTTDLNGI
TrCBG           LPKFSTEESKELTGSFDFLGLNYYSSYYAAKAPRIPNARP--AIQTDSLINAT-FEHNGK
DcBglu1         LPKFTTDQAKLVKGSFDFIGLNYYTTNYATKSDASTCCPP--SYLTDPQVTLL-QQRNGV
SbDhr1          VPYFKEKEQEKLVGSYDMIGINYYTSTFSKHIDLSPNNSP-VLNTDDAYASQETKGPDGN
                :* *. .: : : **:* :*:***:: ::          .      *.  .      .*

PsAH            PIGPMAASGWLYVYPKGIHDLVLYTKEKYNDPLIYITENGVDEFNDPKLSMEEALKDTNR
PsPH            PIGPRAASDWLYVYPKGLYDLVLYTKEKYNDPVMYITENGMDEFNVPKLSLDEALDDANR
TrCBG           PLGPMAASSWLCIYPQGIRKLLLYVKNHYNNPVIYITENGRNSSTIN--------TVTSR
DcBglu1         FIGPVTPSGWMCIYPKGLRDLLLYFKEKYNNPLVYITENGIDEKNDASLSLEESLIDTYR
SbDhr1          AIGPPTGNAWINMYPKGLHDILMTMKNKYGNPPMYITENGMGDIDKGDLPKPVALEDHTR
                 :** : . *: :**:*: .::: *::*.:* :****** ..               *

PsAH            IDFYYRHLCYLQAAIKKGSKVKGYFAWSFLDNFEWDAGYTVRFGINYVDYNDNLKRHSKL
PsPH            IDYYYHHLCYLQAAIKEGANVQGYFAWSLLDNFEWSEGYTVRFGINYVEYDSGLERHSKL
TrCBG           IPF-------------------------------------------------------
DcBglu1         IDSYYRHLFYVRYAIRSGANVKGFFAWSLLDNFEWAEGYTSRFGLYFVNYTT-LNRYPKL
SbDhr1          LDYIQRHLSVLKQSIDLGADVRGYFAWSLLDNFEWSSGYTERFGIVYVDRENGCERTMKR
                :

PsAH            STYWFTSFLKKYERSTKEIQMFVESKLEHQKFESQMMNKVQSSLAVVV
PsPH            SKHWFKSFLKKSSISKKKIRRSGNTNARATKFVYQM------------
TrCBG           --------------------------------------------------
DcBglu1         SATWFKYFLARDQESAKLEILAPKARWSLSTMIKEEKTKPKRGIEGF-
SbDhr1          SARWLQEFNGAAKKVENNKILTPAGQLN--------------------
```

Fig. 3.1. Multiple sequence alignment of the Family 1 glycosyl hydrolase amino acid sequences (Clustal W). Amygdalin hydrolase PsAH (AAA93234) and Prunasin Hydrolase PsPH (AAL39079) from *Prunus serotina*. Cyanogenic β-glucosidaseTrCBG (CAA40057) from *Trifolium*, Dalcochinin β-glucosidase DcBglu1 (AAF04007) from *Dalbergia* and Dhurrinase SbDhr1 (AAC49177) from *Sorghum bicolor* were used. GenBank accession numbers are given in bracket.

3.3.2. Multiple sequence alignment of nucleotide sequences of conserved sequences of the Family 1 Glycosyl hydrolases

The nucleotide sequences of all the five genes, corresponding to the consensus amino acid sequence (DAYRFSISWSRVLP and PFVTIYHWDLPQALEDEYGGFL**)** are selected for forward and reverse primer respectively (Fig 3.1). These nucleotide sequences are aligned using ClustalW program and also manually edited for designing primers.

3.3.2.1 Multiple sequence alignment of nucleotides for forward primers

Multiple sequence alignment of nucleotides for conserved region D A Y R M S I S W P R

```
                    ──────────►
TrCBG           GATGCTTATAGATTCTCCATTTCTTGGCCTAGA
DcBGLUI         GATGCTTATAGAATGTCCATCTCCTGGCCTAGA
PsAHI           GATGCTTATAGGTTTTCTATCTCGTGGTCCAGA
PsPHB           GATGCTTATAGGTTCTCTATCTCATGGTCCAGA
SbDhrI          GACGCCTATAGGTTCTCCATCTCTTGGCCCAGA
                ** ** *****  * ** ** ** *** * ***
```

Forward primer: F1 5′GAT TCG TAC AGA TTC TCA ATC TCT TGG 3′

3.3.2.2 Multiple sequence alignment of nucleotides for reverse primer

Multiple sequence alignment of nucleotides F H W D V P Q A L V D E Y G G L L

```
                                          ◄──────────────────────────────────
PsAHI           TATCATTGGGATCTTCCCCAAGCTTTAGAGGACGAATACGGTGGTTTCTTA
PsPHB           TATCATTGGGATCTTCCCCAAGCTTTAGAGGACGAATACGGTGGTTTCTTA
TrCBG           TTCCATTGGGACGTTCCGCAAGCCTTAGAGGACGAGTATCGCGGTTTTTTA
SbDhrI          TTCCATTGGGACGTTCCGCAAGCCTTAGAGGACGAGTATCGCGGTTTTTTA
DcBGLUI         TTCCACTGGGACACGCCTCAAGCGCTGGTAGACGCGTATGGCGGCTTCTTA
                *  ** *****    ** *****  * *  ****  **  * ** ** ***
```

Reverse primer: R2 5′TAA GAA ACC ACC GTA CTC ATC TTC TAA 3′

3.3.3. Genomic DNA extraction and PCR

Genomic DNA of good quality was isolated from the leaves of *L. leucocephala* (Fig. 3.2) PCR was done with Forward primer F1 (5'GATTCGTACA GATTCTCAAT CTCTTGG 3') and Reverse primer R2 (5'TAA GAA ACC ACC GTA CTC ATC TTC TAA3') using *Leucaena* gDNA as template. A fragment of approximately 400 bp was observed on 1% agarose gel (Fig 3.3). The amplified fragment was eluted from the agarose gel using gel extraction kit (Axyprep, USA) and cloned into pGEMT-Easy vector (Promega, USA). Positive clone was confirmed by *EcoRI* digestion which has

released a ~400 bp insert. The clone was referred as gCF1R2 and sequenced. Nucleotide sequence was edited manually for intron exon junctions Fig.3.4) with the help of other known sequences (example *Cicer*. Accession number (AJ630653)

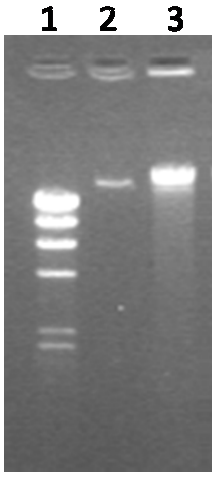


Fig.3.2. Genomic DNA of *L. leucocephala*: λDNA *Hind* III digest (lane 1), undigested phage DNA (lane 2), *Leucaena* genomic DNA (lane 3)

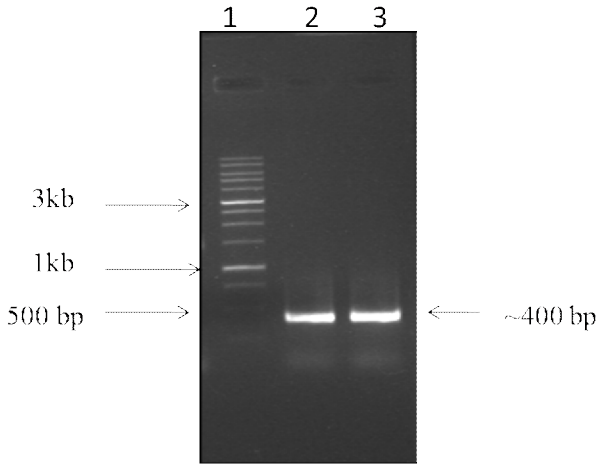


Fig.3.3. PCR with *L. leucocephala* genomic DNA template and primers F1 andR2. DNA size marker (lane 1), PCR amplified product (lane 2)

The coding region was translated into amino acid sequence using program Expasy translate tool. The amino acid sequence was searched against non-redundant protein sequences at NCBI using hprogram BLAST. It showed homology with known Family1 glycosyl hydrolases and maximum homology (79 % identity) with non-cyanogenic β-glucosidase from *Cicer (*Accession number CAG14979). The sequence was submitted to NCBI GenBank database under accession number (DQ417200).

```
  1 GATTCGTACA GATTCTCAAT CTCTTGGTCA AGAATACTGC CAAGTAAGTT
 51 AAATGTTACT TCATGTGCAT ATGAAAATGT TGTTTACTAG TTCCCCGTTT
101 CACTTTGACC CTTTTAATTT TTTATGAACT GTCTGCAGAA GGAAAGGTCG
151 GAGGAGGTAT TAATCAAGAA GGAGTTAAGT ATTACAACAA CCTCATCGAC
201 GAGCTATTGG CTAATGGTCA AAATTAAATT TTTATTTCGT TTTATCTCAT
251 TATTAGGTAG CATCATATTT ATTTCGAAGT TATAAACATG ACAATTAACT
301 TTTGAAATTA ATTCAGGTAT GAAACCATTT GTGACCATCT TCCATTGGGA
351 TGTTCCCCAA GCTTTAGAAG ATGAGTACGG TGGTTTCTTA
```

Fig.3.4. Nucleotide sequence of gCF1R2: Forward and reverse primers are in red colour and underlined. The sequence letters in black are exons and Sequence letters in pink are introns.

Nucleotide sequence of coding region of gF1R2 after manual editing

```
  1 GATTCGTACA GATTCTCAAT CTCTTGGTCA AGAATACTGC CAAAAGGAAA
 51 GGTCGGAGGA GGTATTAATC AAGAAGGAGT TAAGTATTAC AACAACCTCA
101 TCGACGAGCT ATTGGCTAAT GGTATGAAAC CATTTGTGAC CATCTTCCAT
151 TGGGATGTTC CCCAAGCTTT AGAAGATGAG TACGGTGGTT TCTTA
```

Deduced amino acid sequences using Just Bio Translate tool

```
  1 - GATTCGTACAGATTCTCAATCTCTTGGTCAAGAATACTGCCAAAAGGAAAGGTCGGAGGA - 60
  1 - D  S  Y  R  F  S  I  S  W  S  R  I  L  P  K  G  K  V  G  G   - 20

 61 - GGTATTAATCAAGAAGGAGTTAAGTATTACAACAACCTCATCGACGAGCTATTGGCTAAT - 120
 21 - G  I  N  Q  E  G  V  K  Y  Y  N  N  L  I  D  E  L  L  A  N   - 40

121 - GGTATGAAACCATTTGTGACCATCTTCCATTGGGATGTTCCCCAAGCTTTAGAAGATGAG - 180
 41 - G  M  K  P  F  V  T  I  F  H  W  D  V  P  Q  A  L  E  D  E   - 60

181 - TACGGTGGTTTCTTA - 195
 61 - Y  G  G  F  L  X
```

3.3.3. cDNA synthesis and partial cDNA amplification

3.3.3.1 Preparation of first strand

Total RNA was extracted from *Leucaena* seedlings by TRIZOL (Sigma,USA) (Described in Chapter 2; section 2.8.4.5). First strand was synthesized using total RNA by AMV Reverse Transcriptase (BD Clonetech,USA) (described in Chapter 2;section 2.8.4.8).

3.3.3.2 PCR amplification of partial cDNA fragments of *Leucaena* β-glucosidase.

PCR amplification using *Leucaena* cDNA as template and primers F1 and R2 (described in section 3.3.3) has resulted into amplification of expected size i.e. approximately 200 bp amplicon.(Fig3.5) It was cloned into pGEMT-Easy vector and confirmed by Restriction digestion with *Eco*RI (Fig.3.6) and sequencing. Now the clone was referred as cF1R2

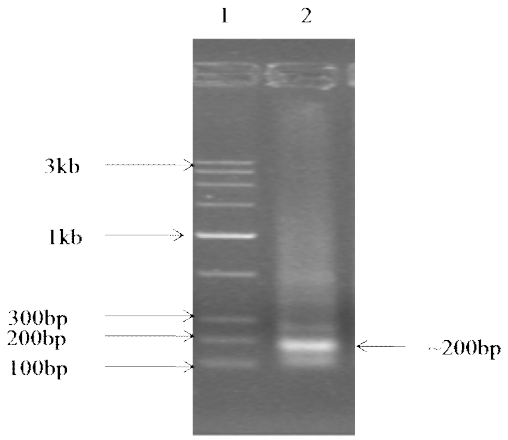


Fig. 3.5. PCR with *L. leucocephala* c DNA template and primers F1 and R2.
DNA size marker (lane 1), PCR amplified product (lane 2).

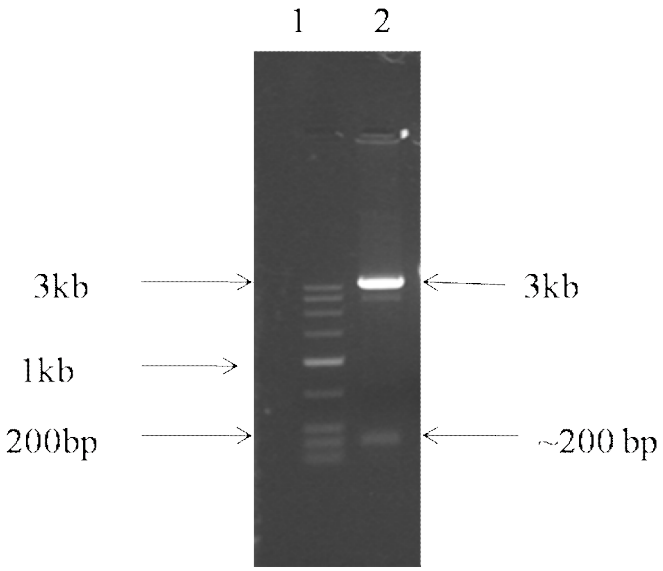


Fig. 3.6. Restriction digestion of cF1R2. DNA size marker (Lane 1), restriction digestion of cF1R2 with enzyme *EcoR* I (Lane2)

Nucleotide sequence of cF1R2

GATTCGTACA GATTCTCAAT CTCTTGGTCA AGAATCCTGC CAAAGGGGAA

GCTAAGCGGA GGTATAAACC AAGAAGGAAT CAAATATTAC AACAACCTCA

TCAACGAGCT ATTGGCTAAC GGTTTAAAAC CATTTGTAAC ACTGTTTCAT

TGGGATCTTC CGCAAGCCTT AGAAGATGAG TACGGTGGTT TCTTA

After removing the primer sequence the remaining nucleotide sequence (141bp) was translated into amino acid sequence and searched at NCBI using BLAST. It showed significant homology (81%) with *Trifolium* cyanogenic β-glucosidase (CAA40057).The sequence was submitted to NCBI GenBank database and later updated (accession number DQ883447).

3.3.3.4 Comparison of amino acid sequences derived from gDNA and cDNA

When amino acid sequences derived from gF1R2 and cF1R2 (after deleting primer sequences) aligned pair wise, mismatches were observed (Shown in Fig 3.7). These differences may be owing to Family 1 Glycosyl hydrolases being a multigene family, these sequences seem to have amplified from different members.

```
gF1R2                    SRILPKGKVGGGINQEGVKYYNNLIDELLANGMKPFVTIFHWDVPQA
cF1R2                    SRILPKGKLSGGINQEGIKYYNNLINELLPNGLKPFVTLFHWDLPQA
                         ********:.*******:*******:***.**:*****:****:***
```

Fig. 3.7. Pair wise alignment of amino acid sequences of gF1R2 and cF1R2

3.3.4. Rapid Amplification of cDNA Ends (RACE, 3' RACE and 5' RACE)

cDNA synthesis and 3'RACE was done using SMART RACE kit (BD Clonetech, Japan) as described earlier (Chapter 2,section 2.8.4.10). Primary PCR was done using Gene specific primers and Universal Primer Mix provided with the kit. Secondary PCR was done using nested primers and dilution of Primary PCR product as template.

3.3.4.1 3' Rapid Amplification of cDNA Ends (3'RACE)

For primary PCR, cDNA, Gene specific forward primer LRF1 (5'AAGAATCCTG CCAAAGGGGA AGC 3') and 3' Universal Primer Mix (provided with the kit) was used. The reaction mixture was diluted and secondary PCR was done using LRF1 and Nested Universal Primer provided with kit. An amplicon of approximately 1.3 kb was amplified (Fig 3.8).

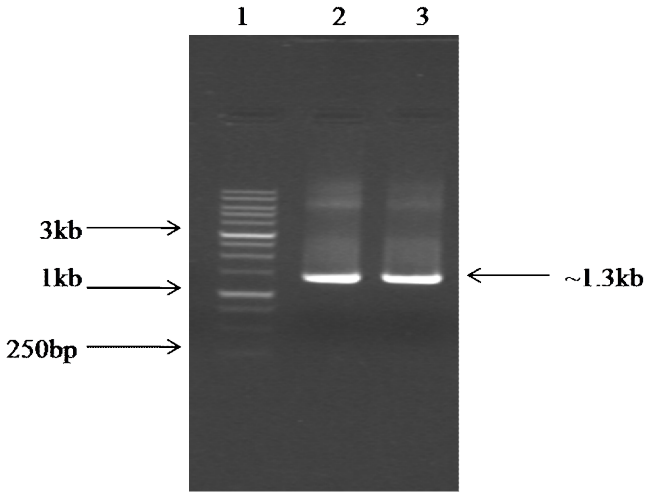


Fig. 3.8. Secondary PCR of 3' RACE with *L. leucocephala* c DNA template and primers LRF1 and NUP. DNA size marker (lane 1), PCR amplified product (lane 2 and 3)

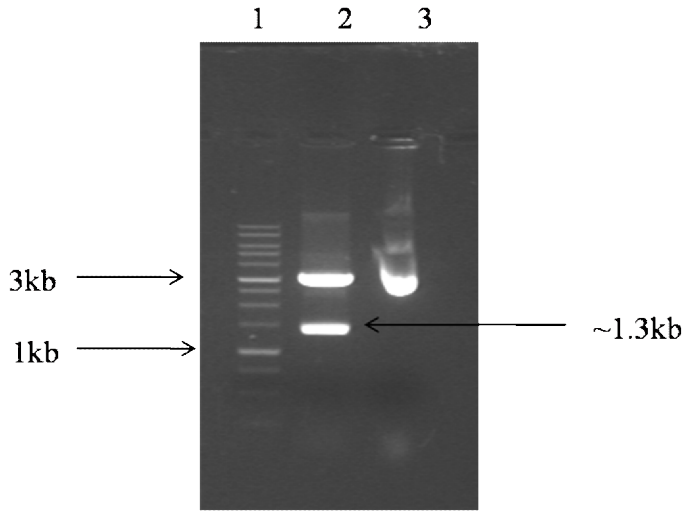


Fig. 3.9. Restriction digestion of 3LRF1. DNA size marker (Lane 1) ,restriction digestion with enzyme *EcoR* I (Lane2) and uncut plasmid c3LRF1 (Lane 3)

The fragment was gel eluted and cloned into pGMT-Easy vector. The clone c1.3KB was confirmed by restriction digestion analysis (Fig 3.9) and sequencing.

```
   1 AAGAATCCTG CCAAAGGGGA AGCTAAGCGG AGGTATAAAC CAAGAAGGAA
  51 TCAAATATTA CAACAACCTC ATCAACGAGC TATTGGCTAA CGGTTTAAAA
 101 CCATTTGTAA CACTGTTTCA TTGGGATCTT CCGCAAGCCT TAGAGGATGA
 151 GTACGGGGGC TTCTTGAGCT CTGACATAGT GAAGGACTAT GGGGATTACG
 201 CGGAACTATG TTTCAAAGAG TTCGGGGACA GAGTGAAGCA TTGGATAACA
 251 CTGAACGAGC CATGGACTTA CAGCAATGGT GGCTATGCGA TGGGGCAGCA
 301 GGCACCGGGT AGATGTTCTG CTTGGCTCCG TCTTAATTGC ACCGGCGGCG
 351 ATTCCTCCAC CGAGCCCTAT CTTGTTGCTC ACCACCTCCT TCTTTCTCAT
 401 GCTTCCGCCG TGCAAATCTA CAAGTCTAAG TTTCAGGCAT CTCAAAAGGG
 451 TGTCATAGGC ATAACTCTGG TGTGCCACTG GTTCGTGCCC TTGTCAGACA
 501 AGAAATCCGA CCAAAATGCT GCCGCACGAG CCGTTGATTT CATGTTTGGA
 551 TGGTTTATGG GACCGTTGAC CGAAGGAGAG TACCCGAAGA GCATGCGAGC
 601 ATTAGTGGGA AGTCGATTGC CAAAGTTCTC AAAGAAAGAA TCCAGCCTTG
 651 TCAAAGGTTC TTTCGATTTT CTTGGGCTTA ACTACTACAC TGCTAATTAT
 701 GCTGCCAATG CACCTTCCCT CCGCAATGCC AGACCCTCGT ATCAAACTGA
 751 TTCTCATGCC AATCTTACAA CTGAGCGCAA CGGGACACCC ATTGGTCCAA
 801 GGGCAGCATC TGATTGGTTA TATGTTTATC CCAGAGGAAT TAGAGACTTA
 851 CTGCTCTATG TCAAGACTAA GTATAACAAT CCTTTGATTT ACATCACTGA
 901 AAACGGTATA GACGAGTTCG ATGATCCAAC ACTGACATTG GAAGAAGCCC
 951 TCATAGATAC TTTTAGAATT GATTACTATT TTCGTCATCT TTATTATCTT
1001 CAATCTGCAA TCAAGGATGG CGCCAATGTG AAAGGATACT TTGCGTGGTC
1051 ATTACTGGAC AACTTTGAAT GGGCTAGCGG CTACACCGTG CGTTTTGGAA
1101 TCAACTTCGT GGATTATAAA CATGGCAACC AAAGATACCA CAAGCTCTCA
1151 GCTCAATGGT TCAGGAACTT CCTTCAAAAA TATTAATAAA TCTTCTGTAC
1201 CAAATGATTC CCAATAAACA ATGTCATGAA TATATTTCTG TTAACAGTGG
1251 CAGTACTCTT TGTAATATAT TCTTATTCAT AGTATATATA CCTGGTCTTC
1301 TCGTCAAAAA AAAAAAAAAA AAAAAAAAAA AAA
```

Fig. 3.10. Nucleotide sequence of c1.3KB. Forward primer LRF1 shown in red. Stop codon shown in pink. 3' UTR was shown in green and poly A tail in blue.

Nucleotide sequence of c1.3KB has revealed that it has a coding sequence of 1183 bp followed by stop codon TAA. The sequence has a 3'untranslated region(3'UTR) of 119 bp followed by a poly A tail (Fig. 3.10). The nucleotide sequence was translated to amino acid sequence (Fig: 3.11) the latter was searched in BLAST in NCBI database. It was found to contain a putative conserved domain found in Family 1 Glycosyl hydrolases. Conserved catalytic motifs ENG and NEP characteristics of this family were also found. The sequence was submitted to NCBI GenBank database (accession number DQ883447)

R I L P K G K L S G G I N Q E G I K Y Y N N L I N E L L A N G L K P F V T
L F H W D L P Q A L E D E Y G G F L S S D I V K D Y G D Y A E L C F K E F
G D R V K H W I T L **N E P** W T Y S N G G Y A **M** G Q Q A P G R C S A W L R L
N C T G G D S S T E P Y L V A H H L L L S H A S A V Q I Y K S K F Q A S Q
K G V I G I T L V C H W F V P L S D K K S D Q N A A A R A V D F **M** F G W F
M G P L T E G E Y P K S **M** R A L V G S R L P K F S K K E S S L V K G S F D
F L G L N Y Y T A N Y A A N A P S L R N A R P S Y Q T D S H A N L T T E R
N G T P I G P R A A S D W L Y V Y P R G I R D L L L Y V K T K Y N N P L I
Y I T **E N G** I D E F D D P T L T L E E A L I D T F R I D Y Y F R H L Y Y L
Q S A I K D G A N V K G Y F A W S L L D N F E W A S G Y T V R F G I N F V
D Y K H G N Q R Y H K L S A Q W F R N F L Q K Y

Fig. 3.11. Deduced amino acid sequence of c1.3KB.

3.3.4.2. 5' Rapid Amplification of cDNA Ends (5'RACE)

3 gene specific reverse primers were designed using Primer 3 program from c1.3KB for 5' RACE reaction. Primary PCR was done with primers 5'UPM and CBGR2. The secondary PCR was done using gene specific nested primer CBGR1 and 5'NUP. An amplicon of approximately 1.0 kb was amplified (Fig. 3.12).

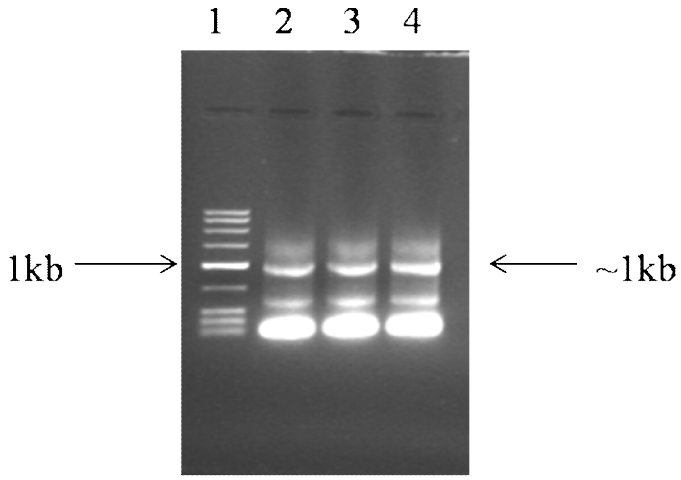


Fig. 3.12.Secondary PCR of 5' RACE with *L. leucocephala* c DNA template and primers CBGR1 and 5'NUP. Lane 1: DNA size marker, lane 2- 4: PCR amplified product

The fragment was gel eluted and cloned into pGEMT-Easy vector. The clone c5CBGR1 was confirmed by restriction digestion analysis and sequencing (Fig. 3.13).

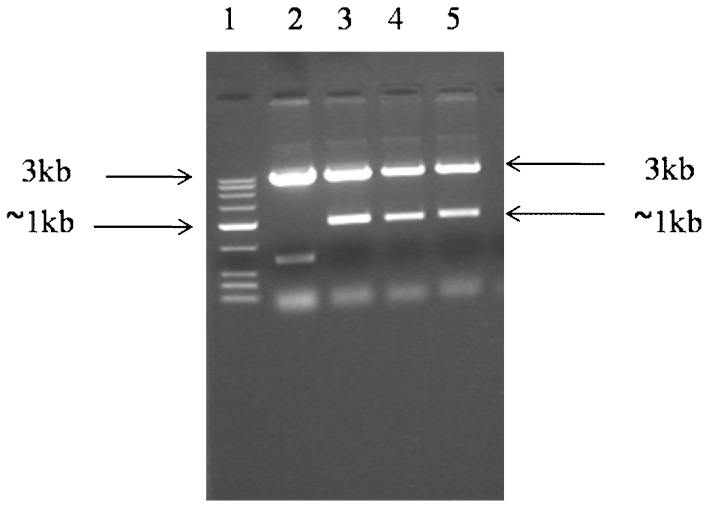


Fig.3.13. Restriction digestion of c5CBGR1. Lane 1: DNA size marker, Lane: 2-5 Restriction digestion of CBGR1 with enzyme *EcoR* I

The nucleotide sequence of c5CBGR1 was analyzed using bioinformatics tools. A 18 bp 5'UTR was observed followed by start codon ATG shown in Blue (Fig.3.14).

Deduced amino acid sequence showed in Fig 3.15.The amino acid sequence shows homology with Family 1 glycosyl hydrolase in NCBI database. The sequence was submitted to NCBI GenBank database (Accession no. EF061245**).**

```
  1 GAAAATATAG ATTGGAGGAT GATGAAGAAG GTGATGGTAG TAGCCGCCGT
 51 CATGTGGGCT CTCATAACTG TTGCTGCAGC TGATGCAACA AATGATATTT
101 CCTCTCTCAG TCGCCGGAGT TTCGCCCCGG CCTTCATTTT CGGCACCGCC
151 TCCGCCTCCT ACCAGTACGA AGGTGCAGCA AAGGAAGGCG GCAGAGGACC
201 CAGCATATGG GATACCTTCA CCCACAAATA CCCAGAAAAA ATATCGGATC
251 GCAGCAACGG AGATGTAGCT AACGACGAAT ATCATCGGTA TAAGGAGGAT
301 GTTGGAATCA TGAAGTATAT GAATTTAGAT GCTTACAGAT TCTCCATTTC
351 TTGGTCAAGA ATCCTGCCAA AGGGGAAGCT AAGCGGAGGT ATAAACCAAG
401 AAGGAATCAA ATATTACAAC AACCTCATCA ACGAGCTATT GGCTAACGGT
451 TTAAAACCAT TTGTAACACT GTTTCATTGG GATCTTCCGC AAGCCTTAGA
501 GGATGAGTAC GGGGGCTTCT TGAGCTCTGA CATAGTGAAG GACTATGGGG
551 ATTACGCGGA GCTATGTTTC AAAGAGTTCG GGGACAGAGT GAAGCATTGG
601 ATAACACTGA ACGAGCCGTG GACTTACAGC AATGGTGGCT ATGCGATGGG
651 GCAGCAGGCA CTGGGTAGAT GTTCTGCTTG GCTCCGTCTT AATTGCACCG
701 GCGGCGATTC CTCCACCGAG CCCTATCTTG TTGCTCACCA CCTCCTTCTT
751 TCTCATGCTT CCGCCGTGCA AATCTACAAG TCTAAGTTTC AGGCATCTCA
801 AAAGGGTGTC ATAGGCATAA CTCTGGTGTG CCACTGGTTC GTGCCCTTGT
851 CAGACAAGAA ATCCGACCAA AATGCTGCCG CACGAGCCGT TGATTTCATG
901 TTTGGATGGT TTATGGGACC GTTGACCGAA
```

Fig. 3.14. Nucleotide sequence of c5CBGR1

```
Met M K K V M V V A A V M W A L I T V A A A D A T N D I S S L S R R S F A
P A F I F G T A S A S Y Q Y E G A A K E G G R G P S I W D T F T H K Y P E
K I S D R S N G D V A N D E Y H R Y K E D V G I M K Y M N L D A Y R F S I
S W S R I L P K G K L S G G I N Q E G I K Y Y N N L I N E L L A N G L K P
F V T L F H W D L P Q A L E D E Y G G F L S S D I V K D Y G D Y A E L C F
K E F G D R V K H W I T L N E P W T Y S N G G Y A M G Q Q A L G R C S A W
L R L N C T G G D S S T E P Y L V A H H L L L S H A S A V Q I Y K S K F Q
A S Q K G V I G I T L V C H W F V P L S D K K S D Q N A A A R A V D F M F
G W F M G P L T E
```

Fig. 3.15. Deduced amino acid sequence of c5CBGR1.

The complete nucleotide sequence of *L. leucocephala* β-glucosidase along with UTR sequences is shown in figure 3.16.

```
   1 GAAAATATAG ATTGGAGGAT GATGAAGAAG GTGATGGTAG TAGCCGCCGT
  51 CATGTGGGCT CTCATAACTG TTGCTGCAGC TGATGCAACA AATGATATTT
 101 CCTCTCTCAG TCGCCGGAGT TTCGCCCCGG CCTTCATTTT CGGCACCGCC
 151 TCCGCCTCCT ACCAGTACGA AGGTGCAGCA AAGGAAGGCG GCAGAGGACC
 201 CAGCATATGG GATACCTTCA CCCACAAATA CCCAGAAAAA ATATCGGATC
 251 GCAGCAACGG AGATGTAGCT AACGACGAAT ATCATCGGTA TAAGGAGGAT
 301 GTTGGAATCA TGAAGTATAT GAATTTAGAT GCTTACAGAT TCTCCATTTC
 351 TTGGTCAAGA ATCCTGCCAA AGGGGAAGCT AAGCGGAGGT ATAAACCAAG
 401 AAGGAATCAA ATATTACAAC AACCTCATCA ACGAGCTATT GGCTAACGGT
 451 TTAAAACCAT TTGTAACACT GTTTCATTGG GATCTTCCGC AAGCCTTAGA
 501 GGATGAGTAC GGGGGCTTCT TGAGCTCTGA CATAGTGAAG GACTATGGGG
 551 ATTACGCGGA GCTATGTTTC AAAGAGTTCG GGGACAGAGT GAAGCATTGG
 601 ATAACACTGA ACGAGCCGTG GACTTACAGC AATGGTGGCT ATGCGATGGG
 651 GCAGCAGGCA CTGGGTAGAT GTTCTGCTTG GCTCCGTCTT AATTGCACCG
 701 GCGGCGATTC CTCCACCGAG CCCTATCTTG TTGCTCACCA CCTCCTTCTT
 751 TCTCATGCTT CCGCCGTGCA AATCTACAAG TCTAAGTTTC AGGCATCTCA
 801 AAAGGGTGTC ATAGGCATAA CTCTGGTGTG CCACTGGTTC GTGCCCTTGT
 851 CAGACAAGAA ATCCGACCAA AATGCTGCCG CACGAGCCGT TGATTTCATG
 901 TTTGGATGGT TTATGGGACC GTTGACCGAA GGAGAGTACC CGAAGAGCAT
 951 GCGAGCATTA GTGGGAAGTC GATTGCCAAA GTTCTCAAAG AAAGAATCCA
1001 GCCTTGTCAA AGGTTCTTTC GATTTTCTTG GGCTTAACTA CTACACTGCT
1051 AATTATGCTG CCAATGCACC TTCCCTCCGC AATGCCAGAC CCTCGTATCA
1101 AACTGATTCT CATGCCAATC TTACAACTGA GCGCAACGGG ACACCCATTG
1151 GTCCAAGGGC AGCATCTGAT TGGTTATATG TTTATCCCAG AGGAATTAGA
1201 GACTTACTGC TCTATGTCAA GACTAAGTAT AACAATCCTT TGATTTACAT
1251 CACTGAAAAC GGTATAGACG AGTTCGATGA TCCAACACTG ACATTGGAAG
1301 AAGCCCTCAT AGATACTTTT AGAATTGATT ACTATTTTCG TCATCTTTAT
1351 TATCTTCAAT CTGCAATCAA GGATGGCGCC AATGTGAAAG GATACTTTGC
1401 GTGGTCATTA CTGGACAACT TTGAATGGGC TAGCGGCTAC ACCGTGCGTT
1451 TTGGAATCAA CTTCGTGGAT TATAAACATG GCAACCAAAG ATACCACAAG
1501 CTCTCAGCTC AATGGTTCAG GAACTTCCTT CAAAAATATT AATAAATCTT
1551 CTGTACCAAA TGATTCCCAA TAAACAATGT CATGAATATA TTTCTGTTAA
1601 CAGTGGCAGT ACTCTTTGTA ATATATTCTT ATTCATAGTA TATATACCTG
1651 GTCTTCTCGT CAAAAAAAAA AAAAAAAAAA AAAAAAAAA
```

Fig. 3.16. The complete nucleotide sequence of *L. leucocephala* β-glucosidase.

5'UTR shown in pink, 3'UTR in blue and overlapping region in green. Start and stop codons shown in red where as the gene specific primers used for RACE were shown in red with underline (Fig 3.16).

3.3.5. Isolation of Full-length cDNA of the family 1 glycosyl hydrolase from *Leucaena leucocephala*

Gene specific forward and reverse primers were designed (with *Kpn I* and *Xho I*) from the start and stop codon of 5'and 3'RACE sequences respectively and designated as AtgN(5'GGT ACC ATG ATG AAG AAG GTG ATG GTA GTA3') and NpTR3(5'CTC GAG TTA ATA TTT TTG AAG GAA GTT CCT G 3'). PCR was done using cDNA as template. An expected size of approximately 1.5 kb amplicon was amplified (Fig. 3.17). The amplicon was cut, eluted from gel and ligated into pGEMT-Easy vector. It was confirmed by restriction digestion with *Eco*RI which released 1.5 kb fragment (Fig.3.18). The clone now referred as cLlbglu1 was sequenced. The sequence showed 77% identity with *Lotus* β-glucosidase (EU710846) and 78% identity with Rose β-glucosidase (AB426520). The sequence was submitted to NCBI GenBank database (EU328158).

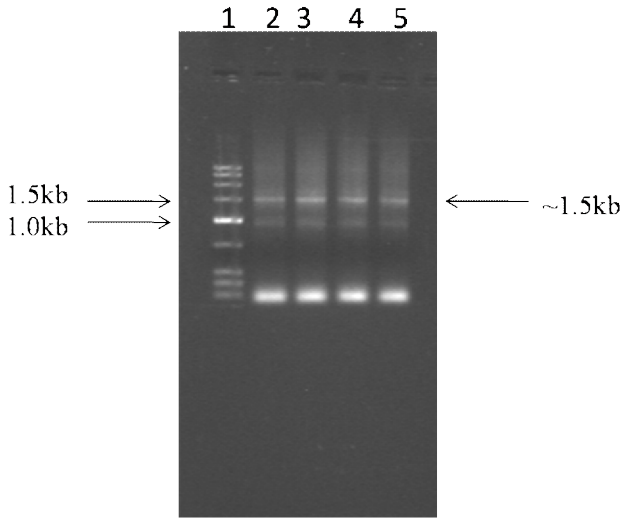


Fig. 3.17. PCR with *L. leucocephala* c DNA template and primers AtgN and NpTR3. lane 1: DNA size marker, lane 2-4: PCR amplified product

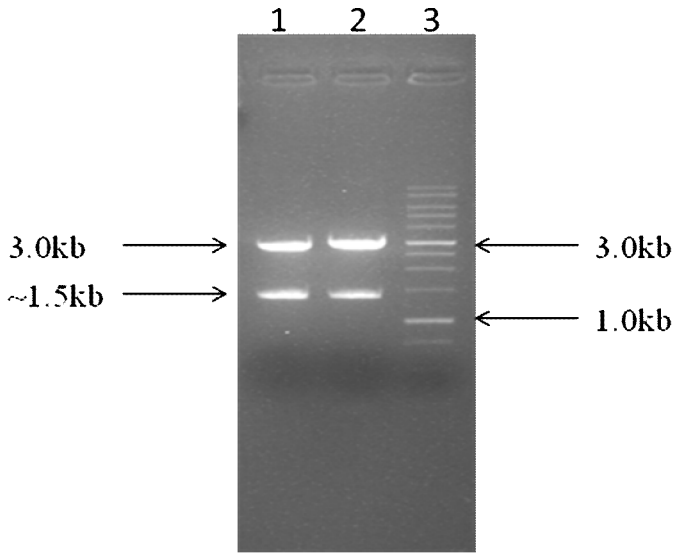


Fig. 3.18. Restriction digestion of cLlbglu1. Lane1 and 2: Restriction digestion ofcLlbglu1 with enzyme *EcoR* I and Lane 3: DNA size marker.

Nucleotide sequence of cDNA clone (cLlbglu1) encoding a full length β-Glucosidase

```
   1 ATGATGAAGA AGGTGATGGT AGTAGCCGCC GTCATGTGGG CTCTCATAAC
  51 TGTTGCTGCA GCTGATGCAA CAAATGATAT TTCCTCTCTC AGTCGCCGGA
 101 GTTTCGCCCC GGCCTTCATT TTCGGCACCG CCTCCGCCTC CTACCAGTAC
 151 GAAGGTGCAG CAAAGGAAGG CGGCAGAGGA CCCAGCATAT GGGATACCTT
 201 CACCCACAAA TACCCAGAAA AAATATCGGA TCGCAGCAAC GGAGATGTAG
 251 CTAACGACGA ATATCATCGG TATAAGGAGG ATGTTGGAAT CATGAAGTAT
 301 ATGAATTTAG ATGCTTACAG ATTCTCCATT TCTTGGTCAA GAATCCTGCC
 351 AAAGGGGAAG CTAAGCGGAG GTATAAACCA AGAAGGAATC AAATATTACA
 401 ACAACCTCAT CAACGAGCTA TTGGCTAACG GTTTAAAACC ATTTGTAACA
 451 CTGTTTCATT GGGATCTTCC GCAAGCCTTA GAGGATGAGT ACGGGGGCTT
 501 CTTGAGCTCT GACATAGTGA AGGACTATGG GGATTACGCG GAGCTATGTT
 551 TCAAAGAGTT CGGGGACAGA GTGAAGCATT GGATAACACT GAACGAGCCG
 601 TGGACTTACA GCAATGGTGG CTATGCGATG GGGCAGCAGG CACTGGGTAG
 651 ATGTTCTGCT TGGCTCCGTC TTAATTGCAC CGGCGGCGAT TCCTCCACCG
 701 AGCCCTATCT TGTTGCTCAC CACCTCCTTC TTTCTCATGC TTCCGCCGTG
 751 CAAATCTACA AGTCTAAGTT TCAGGCATCT CAAAAGGGTG TCATAGGCAT
 801 AACTCTGGTG TGCCACTGGT TCGTGCCCTT GTCAGACAAG AAATCCGACC
 851 AAAATGCTGC CGCACGAGCC GTTGATTTCA TGTTTGGATG GTTTATGGGA
 901 CCGTTGACCG AAGGAGAGTA CCCGAAGAGC ATGCGAGCAT TAGTGGGAAG
 951 TCGATTGCCA AAGTTCTCAA AGAAAGAATC CAGCCTTGTC AAAGGTTCTT
1001 TCGATTTTCT TGGGCTTAAC TACTACACTG CTAATTATGC TGCCAATGCA
1051 CCTTCCCTCC GCAATGCCAG ACCCTCGTAT CAAACTGATT CTCATGCCAA
1101 TCTTACAACT GAGCGCAACG GGACACCCAT TGGTCCAAGG GCAGCATCTG
1151 ATTGGTTATA TGTTTATCCC AGAGGAATTA GAGACTTACT GCTCTATGTC
```

```
1201 AAGACTAAGT ATAACAATCC TTTGATTTAC ATCACTGAAA ACGGTATAGA
1251 CGAGTTCGAT GATCCAACAC TGACATTGGA AGAAGCCCTC ATAGATACTT
1301 TTAGAATTGA TTACTATTTT CGTCATCTTT ATTATCTTCA ATCTGCAATC
1351 AAGGATGGCG CCAATGTGAA AGGATACTTT GCGTGGTCAT TACTGGACAA
1401 CTTTGAATGG GCTAGCGGCT ACACCGTGCG TTTTGGAATC AACTTCGTGG
1451 ATTATAAACA TGGCAACCAA AGATACCACA AGCTCTCAGC TCAATGGTTC
1501 AGGAACTTCC TTCAAAAATA TTAA
```

3.3.5.2. Restriction map of *Leucaeana* β-glucosidase

Frequently used restriction enzymes having 6bp length sites were selected using program pDRAW and the nucleotide sequence of full length cDNA of *Leucaena leucocephala* was analyzed for the restriction sites. Details of position of restriction site and name of restriction enzymes for entire sequence is shown in figure 3.19.

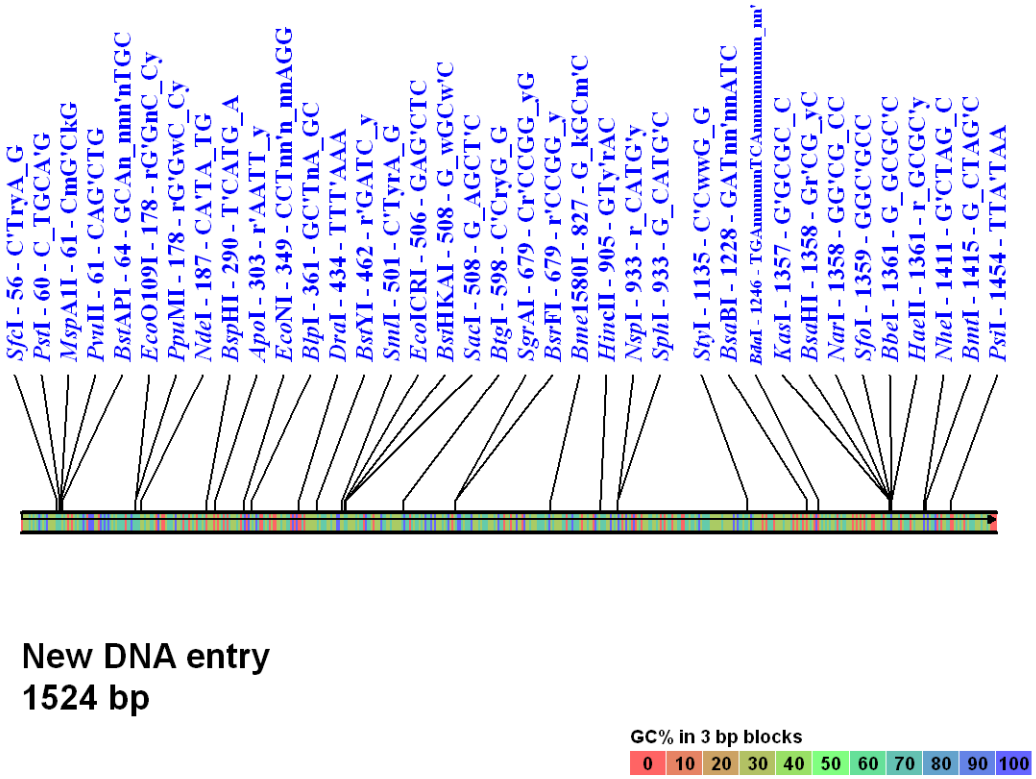


Fig: 3.19. Image showing the restriction enzyme sites in the nucleotide sequence of *Leucaena leucocephala* glycosyl hydrolase

3.3.6. Slot blot hybridization

The genome size of *L.leucocephala* has ~1.81 pg of DNA per haploid genome. 1.81 pg corresponds to a single representation of *L.leucocephala* genomic DNA, 0.18 µg, 0.36 µg and 0.54 µg representing the genome $1x10^5$, $2x10^5$ and $3x10^5$ times respectively was spotted on Hybond N+ membrane (Amersham, USA). The plasmid DNA of the clone cCBGR1 (Accession no. EF061245) was spotted as standard dilutions representing $1x10^5$, $2x10^5$ and $3x10^5$ copies in duplicate. The blot was probed with radiolabelled ~1.0kb amplicon of cCBGR1. Positive signal obtained under high stringency hybridization conditions was indicative of presence of the β-glucosidase gene in *L. leucocephala*. Based on signal intensity it was also inferred that in *L. leucocephala* the β-glucosidase gene represented possibly by two members (Fig 3.20)

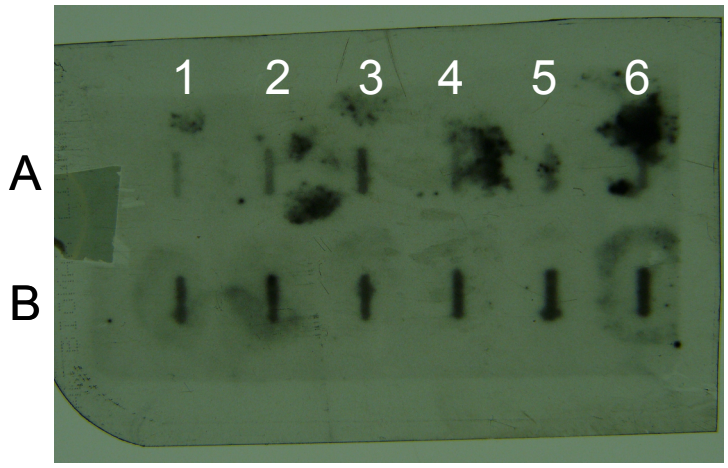


Fig. 3.20. Slot blot hybridization of Bglu1 gene in *l. Leucocephala*

Lane A1, A2, A3, A4, A5 and A6; Signals from $1x10^5$, $2x10^5$ and $3x10^5$ copies of Plasmid DNA representations of *L. leucocephala* clone (cCBGRI) in duplicates.

Lane B1, B 2, B 3, B 4, B 5 and B 6; Signals from $1x10^5$, $2x10^5$ and $3x10^5$ copies of genomic DNA representations of *L. leucocephala* in duplicates.

3.3.7.1. Multiple sequence alignment of nucleotide sequences of selected Family1 glycosyl hydrolases.

The *Leucaena* β-glucosidase (EU328158) nucleotide sequence was used for homology search using BLASTN program (available at www.ncbi.nlm.nih) against somewhat similar sequences available in database. Five nucleotide sequences of family 1 glycosyl hydrolases namely *Lotus japonicus* β-glucosisidase D2 (EU710845), *Medicago truncatula* β-glucosidase G3(EU078903), cyanogenic β glucosidase from *Trifolium repens* (X56733), *Glycine max*Gm1CHG (AB259819), Prunus hydrolase from *Prunus* (AF221526) were selected for multiple sequence alignment. These sequences have 70-75 % identity and approximately 90% query coverage with the *Leucaena* β-glucosidase (EU328158) nucleotide sequence .They are used for multiple sequence alignment using program Clustal X. The conserved sequences are shown in figure 3.21.

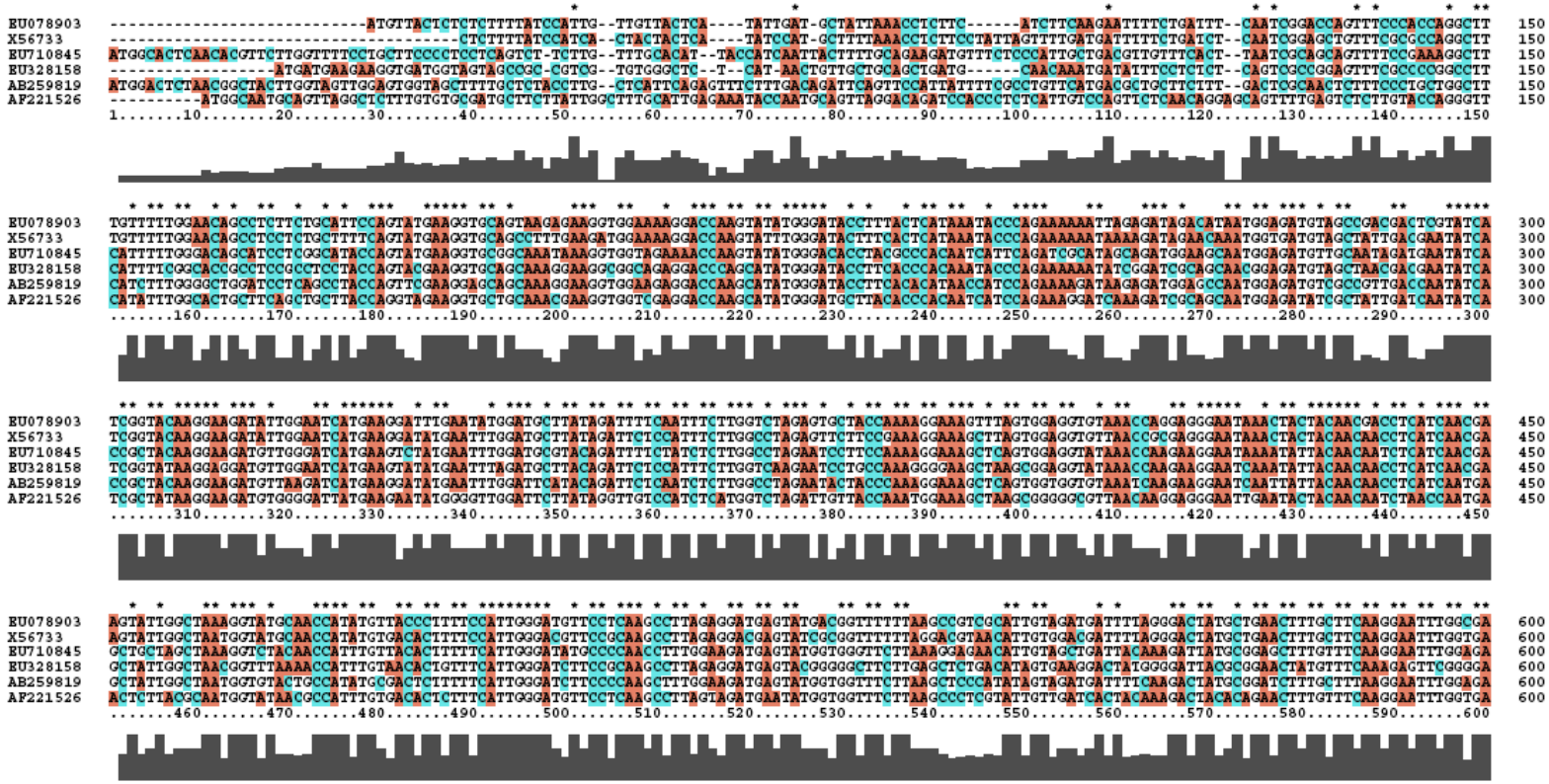


Fig 3.21

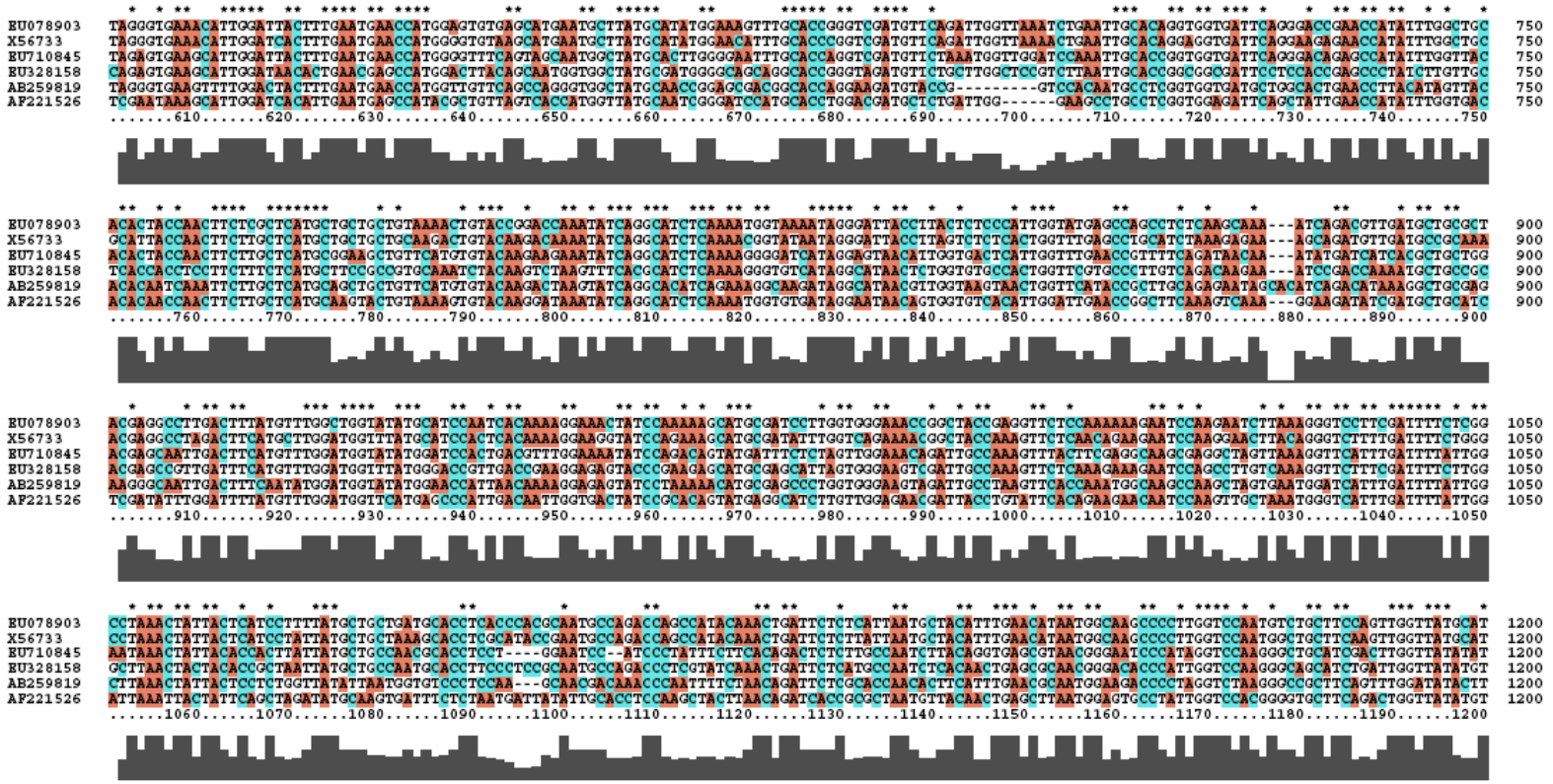


Fig 3.21 continued

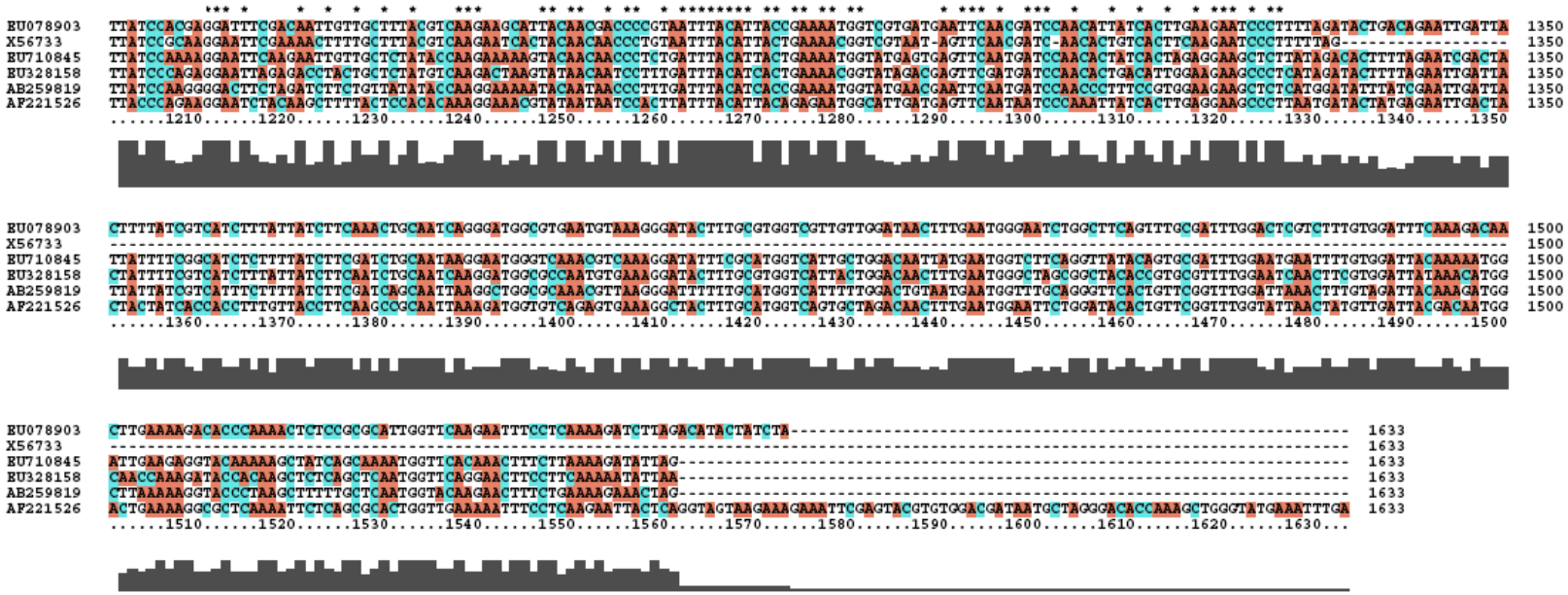


Fig 3.21. Multiple sequence alignment of nucleotide sequence of selected family 1 Glycosyl hydrolases with *Leucaena* β-glucosidase .

Deduced amino acid sequence of cLlbglu1 nucleotide sequence.

```
M M K K V M V V A A V M W A L I T V A A A D A T N D I S S L S R R S F A
P A F I F G T A S A S Y Q Y E G A A K E G G R G P S I W D T F T H K Y P
E K I S D R S N G D V A N D E Y H R Y K E D V G I M K Y M N L D A Y R F
S I S W S R I L P K G K L S G G I N Q E G I K Y Y N N L I N E L L A N G
L K P F V T L F H W D L P Q A L E D E Y G G F L S S D I V K D Y G D Y A
E L C F K E F G D R V K H W I T L N E P W T Y S N G G Y A M G Q Q A L G
R C S A W L R L N C T G G D S S T E P Y L V A H H L L L S H A S A V Q I
Y K S K F Q A S Q K G V I G I T L V C H W F V P L S D K K S D Q N A A A
R A V D F M F G W F M G P L T E G E Y P K S M R A L V G S R L P K F S R
K N P S L V K G S F D F L G L N Y Y T A N Y A A N A P S L R N A R P S Y
Q T D S H A N L T T E R N G T P I G P R A A S D W L Y V Y P R G I R D L
L L Y V K T K Y N N P L I Y I T E N G I D E F D D P T L T L E E A L I D
T F R I D Y Y F R H L Y Y L Q S A I K D G A N V K G Y F A W S L L D N F
E W A S G Y T V R F G I N F V D Y K H G N Q R Y H K L S A Q W F R N F L
Q K Y Stop
```

Fig. 3.22. Deduced amino acid sequence of cLlbglu1 nucleotide sequence

Homology search was done for the amino acid sequence using BLASTP program (available at www.ncbi.nlm.nih) against non-redundant database. Llbglu1 has significant homology with hydroxynitrile glucoside cleaving β-glucosidases from *Lotus japonicus* β-glucosidase D7 (ACD65511) with 77%, D2 (ACD65510) with 74% and D4 (ACD65509) with 73% identity respectively. The sequence also showed homology with isoflavonoid hydrolyzing β-glucosidases from *Medicago truncatula* β-glucosidase G2 (ABW76287) with 73% and G3 (ABW76288) with 68% identity.

It showed 67% identity with amygdalin hydrolase isoform AH I precursor (AAA93234) [*Prunus serotina*] and cyanogenic β-glucosidase [*Trifolium repens*] 70% identity. The percent identity was calculated based on alignment length, which is a 500 amino acid long. The amino acid sequence of this β-glucosidase also showed 98% identity with un characterized coniferrin β-glucosidase from *L. leucocephala* (Accession no.ABY84677), however the alignment length is 400 amino acids only.

3.3.7.2. Multiple sequence alignment of *Leucaena leucocephala* β-glucosidase aminoacid sequence with other defensive β-glucosidases

Multiple sequence alignment of the Family 1 glycosyl hydrolases amino acid sequences (Just Bio, Fig. 3.23) was done using : Amygdalin hydrolase PsAH (AAA93234)) Prunasin hydrolase PsPH(AF411009) from *Prunus serotina*. Cyanogenic β-glucosidase TrCBG (1CBG) from *Trifolium*, Dalcochinin β-glucosidase DcBglu1

(AAF04007) from *Dalbergia* and Dhurrinase SbDhr1 (AAC49177) from *Sorghum bicolor* along with *L. leucocephala* β-glucosidase (ABY48758). GenBank aceesion numbers are given in bracket.

Leucaena β-glucosidase (Llbglu1) contains several sequence elements that are highly conserved among many family 1 glucosidases. Between residues 41 and 52 it carries the N-terminal signature sequence F, X, (FYWM), (GSTA), X, (GSTA), X, (GSTA), (GSTA), (FYN),X,E,X(GSTA) as characteristic of family 1 glycosyl hydrolases (Henrissat, 1991). Two of the four cysteine residues found in Llbglu1 (Cys-218 and Cys-271) are also conserved in these homologous β-glucosidases, suggesting that they may be involved in forming important intra molecular disulfide bridges. Other conserved sequence elements include the sequence -ENG- at residues 413–415 within the C-terminal signature, and the sequence -NEP- at residues 198–200 (Figure 3.23). These sequence motifs are thought to be important for enzyme activity (Baird *et al.*, 1990; Withers *et al.*, 1990) and this region may be involved in binding of the pyranose ring during catalysis (Baird *et al.*, 1990). The hydrolytic mechanism of the family 1 β-glucosidases is considered to be general acid catalysis (Sinnott, 1990) with Glu and Asp residues in conserved motifs serving as active-site nucleophile and acid catalyst. Evidence from inhibitor and site-directed mutagenesis studies strongly suggest the Glu-413 within the conserved ENG motif to be the active-site nucleophile (Withers *et al.*, 1990; Trimbur *et al.*, 1992). This carboxylate side-chain has been suggested to play the role of acid-base catalyst during hydrolysis of the glycosidic linkage (Trimbur *et al.*, 1992). Multiple sequence alignment of aminoacid sequences of β-glucosidases involved in defense along with *Leucaena leucocephala* β-glucosidase (ABY48758) is shown in Fig.3.23.

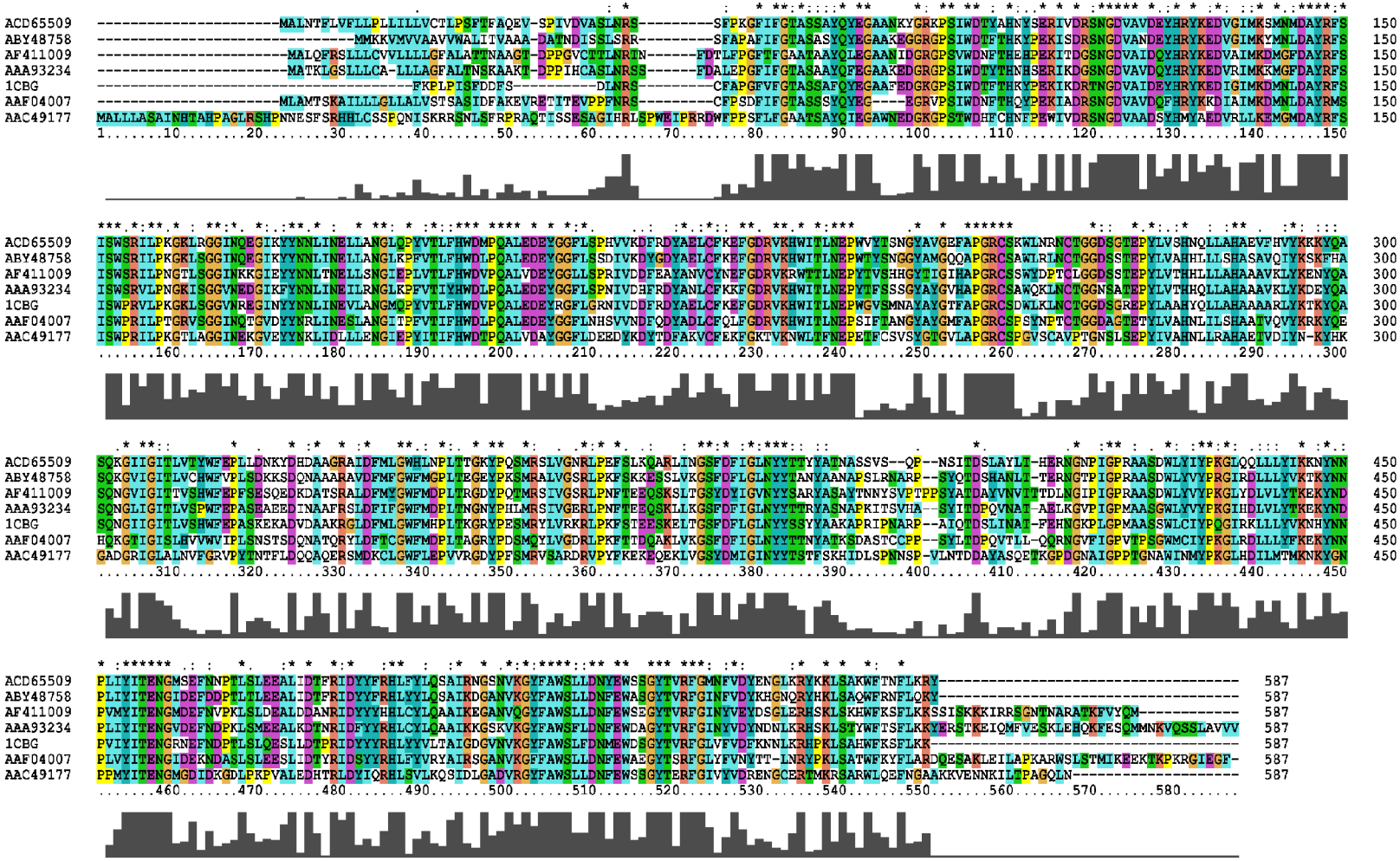

ACD65509
ABY48758
AF411009
AAA93234
1CBG
AAF04007
AAC49177

Fig.3.23

3.3.7.3 Analysis of amino acid sequence of the *Leucaena* β-glucosidase for conserved domain

Amio acid sequence of Llbglu1 was searched against the sequences at the NCBI database at http://www.ncbi.nlm.nih.gov/blast/Blast.cgi. In the protein a Putative conserved domains has been detected and it was found to be a putative glycosyl hydrolase family 1 conserved domain (Fig 3.24).

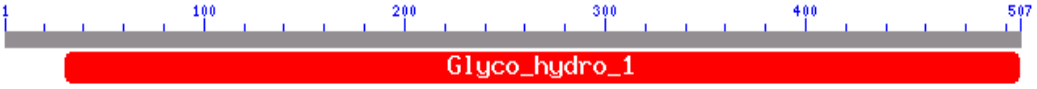


Fig.3.24.Image showing the putative glycosyl hydrolase conserved domain.

3.3.7.4. Theoretical molecular weight (MW) and Isoelectric Point of the β-glucosidase.

Theoretical molecular weight (MW) and Isoelectric Point of the Llbglu1 were analyzed at http://www.expasy.ch/cgi-bin/protparam and found to be 57 KD and 8.33 respectively.

3.3.7.6. Signal sequence prediction of the *Leucaena* β-glucosidase and Hydropathy plot

Signal sequence in the β-glucosidase was predicted at http://www.cbs.dtu.dk/cgi-bin using neural networks (NN) trained on eukaryotes and most likely cleavage site found between amino acid position 21 and 22 (Fig 3.25)

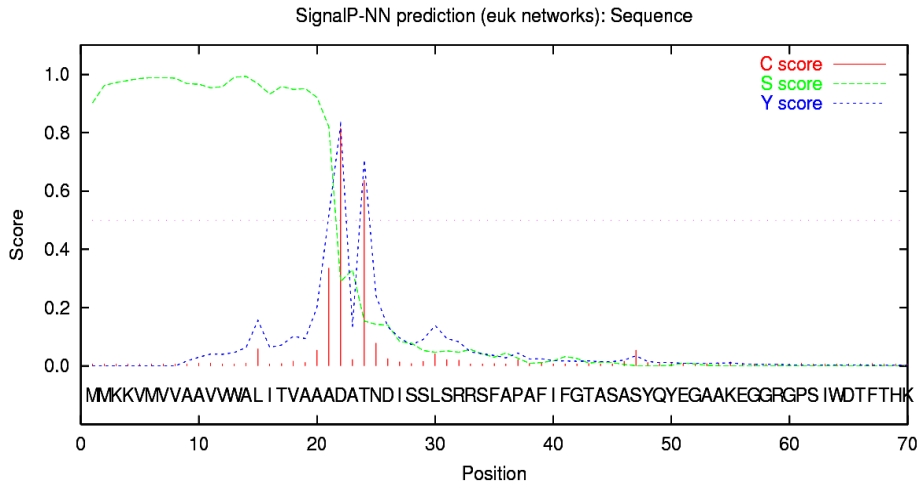


Fig 3.25. Image showing the probable cleavage site at signal peptide pos. 21 and 22.

When the Hydropathy plot of the *Leucaena* β-glucosidase was analyzed using Kyte-Doolittle Hydropathy plot at http://gcat.davidson.edu/rakarnik/kd.cgi it is having a transmembrane sequence at the N-terminal end supporting the view that of being a signal sequence (Fig. 3.26).

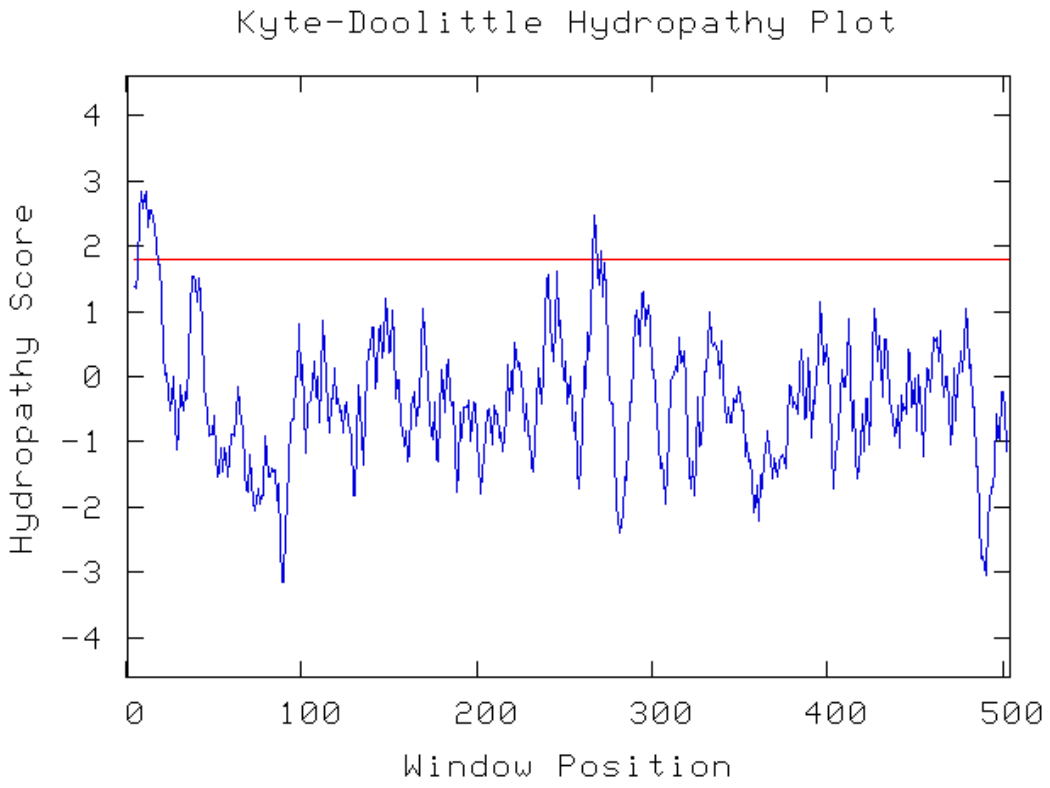


Fig 3.26. Hydropathy plot of the *Leucaena* β-glucosidase

3.3.7.7 Phylogenetic and functional relationship of *Leucaena leucocephala* glycosyl hydrolase with other Family 1 glycosyl hydrolases.

To determine the relationship of *L. leucocephala* glycosyl hydrolase with other glycosyl hydrolases genes, phylogenetic analysis was undertaken. Such an approach found to be fruitful in defining the functional relationship among *Arabidopsis BGLU* genes (Xu *et al* 2004, Escamilla-trevino *et al* 2006). Xu *et al* constructed a phylogenetic tree with all the 47 *Arabidopsis* BGLU genes along with β-glucosidases from other sources (Chapter

1, Section 1.1.1.1, Fig1.1) *BGLU* 44 clustered with mannosidases in phylogenetic tree later it was proved that it has substrate specificity for mannose. Similarly BGLU 45 and *BGLU*46 cluster with coniferrin β-glucosidases when these genes heterologously expressed, they hydrolysed monolignals. We hypothesize that the closely clustered members exhibit similar, if not identical, substrate specificities. A phylogenetic tree depicting the relationships between GH Family 1 members from different species along with *Leucaena leucocephala* β-glucosidase (Llbglu1) was constructed (Fig 3.27). Amino acid sequences of members of the glycoside hydrolase family 1 used for alignment and construction of phylogenetic tree along with *Leucaena leucocephala* β-glucosidase are given in table 3.2.In the phylogenetic tree all myrosinases clustered into a branch supporting the above hypothesis. These hydrolases are unique in having a glutamine residue to replace the acid/base catalyst glutamate shown by the other family members. Similarly all the mannosidases are grouped together whereas all monocot defence β-glucosidases fall into a branch. In the phylogenetic tree the *L. leucocephala* glycosyl hydrolase (Llbglu1) closely placed with the other dicot defensive β-glucosidases (Fig.3.27). The putative biological role of this β-glucosidase may be defense in function. However it needs further study to prove the hypothesis.

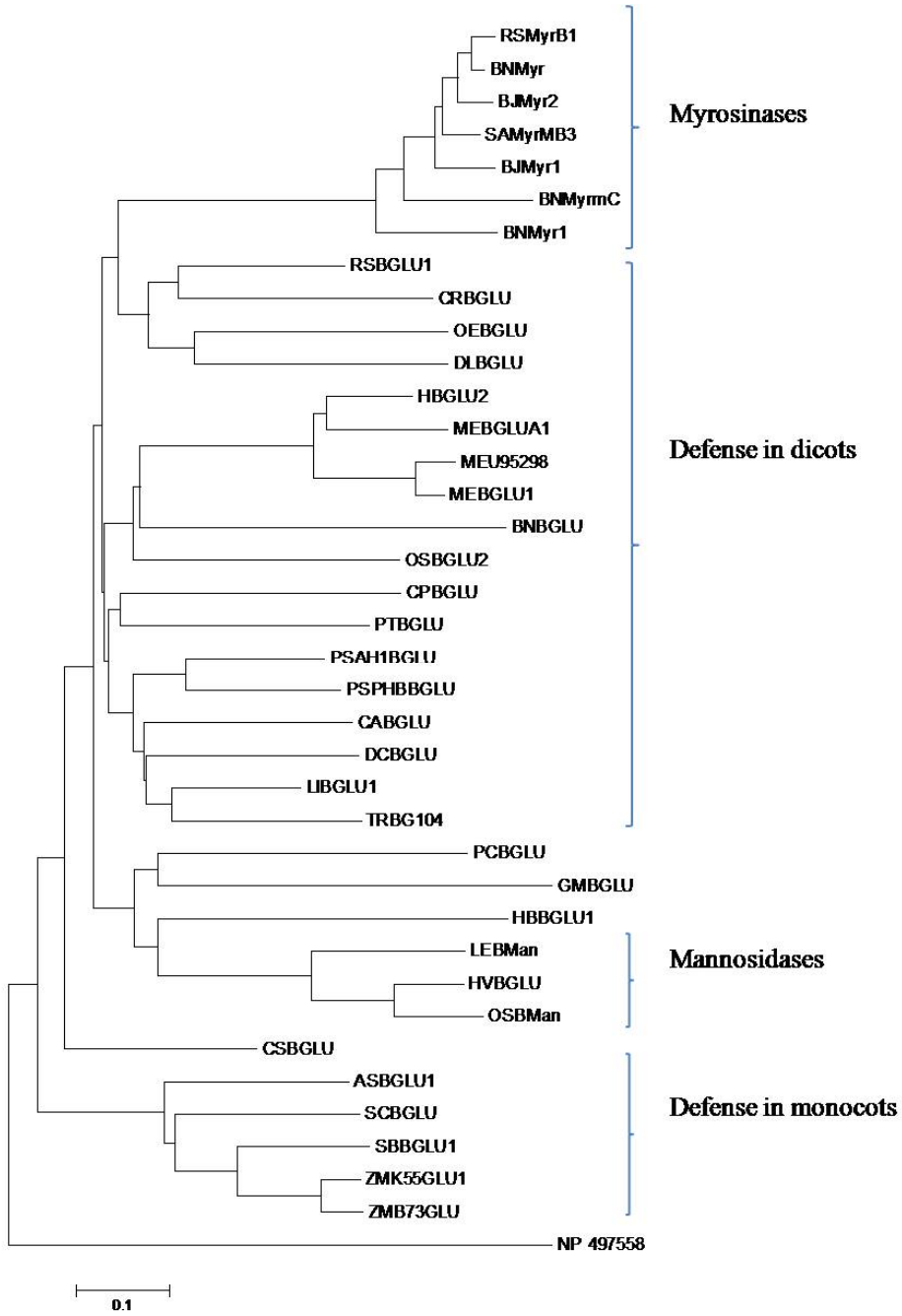


Fig:3.27 . A phylogenetic tree depicting the relationships between GH Family 1 members from different species. This neighbor-joining tree was constructed using the program PAUP 4.0b2 after alignment of amino acid sequences by ClustalX followed by manual editing. The *C. elegans* β-glucosidase (NP_497558) was used to root the tree. All branches are drawn to scale as indicated by the scale bar (= 0.05 substitutions/site).

Table 3.2 Amino acid sequences of members of the glycoside hydrolase family 1 used for alignment and construction of phylogenetic tree along with *Leucaena leucocephala* β-glucosidase

S.No.	**Symbol**	**Enzyme**	**Organism**	**Genbank accession no.**
1	BNMyr	Myrosinase	*Brassica napus*	CAA42775
2	RSMyrB1	Myrosinase	*Raphanus sativus*	BAB17226
3	BJMyr2	Myrosinase,	*Brassica juncea*	CAA11412
4	SAMyrMB3	Myrosinase	*Sinapis alba*	CAA42534
5	BJMyr1	Myrosinase	*Brassica juncea*	AAG54074
6	BNMyrmC	myrosinase	*Brassica napus*	CAA55685
7	BNMyr1	Myrosinase,	*Brassica napus*	CAA79989
8	BNBGLU	B-glucosidase	*Brassica napus*	CAA57913
9	ASBGLU1	B-D-glucosidase	*Avena sativa*	CAA55196
10	SCBGLU	B-glucosidase	*Secale cereal*	AAG00614
11	ZMK55GLU1	B-D-glucosidase	*Zea mays*	AAB03266
12	ZMB73GLU	B-glucosidase	*Zea mays*	AAD09850
13	SBBGLU1	Dhurrinase	*Sorghum bicolor*	AAC49177
14	LEBMan	B-mannosidase	*Lycopersicon esculentum*	AAL37714
15	HVBGLU	B-glucosidase	*Hordeum vulgare*	AAA87339
16)	OSBMan	B-glucosidase	*Oryza sativa*	BAC20802
17	HBBGLU1	Latex cyanogenic β glucosidase	*Hevea brasiliensis*	AAP51059
18	PCBGLU	B-glucosidase	*Pinus contorta*	AAC69619

19	MEBGLU1	Linamarase	*Manihot esculenta*	AAB22162
20	MEBGLUA1	B glucosidase	*Manihoe esculenta*	CAA64442
21	HBGLU2	P66 protein	*Hevea brasiliensis*	AAO49267
22	OSBGLU2	B-glucosidase	*Oryza sativa*	AAL14713
23	CSBGLU	Furostanol glycoside 26-O-β-glucosidase	*Costus sp*	BAA11831
24	PSPHBBGLU	Prunasin hydrolase PHB	*Prunus serotina*	AAL39079
25	PSAH1BGLU	Amygdalin hydrolase isoform AH I	*Prunus serotina*	AAA93234
26	TrCBG	Cyanogenic B-Glucosidase	*Trifolium repens*	1CBG
27	CABGLU	B-glucosidase	*Cicer arietinum*	CAC08209
28	DCBGLU	Dalcochinin 8'-O-β-glucosidase	*D.cochinchinensis*	AAF04007
29	PTBGLU	B-glucosidase	*Polygonum tinctorium*	BAA78708
30	DLBGLU	Cardenolide16-glucohydrolase	*Digitalis lanata*	CAB38854
31	OEBGLU	B-glucosidase	*Olea europaea*	AAL93619
32	CRBGLU	Strictosidine β-glucosidase	*Catharanthus roseus*	AAF28800
33	RSBGLU1	Raucaffricine-O-β-D-glucosidase	*Rauvolfia serpentina*	AAF03675
34	CPBGLU	Silverleaf whitefly-induced protein 3	*Cucurbita pepo*	AAG25897
35		hypothetical protein E02H9.5	*Caenorhabditis elegans*	NP_497558
36	GMBGLU	Hydroxyisourate hydrolase	*Glycine max*	AAL92115
37	LlBGLU1	Glycosylhydrolase 1	*Leucaena leucocephala*	ABY48758

3.3.7.8 Phylogenetic analysis of plant β-glucosidases involved in the bioactivation of defense compounds

A phylogenetic tree was constructed after multiple sequence alignment of plant b-glucosidases involved in the bioactivation of defense compounds along with *L. leucocephala* glycosyl hydrolase. In the phylogenetic tree, the β-glucosidases known from the literature to hydrolyze α-hydroxynitrile glucosides form separate clades in monocotyledons and eudicotyledons. This argues that the ability to hydrolyze α-hydroxynitrile glucosides has evolved independently in monocotyledons and eudicotyledons (Morant *et al* 2008).Phylogenetic analysis reveals that *L. leucocephala* β-glucosidase clustered with Hydroxynitrile Glucoside-Cleaving β-Glucosidases LjBGD2, LjBGD4 and LjBGD7 from *L. japonicas* and TrCBG, a β-glucosidase from white clover that hydrolyzes linamarin (Fig 3.28). In agreement with the phylogenetic analysis provided by Chuankhayan et al. (2007b), the isoflavonoid glucoside-cleaving β-glucosidases from *Glycine max* and *Dalbergia cochinchinensis* were found to cluster between hydroxynitrile glucoside-specific *β*-glucosidases from legumes and *P. serotina*. Based on the phylogenetic analysis, isoflavonoid glucosidecleaving β-glucosidases in legumes appear to have evolved from the β-glucosidases involved in the hydrolysis of hydroxynitrile glucosides. From this analysis it can be inferred that *L. leucocephala* glycosyl hydrolases (lu1Llbg) is involved in defense probably hydrolysig the Hydroxynitrile Glucosides.

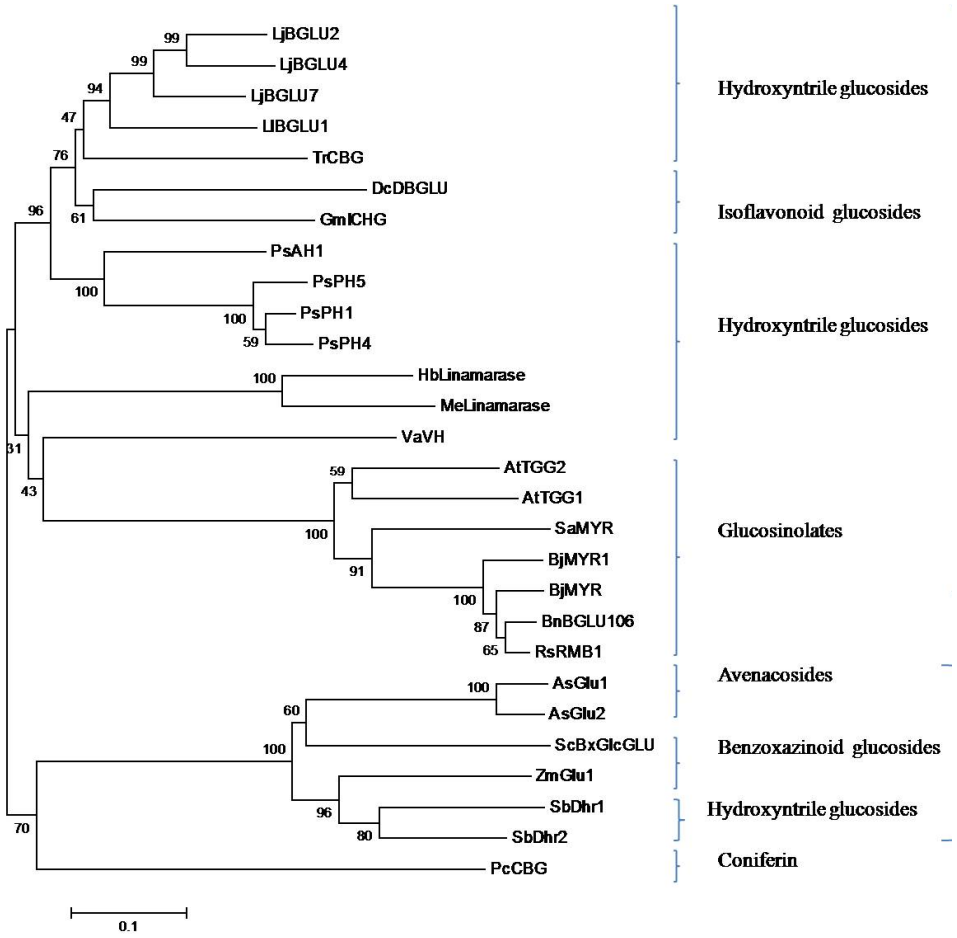


Fig. 3.28. Phylogenetic analysis of selected plant β-glucosidases involved in the bioactivation of defense compounds. The phylogenetic tree includes hydroxynitrile and isoflavonoid glucoside-cleaving β-glucosidases from eudicotyledons, glucosinolate degrading myrosinases (Brassicales), and selected β-glucosidases involved in the bioactivation of defense compounds in monocotyledons. The amino acid sequences used for the phylogenetic tree are shown in Table 3.3.

Table 3.3. Amino acid sequences of members of the glycoside hydrolase family 1 used for alignment and construction of phylogenetic tree along with *Leucaena leucocephala* β-glucosidase

S.No	Symbol	Enzyme	Organism	GenBank Accession No
1	AtTGG2	GLUCOSIDE GLUCOHYDROLASE2	*Arabidopsis thaliana*	NP_568479

2	AtTGG1	THIOGLUCOSIDE GLUCOHYDROLASE 1	*Arabidopsis thaliana*	NP_851077
3	AsGlu1	β-D-glucosidase	*Avena sativa*	CAA55196
4	AsGlu2	β-D-glucosidase	*Avena sativa*	AAD02839
5	BjMYR1	myrosinase	*Brassica juncea*	AAG54074
6	BjMYR	myrosinase	*Brassica juncea*	CAA11412
7	BnBGLU106	myrosinase	*Brassica napus*	CAA42775
8	VaVH	vicianin hydrolase	*Vicia sativa* subsp. nigra	ABD03937
9	LjBGLU2	β-glucosidase D2	*Lotus japonicus*	ACD65510
10	LjBGLU4	β-glucosidase D4	*Lotus japonicus*	ACD65509
11	LjBGLU7	β-glucosidase D7	*Lotus japonicus*	ACD65511
12	PCBGLU	B-glucosidase	*Pinus contorta*	AAC69619
13	PsPH1	prunasin hydrolase isoform PH I	*Prunus serotina*	AAA93032
14	PsPH4	prunasin hydrolase isoform PH B precursor	*Prunus serotina*	AAL39079
15	PsPH5	prunasin hydrolase isoform PH C precursor	*Prunus serotina*	AAL35324
16	PsAH1 precursor	amygdalin hydrolase isoform AH I	*Prunus serotina*	AAA93234
17	RsRMB1	myrosinase	*Raphanus sativus*	BAB17226
18	RsRMB1	myrosinase	*Raphanus sativus*	BAB17227
19	ScBxGlcGLU	β-glucosidase	*Secale cereale*	AAG00614
20	ZmGlu1	β glucosidase1	*Zea mays*	NP_001105454
21	TrCBG	linamarase	*Trifolium repens*	1CBG
22	SbDhr1	dhurrinase	*Sorghum bicolor*	AAC49177
23	SbDhr2	dhurrinase-2	*Sorghum bicolor*	AAK49119
24	SaMYR	myrosinase	*Sinapis alba*	1MYR_A
25	MeLinamarase	linamarase	*Manihot esculenta*	AAB22162
26	HbLinamarase	β glucosidase	*Hevea brasiliensis*	ABL01537
27	DcDBGLU	dalcochinin 8'-O-β-glucoside β-glucosidase	*Dalbergia cochinchinensis*	AAF04007
28	GmICHG	isoflavone conjugate-specific β-glucosidase	*Glycine max*	BAF34333
29	LlBGLU1	Glycosylhydrolase 1	*Leucaenaleucocepha*	ABY48758

3.4. Conclusions

A PCR based approach has been used here to isolate a family 1 glycosyl hydrolase gene from *Leucena leucocephala*. Primers were designed from the selected Family 1 glycosyl hydrolases that are involved in defense, from the NCBI database. Leucaena gDNA and cDNA was used to amplify partial sequences of the β-glucosidase. A 390 bp sequence was amplified with gDNA as template. This sequence was having two introns. It showed homology with known Family1 glycosyl hydrolases and maximum homology (79% identity) with non-cyanogenic β-glucosidase from *Cicer (*Accession number CAG14979) The sequence was submitted to NCBI GenBank database under accession number (DQ417200). Using cDNA as template 141 bp sequence was amplified. The coding region was translated into amino acid sequence using program Expasy translate tool. The amino acid sequence was searched against non-redundant protein sequences at NCBI using hprogram BLAST. It showed significant homology (81%) with *Trifolium* cyanogenic β-glucosidase (CAA40057).The sequence was submitted to NCBI GenBank database and later updated (accession number DQ883447).

Rapid Amplification of cDNA Ends (RACE, 3' RACE and 5' RACE) was done to obtain the ful-llength cDNA clone of the β-glucosidase.3'RACE has resulted into amplification of ~1.2 kb fragment, it was cloned into pGEMT-Easy vector. Nucleotide sequence of the clone (c1.3KB) has revealed that it has a coding sequence of 1183 bp followed by stop codon TAA. The sequence has a 3'untranslated region (3'UTR) of 118 bp followed by a poly A tail. The nucleotide sequence was translated in to amino acid sequence; the latter was searched in BLAST in NCBI database. It was found to contain a putative conserved domain found in Family 1 Glycosyl hydrolases. Conserved catalytic motifs ENG and NEP characteristics of this family were also found. The sequence was submitted to NCBI GenBank database (accession number DQ883447). In 5'RACE an amplicon of approximately 1.0kb was amplified and cloned. The nucleotide sequence of the clone (c5CBGR1) was analyzed using bioinformatics tools. A 18 bp 5'UTR was observed followed by start codon ATG. Deduced amino acid sequence has homology with Family 1 glycosyl hydrolase in NCBI database. The sequence was submitted to NCBI GenBank database (Accession no. EF061245**).** Gene specific forward and reverse primers were designed from the start and stop codon of 5'and

3'RACE sequences respectively. PCR was done using cDNA as template. An expected size of approximately 1.5 kb amplicon was amplified. The amplicon was cut, eluted from gel and ligated into pGEMT-Easy vector. It was confirmed by restriction digestion with *Eco*RI which released 1.5 kb fragment (Fig.3.18). The clone now referred as cLlbglu1I was sequenced. The sequence showed 77% identity with lotus β-glucosidase (EU710846) and 78% identity with Rose β-glucosidase (AB426520). The sequence was submitted to NCBI GenBank database (EU328158).

Slot blot hybridization showed the presence of at least two copy number of the gene. Multiple sequence alignment of nucleotide and deduced amino acid sequence of Llbglu1 with other Family1 GHs involved in defence was done using ClustalX. Amio acid sequence of Llbglu1 was searched against the sequences at the NCBI database at http://www.ncbi.nlm.nih.gov/blast/Blast.cgi. In the protein a Putative conserved domains has been detected and it was found to be a putative glycosyl hydrolase family 1 conserved domain. Theoretical molecular weight (MW) and Isoelectric Point of the Llbglu1 were analyzed at http://www.expasy.ch/cgi-bin/protparam and found to be 57 KD and 8.33 respectively. Signal sequence in the β-glucosidase was predicted at http://www.cbs.dtu.dk/cgi-bin using neural networks (NN) trained on eukaryotes and most likely cleavage site found between amino acid position 21 and 22. When the Hydropathy plot of the *Leucaena* β-glucosidase was analyzed using Kyte-Doolittle Hydropathy plot at http://gcat.davidson.edu/rakarnik/kd.cgi it is having a transmembrane sequence at the N-terminal end supporting the view that of being a signal sequence. Phylogenetic and functional relationship of *L. leucocephala* glycosyl hydrolase with other Family 1 glycosyl hydrolases revealed that it is a β-glucosidase involved in defense. It clustered with *Lotus japonicas* β-glucosidases which hydrolyses the hydroxynitrile glucosidases.

4.1. Introduction

Carbohydrates and their glycoconjugates are one of the most diverse groups of organic molecules in the biosphere. The selective cleavage of glycosidic bonds is crucial in a variety of fundamental biological processes for all living organisms (Verdoucq *et al*., 2004). β-Glucosidases (E.C. 3.2.1.21) represent a group of ubiquitously expressed, hydrolytic enzymes, which catalyze the hydrolysis of β-*O*-glucosidic linkages between β- D-glucose and an aglycone or another sugar (Esen, 1993). β-Glucosidases exhibit similar specificity for a β-glucoside substrate, but distinct specificities for the aglycone linked to the glucosyl group (Esen, 1993), suggesting their diverse biological functions. In plant physiology, β-glucosidases are implicated in growth regulation, stress response, cellobiose degradation, lignification, and defence (Cairns *et al*., 2000).

Naturally occurring glucosidic substrates contain a broad range of aglycone groups, including cyanogenic glucosides (Barette *et al*., 1995; Eksittukul and Chulavatanatol 1988), cellobiose (Ferriera and Terrra, 1983), phenolic glucosides (Podstolski and Levak, 1970), thioglucosides (Durham and Poulton, 1989), and isoflavonoid glucosides (Svasti *et al*., 1999). To date, 3D-structures have been solved for 10 GH Family 1 enzymes, including four from plant sources (Barrett *et al*., 1995; Burmeister *et al*., 1997; Czjzek *et al*., 2000; Verdoucq *et al*., 2004). Although they share only 17%–70% sequence identity, all 10 enzymes exhibit essentially the same $(\beta/\alpha)_8$-barrel fold, in which the active site is located at the barrel's C terminus (Xu *et al*., 2004).

The identification of β-glucosidases, their substrates, and the nature of their interactions will not only shed light on the structure and function of the enzymes, but also help define their biological significance *in vivo*. To investigate the molecular basis of enzymatic catalysis and to engineer it to achieve desirable catalytic properties, a reliable expression and purification system was developed in this study.

4.2.Materials and methods

4.2.1. Cloning of *Leucaena leucocephala* β-glucosidase into prokaryotic expression system

The pET-28 a (+) vector (Novagen, USA) was used for expression of LlBGLU1 gene in *E.coli* BL 21 (DE3). This vector carries a N-terminal His•Tag®/thrombin/S•Tag™/

enterokinase configuration plus an optional C-terminal His•Tag sequence. This vector has unique restriction sites for cloning ease (Fig. 4.1). The cloned gene is expressed under the T7 RNA polymerase promoter.The signal sequence from the clone Llbglu1 was eliminated and restriction sites *Eco*RI and *Not* I were incorporated into forward and reverse primer respectively.

Forwad primer NPF: 5'GAATTCGATGCAACAAATGATATTTCC 3'

Reverse primer NPR: 5'GCGGCCGCATATTTTGAAGGAAGTTCCTG 3'

PCR was performed with pLlbglu1 as template. A 1.4 Kb fragment was amplified. The amplicon was ligated into pGEMT vector. Positive clone was sequenced. The insert was released by restriction digestion with *Eco*RI and *Not* I and ligated into pET-28a+ vector (Fig 4.1), which is also cut with the same restriction enzymes. The ligation mixture was transformed into *E.coli* XL1. The positive clones were sequenced for confirmation of His tag.

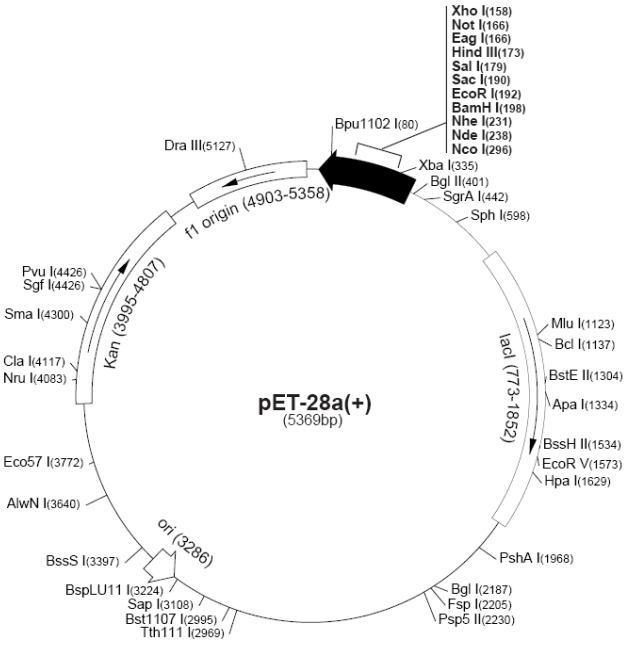


Fig.4.1. pET 28 a (+) vector map with multiple cloning sites.

4.2.2. Heterologous expression and purification of *Leucaena* β-glucosidase in *E.coli*

The plasmid (pET-Llbglu-His) was mobilized in to *E.coli* BL 21(DE3). Individual colonies were screened for protein expression. The positive colonies were grown in LB media for protein expression for 4-6 h at 37^{o}C and cells were harvested and lysed, both soluble and in-soluble fractions were analysed on SDS-PAGE.

Optimization of culture conditions:

Culture conditions like temperature, IPTG concentration and Period of incubation were optimized so that expressed protein was in soluble form. In the crude cell lysate the enzyme activity was monitored by assaying for glucosidase activity using PNPG as a substrate

Purification of recombinant enzyme:

Leucaena β-glucosidase was heterologously expressed with 6X Histidine as fusion protein. The recombinant enzyme was purified under native conditions by Immobilized Metal Affinity chromatography, using Ni-NTA agarose beads (Qiagen, USA) as described in Chapter 2, Section 2.8.7.2. The concentration of purified enzyme was determined by Lowry's methods using BSA as standard.

4.2.3. Optimization of enzyme assay conditions for the recombitant *Leucaena* β-glucosidase

The enzyme was assayed spectrophotometrically using p-nitrophenol-β-D-glucopyranoside (PNPG). Appropriately diluted enzyme was incubated with substrate (1mM PNPG) in a final volume of 500 µL at different temperatures (30-55^{o}C) and pH (3-7). The time drive also studied by incubating at different time intervals (10 min-100 min). After incubation the reaction was terminated by the addition of 1.0 ml 1.0 M Na_2CO_3. The paranitrophenol liberated was read ,as the phenolate anion at 420 nm. The concentration of p-nitrophenol was determined using a molar absorption co-efficient of $1.77X10^4$. One unit of the enzyme is defined as the amount of enzyme that liberates 1 µmol of p-nitrophenol /min under the assay conditions.

4.2.3.1. Optimization of pH

The effect of pH was studied by monitoring the enzyme activity at various pH (3.0-7.0) in universal buffers using PNPG as substrate at room temperature (37^{o}C).

4.2.3.2. Optimization of Temp

The effect of temperature was studied by monitoring the enzyme activity at various temperatures (30-55^{o}C), in 100 mM-200 mM Citrate-phosphate buffer pH 4.8, containing 150 mM NaCl using PNP-Glucoside.

4.2.3. 3. Optimization of Time drive

To find the linear range of the time drive, enzyme assays were carried out at 45^{o}C in 100 mM-200 mM Citrate-phosphate buffer, pH 4.8 monitoring the activity at each 10 min for 100 min using pNPglucoside as substrate.

4.2.4. Determination of Glycone specificity of the recombinant enzyme

The glycone and linkage specificities of recombinant Llbglu1 were examined by assaying their activities towards several β- and α-linked p-nitrophenyl (pNP) and ortho-nitrophenyl sugars (Table 4.1).

Table 4.1. β- and α-linked p-nitrophenyl (pNP/oNP)-sugars

S.No	Name of PNP sugar
1	p-Nitrophenyl β-D-glucopyranoside
2	p- Nitrophenyl β-D-glucoronide
3	p-Nitrophenyl-N-acetyl-1-thio-β-glucosaminide
4	p-Nitrophenyl α-D-glucopyranoside
5	p-Nitrophenyl N-acetyl-β-D-glucosaminide
6	p-Nitrophenyl β-D-galactopyranoside
7	p-Nitrophenyl β-D-mannopyranoside
8	p-Nitrophenyl β-D-xylopyranoside
9	p-Nitrophenyl β-L-arabinopyranoside
10	o-Nitrophenyl β-D-glucopyranoside
11	p-Nitrophenyl N-acetyl-α-D-glucosaminide

4.2.5. Determination of *Km* and *Vmax* for glycosides

Different substrates, PNPG sugars (table 4.1) were tested for activity using purified recombitant Llbglu1. The enzyme hydrolysed PNP-glucoside and PNP-galactoside. To find the glucone specificity , kinetic constants like *Km* and *Vmax* were determined for these substrates.

4.2.5.1. Determination of *Km* and *Vmax* for PNP-Glucoside

The kinetic rate constants, *Km* and *Vmax* of the PNP-glucoside were determined at 45^{o}C in 100 mM-200 mM Citrate-phosphate buffer, pH 4.8. A fixed dilution of enzyme *i.e* 10 µL (0.48mg/mL) was incubated with varying concentration of the substrate.

4.2.5.2. Determination of *Km* and *Vmax* for PNP-Galactoside

The kinetic rate constants, *Km* and *Vmax* of the PNP-galactoside were determined at 45^{o}C in 100 mM-200 mM Citrate-phosphate buffer, pH 4.8. A fixed dilution *i.e* 20 µL (0.12 mg/mL) of the enzyme was incubated with varying concentration of the substrate.

4.2.6. Determination of aglycone preference of the recombinant enzyme by synthetic glucosides

Two glucosides one chromogenic substrate, VRA-G (Ammonium 5-[4-(β-D-glucopyronosyloxy)-3-methoxymethylene]-2-thioxothiazolidin-4-one-3-acetate and a fluorogenic substrate, 4-Methyl-umbellifryl-β-D-glucopyranoside (4MUG) were used for hydrolysis under standard assay conditions.

4.2.6.1. Determination of *Km* and *Vmax* for VRA-G

The kinetic rate constants, *Km* and *Vmax* of the VRA-glucoside at at 45^{o}C in 100 mM-200 mM Citrate-phosphate buffer, pH 4.8. A fixed dilution of enzyme *i.e* 3.0 µL (0.52mg/mL)was incubated with varying concentration (0.5 mM-2.0 mM) of the substrate. The liberated red coloured phenol is measured at 490nm (ε 38600, pH 9.5).

4.2.6.2. Determination of *Km* and *Vmax* for 4MUG

The kinetic rate constants, *Km* and *Vmax* of the 4MUG at 45^{o}C in 100 mM-200 mM Citrate-phosphate buffer, pH 4.8. A fixed dilution of enzyme *i.e* 20 µL (0.02 mg/mL) (1 µg) was incubated with varying concentration (10 µM-200 µM) of the substrate. The excitation was done at 360nm and fluorescence was mearused at 448 nm. The 4-Methylumbelliferone standard was used for determination of amount of liberated aglucone.

4.2.7. Bioinformatics analysis of the *Leucaena* β-glucosidase

Glycoside hydrolases (GH; EC 3.2.1) catalyze the selective hydrolysis of glycosidic bonds between carbohydrates and non-carbohydrate moieties. Based on amino acid sequence similarities, GHs are currently classified into 112 families, as described in the CAZy database (http://www.cazy.org) (Coutinho and Henrissat, 1999). Glycoside Hydrolase (GH) Family 1 typically hydrolyze substrates of the type G-O-X (or G-S-X), where G-represents glycone and X-represents aglycone moiety. The nature of the aglycone moiety of substrates is believed to be critical for the physiological functions of these enzymes. Of the 16 members of GH1 with crystal

Structures five are from plant sources (Barrett *et al*., 1995; Burmeister *et al*., 1997; Czjzek *et al*., 2000; Verdoucq *et al*., 2004, Barleben *et al*., 2008, Hill *et al*., 2008). There have been a sufficient number of crystal structures of GH1 enzymes gathered to clearly establish the links between active-site amino acid residues and ligand components. So, we decided to use this data and computational methods such as secondary structure prediction, homology modeling and structural superimposition, in an attempt to better understand the catalytic properties of *Leucaena* β-glucosidase enzyme and to identify its amino acid residues that help to determine its specificities. Docking studies were performed to predict its natural substrate in vivo. This computational project is not accompanied by experimental verification of its predictions, a totally separate endeavor.

4.2.7.1. Homology Modeling of *L. leucocephala* β-glucosidase

The homologous structure of *Leucaena leucocephala* β-glucosidase (Llbglu1) was searched in Protein Data Bank using NCBI Blast program. The cyanogenic β-glucosidase from white clover (T. repens, PDB ID: 1CBG) was found to be closest to Llbglu1, with 70% identity at the amino acid sequence level. The structure of white clover β-glucosidase is used as template to generate the 3D model of the Llbglu1 using the program Modeler (Marti-Renom *et al*., 2000). The modeled structure of Llbglu1 is visualized using program PyMoL (DeLano, WL; 2002).

4.2.7.2. Checking the quality of the model with PROCHECK

The quality of the 3D model of Llbglu1 was examined using the program PROCHECK (Laskowski *et al*., 1993). The Ramachandran (φ, ψ) plot for Llbglu1 is shown in figure 4.16 (Ramachandran and Sasisekhara, 1968).

4.2.7.3. Secondary structure calculation and comparison between *L. leucocephala* β-glucosidase and Trifolium CBG

Pair wise sequence alignment of modeled *L. leucocephala* β-glucosidase (Llbglu1) is done with *Trifolium* cyanogenic β-glucosidase (1CBG) using ClustalW. Secondary structural elements (α helices, β-sheets and turns) of Llbglu are compared with the already known 1CBG. The figure showing the secondary structural elements was generated using web-based program ESPRIPT (Gouet *et al*., 1999).

4.2.7.4. Structural super imposition *L. leucocephala* β-glucosidase (Llbglu1) with *Trifolium* CBG

Pair wise structural superimposition of modeled *Leucaena* β-glucosidase was done with *Trifolium* cbg using combinatorial extension algorithm at SDSC-CE (http://cl.sdsc.edu/ce.html) (Shindyalov, 1998). Overall view of secondary structures and the segment containing residues forming the active site of Llbglu1 and Tr CBG was presented.

4.2.7.5. Analysis of amino acids involved in Glycone binding and catalysis in *Leucaena* β- Glucosidase with other β- glucosidases

CE-MC multiple protein structural alignment at http://pathway.rit.albany.albany.edu/~cemc) has been used for the aligning 10 Family 1 GHs structures from different sources (mostly from plant origin)along with the modeled Llbglu1. The program CE-MC uses C-alfa co-ordinates for superposition using combinatorial extension (CE) and Monte Carlo optimization methods (Guda *et al*., 2001, 2004). By multiple structural alignment, the conserved residues of the binding pocket involved in glycose binding and catalysis in Llbglu1 were observed.

4.2.7.6. Docking of Rhodocyanides into *Leucaena* β-glucosidase

In phylogenetic tree *L. leucocephala* β-glucosidase clustered with *Lotus* β-glucosidases which are involved in defense by hydrolyzing the hydroxynitrle glucosides (chapter 3, Section 3.3.7.7, Fig 3.28). Furthermore Llbglu1 showed upto 74% identity and 86% positives with *Lotus* β-glucosidases (LjBGD2 and LjBGD4) at amino acid level. These β-glucosidases exhibited the lowest K_m for rhodiocyanoside A, followed by dhurrin, prunasin, and lotaustralin (Morant *et al*., 2008). So in the present study rhodiocyanoside A is chosen as asubstrate for docking into the binding pocket of Llbglu1.

Molecular docking was performed using the DS Modeling 1.2-SBD Docking Module by Accelrys™ Software (Accelrys Software Inc., 2003). Modeled Llbglu1 (receptor) and Rhodocyanoside A (ligand) were prepared for docking studies with DS Modeling 1.1 SBD module of Accelrys™ Software Package (Accelrys Software Inc., 2003). Structure Data File (SDF) of ligand molecule was downloaded from the pubchem (http://pubchem.ncbi.nlm.nih.gov). SDF format was converted to MOL2 format using online server http://www.webqc.org/molecularformatsconverter.php. Hydrogens were added to receptor as well as to ligand using the built-in program Add Hydrogen in Accelrys™ Software. To generate partial charges on every atom in ligand and receptor CHARMm force field was run for both separately.

In the docking experiment, using Ligand fit docking module, binding site was assigned by finding receptor cavities. The best cavity with 150 Å3 volume was determined as binding site. The docking was performed to find the most stable and favourable orientation of ligand in the cavity. The docking options consisted of the following steps: (i) Monte Carlo options to perform flexible fit, (ii) thresholds for diversity of saved pose (defined by user to 1.5Å to scan through different conformations), (iii) pose optimization done in two steps: (a) steepest descent minimization and (b) BFGS rigid body minimization, (iv) ligand internal energy optimization and filtering poses with short contacts (VDW and electrostatic energy calculated), and (v) pose filtering and processing: dock scores for conformations above energies of 20.0 kcal mol−1 were accepted. Clustering of poses using leader algorithm was done.

Scoring for the docked poses was determined primarily using Ludi score that included five major contributions: (a) contributions from ideal hydrogen bonds, (b) contributions from perturbed ionic interactions, (c) contributions from lipophilic interactions, (d)

contributions due to the freezing of internal degrees of freedom and (e) contributions due to loss of translational and rotational entropy of the ligands. The best docked ligand pose (based on docking score) was used for further In Situ Ligand Minimization. Smart minimizer algorithm was used. Which combines steepest descents (SD-1000 cycles), conjugate gradient (CONJ-1000 cycles) and adopted basis-set Newton-Raphson (ABRN-1000 cycles) methods. In order to check receptor ligand interactions Ligplot (Wallace *et al*., 1995) was generated through PDBSum (Laskowski R A., 2001) on ebi server (http://www.ebi.ac.uk/pdbsum).

The models obtained of the three- dimensional structures of the Llbglu1, with closeup views of the active sites into which rhodiocyanoside A is docked are presented in Figure (fig 4.20).

4.3. Results and discussions

4.3.1. Cloning of *Leucaena leucocephala* β-glucosidase into prokaryotic expression system

A 23 amino acid long signal sequence was predicted in Llbglu 1 by signalP program at C-terminus (Chapter 3, Section: 3.3.7.6). The Llbglu1 was cloned without signal sequence, into pET 28 a(+) vector at *EcoR I* and *Not I* sites. The positive colonies were confirmed by restriction enzyme digestion with *EcoR I* and *Not I* which released ~1.4 kb fragment (Fig.4.2). The plasmid pET-Llbglu-His was confirmed by sequencing.

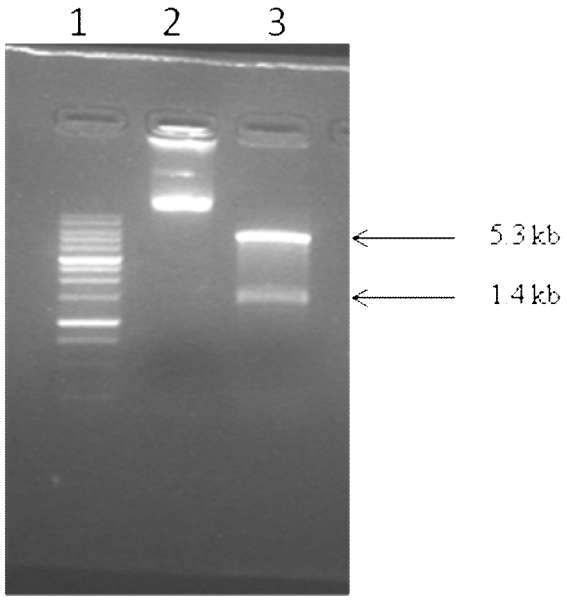


Fig. 4.2. Restriction digestion of plasmid pET-Llbglu-His. Lane 1:DNA size marker, (Lane 2: Undigested plasmid, Lane 3: restriction digestion of the plasmid with enzyme *EcoR* I and *Not* I

4.3.2. Heterologous expression and purification of *Leucaena* β-glucosidase in *E.coli*

The plasmid (pET-Llbglu1-His) was mobilized in to *E.coli* BL 21(DE3). Individual colonies were screened for protein expression. The positive colonies expressing the recombinant protein were grown in LB media for protein expression by 1mM IPTG induction for 4-6 h at 37^{o}C and cells were harvested and lysed, both soluble and in-soluble fractions were analysed on SDS-PAGE. Gradual accumulation of recombinant protein (~55 kD) was observed in insoluble fraction from 1-7 h (Fig.4.3) at the above conditions.

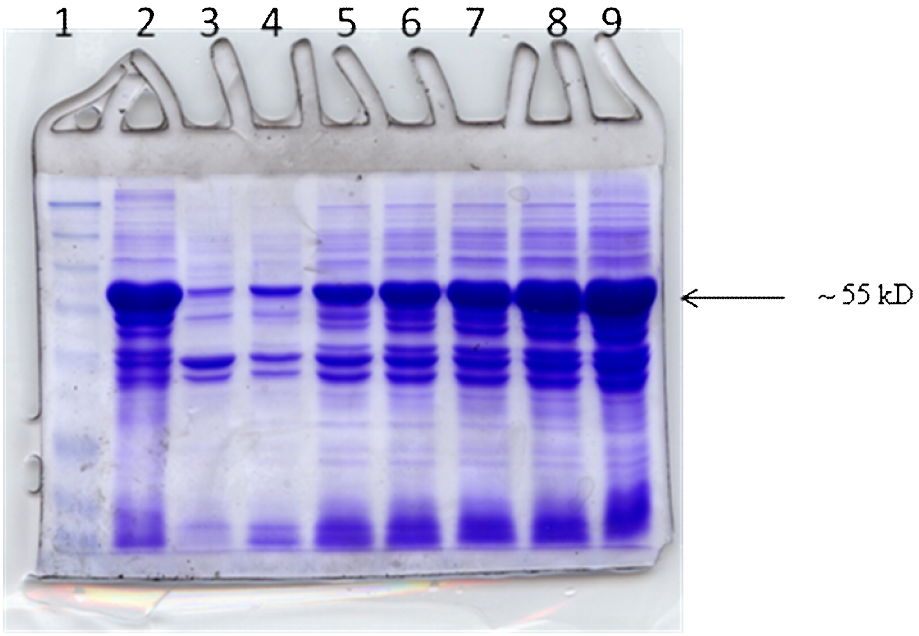


Fig. 4.3. Expression of Llbglu in *E.coli* and SDS-PAGE of inclusion bodies for different time intervals by Coomassie Brilliant Blue-stained gel.

Lane1, Molecular weight marker ;(range7 kD-175 kD) order from top to bottom-175, 80, 58, 46, 30, 25 ,17 and 7kD. (NEB,UK))

Lane 2: inclusion bodies after 4 h culturing. Lane 3-9: inclusion bodies after 1hr-7 h culturing

Optimization of culture conditions

Growth conditions of the culture like temperature, IPTG concentration and period of incubation were optimized so that expressed protein was in soluble form (active form). Activity was monitored in crude lysate by assaying for β-glucosidase activity with PNPG as substrate.

Two temperatures 18^{o}C and 37^{o}C were tested, lower temperature (18^{o}C) found to be favorable for enzyme activity. IPTG was tested from 0.01mM to 1mM to concentration in media and 0.05 mM IPTG found to be optimum for good enzyme activity. IPTG induction was studied for varied period of time *i.e* from 3 h to 12 h at 3 h increment. More enzyme activity observed after 9 h of IPTG induction. The recombinant protein was purified by Immobilized Metal Affinity chromatography, using Ni-NTA agarose

beads (Qiagen, USA). The purified Protein concentration was determined by Lowry's methods using BSA as standard. The purified protein was analyzed on SDS-PAGE and it has a molecular weight of ~55kD (Fig.4.4).

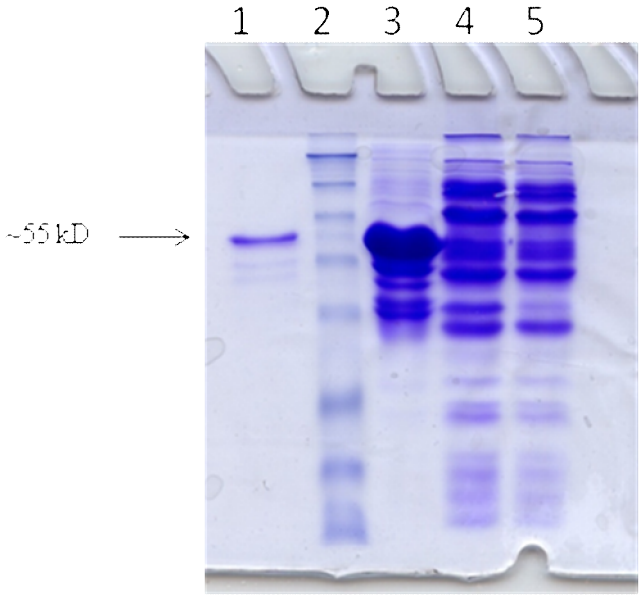


Fig.4. 4. Expression and purifcation of recombinant *Leucaena* β-glucosidase in *E. coli*. SDS-PAGE of Coomassie Brilliant Blue-stained gel. Lane 1: Purified protein, Lane 2: Molecular weight Marker, Lane 3:*E.coli* Lysate with recombinant protein (soluble fraction), Lane 4 and 5 *E.coli* lysate only with pET vector.

4.3.3. Optimization of enzyme assay conditions for the recombinant *Leucaena* β-glucosidase

4.3.3.1. Optimization of pH

Purified *Leucaena* β-glucosidase (~1µg) was assayed at acidic, neutral and alkaline conditions. Activity was found only in acidic conditions hence, various pH (3.0-7.0) were tested using 1 mM PNPG as substrate at room temperature. Maximum acitivity observed at pH 4.8. The realative activities are expressed as percent of activity at pH 4.8. (Fig 4.5)

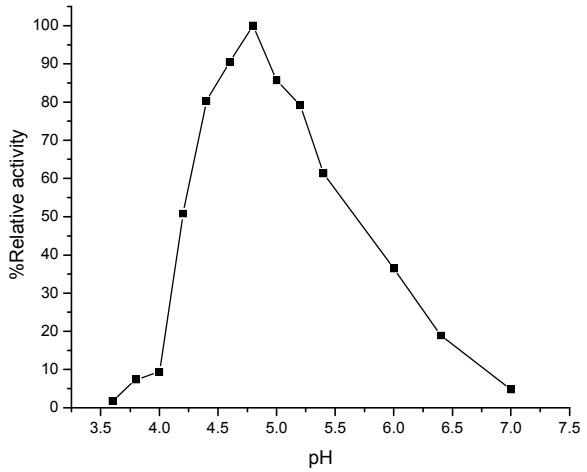


Fig 4.5. The relative activities are expressed as percent of activity at pH 4.8

4.3.3.2. Optimization of Temp

Purified *Leucaena* β-glucosidase (~1μg) was assayed at various temperatures (30^{o}C - 55^{o}C) as described in methods using 1 mM PNPG as substrate.

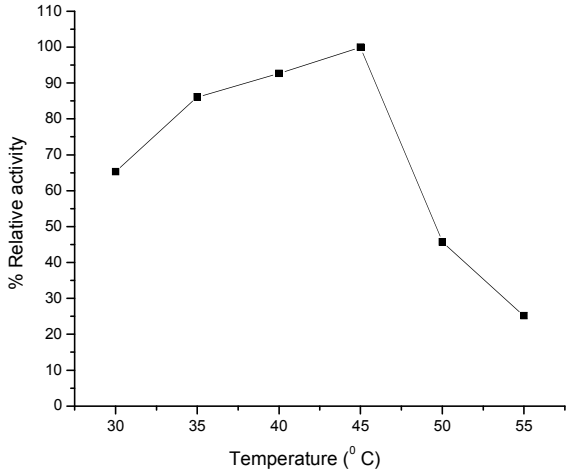


Fig. 4.6. The relative activities are expressed as a percent activity at 45^{o}C

Maximum activity was observed at 45^{o}C. The relative activities are expressed as a percent activity at 45^{o}C (Fig 4.6).

4.3.3. 3.Optimization of Time course

Purified *Leucaena* β-glucosidase was assayed in 100-200 mM Citrate-Phosphate buffer pH 4.8 at 45^{o}C by incubating with 1 mM PNPG as substrate for different time intervals. O.D at 420 was plotted against time. Linearity in time drive observed upto 40 min (Shown in Fig.4.7).

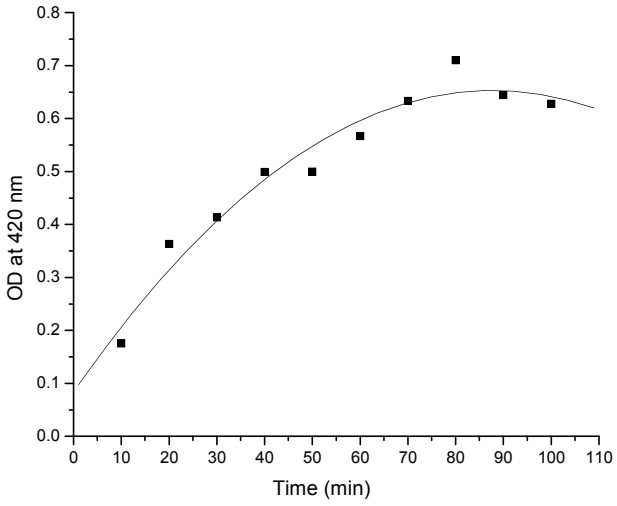


Fig.4.7.Time drive of *Leucaena* β-glucosidase

4.3.4. Determination of Glycone specificity of the recombinant enzyme

The purified recombinant Llbglu1 hydrolase was incubated at its optimum pH (pH 4.8) with potential substrates provided at 2 mM final concentration. Enzyme activity was determined by measuring the rate of PNP (or ONP) production spectrophotometrically at 420 nm with subsequent use of standard curves. Reaction rates are expressed as a percentage of that observed with PNPG (Table 4.2). The enzyme hydrolysed PNP-glucoside and PNP-galactoside.

Table 4.2. Glycone and linkage specificities of recombinant Llbglu1

S.No	Name of PNP sugar	Llbglu1 relative activity (%)
1	p-Nitrophenyl β-D-glucopyranoside	100
2	p- Nitrophenyl β-D-glucoronide	0.8
3	p-Nitrophenyl-N-acetyl-1-thio-β-glucosaminide	0.5
4	p-Nitrophenyl α-D-glucopyranoside	0.9
5	p-Nitrophenyl N-acetyl-β-D-glucosaminide	4.0
6	p-Nitrophenyl β-D-galactopyranoside	53.0
7	p-Nitrophenyl β-D-mannopyranoside	2.2
8	p-Nitrophenyl β-D-xylopyranoside	0.6
9	p-Nitrophenyl β-L-arabinopyranoside	6.0
10	o-Nitrophenyl β-D-glucopyranoside	42.0
11	p-Nitrophenyl N-acetyl-α-D-glucosaminide	0.4

4.3.5. Determination of *Km* and *Vmax* for glycosides

Different substrates, PNPG sugars (Table 4.1) were tested for activity using purified recombinant Llbglu1. The enzyme hydrolysed PNP-glucoside and PNP-galactoside. To find the glucone specificity, kinetic constants like *Km* and *Vmax* were determined for these substrates.

4.3.5.1. Determination of *Km* and *Vmax* for PNPG

The kinetic rate constants, *Km* and *Vmax* of the PNP-glucoside were determined. Enzyme assays were carried out with varying concentration (10 µM-120 µM) of PNPG with fixed dilution of enzyme *i.e* 10 µL (0.48mg/mL)) at optimum assay conditions *i.e* at 45^{o}C in 100 mM-200 mM Citrate-phosphate buffer, pH 4.8. Released PNP was read at 420 nm. One unit was defined as µmol para-nitrophenol liberated per min. The velocity (U/ml/min) was plotted against substrate concentration. *Km* and *Vmax* were determined by Michaelis -Menten curve by Graph pad 5 program Fig.4.8. *Km* was found to be 38.59 µM and *Vmax* 0.8237µM/mg/min

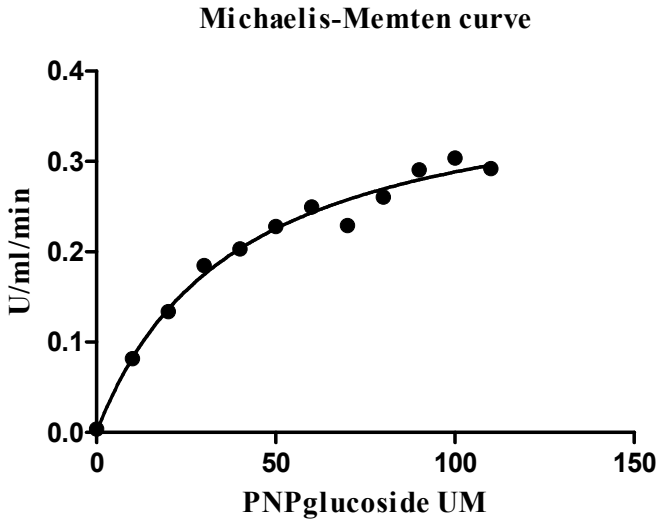


Fig.4.8. Michaelis -Menten curve for the recombinant Llbglu enzyme with PNPG.

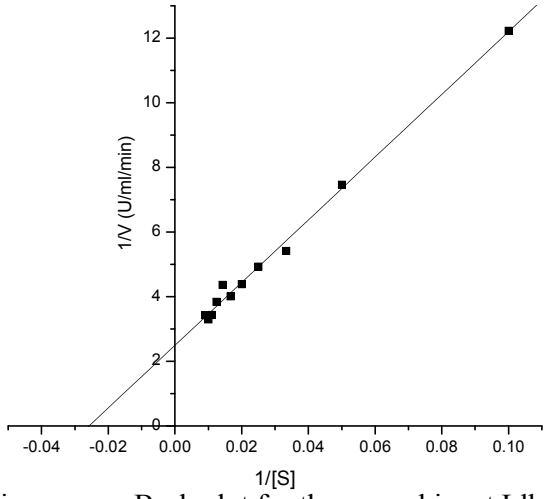


Fig 4.9.Lineweaver-Burk plot for the recombinant Llbglu enzyme with PNPG

Lineweaver-Burk plot (Fig.4.9) was also drawn using program Origin and *Km* was found to be 40 μM and *Vmax* 0.8278 μM/mg/min

4.3.5.2. Determination of *Km* and *Vmax* for PNPGalactoside

To determine the energy of activation of the enzyme , the kinetic rate constants *Km* and *Vmax* of the PNP-galactoside were determined. The enzyme was incubated with varying concentrations of PNP-Galactoside (0.1mM-1.5 mM) at 45^{o}C in 100 mM-200 mM Citrate-phosphate buffer ,pH 4.8, with fixed dilution *i.e* 20 μL (0.12 mg/mL) of the enzyme . The velocity (U/ml/min) was plotted against substrate concentration. *Km* and *Vmax* were determined by Michaelis -Menten curve by Graph pad 5 program Fig.4.10. *Km* was found to be 1.85mM and *Vmax* was 0.1037 μM/mg/min

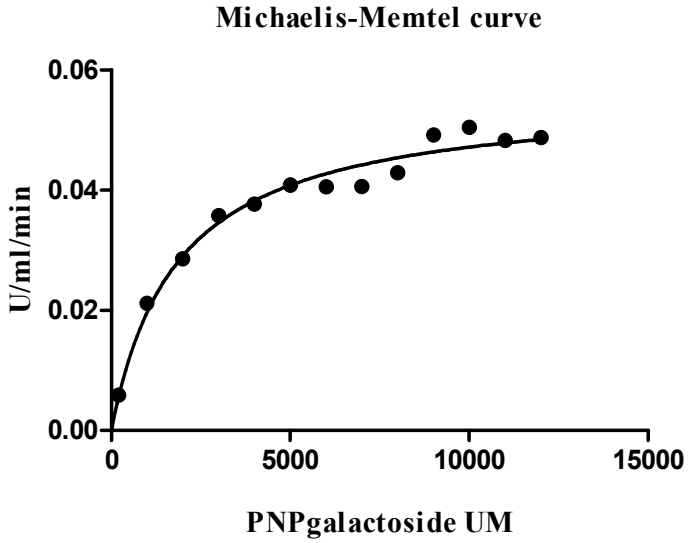


Fig.4.10. Michaelis -Menten curve for the recombinant Llbglu enzyme with PNP-Galactoside.

Lineweaver-Burk plot was also drawn using program Origin and *Km* was found to be 1.75 mM and *Vmax* 0.1008 μM/mg/min (Fig.4.11).

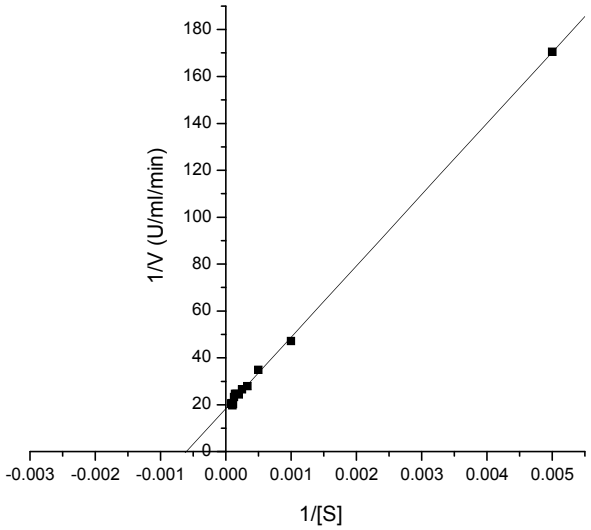
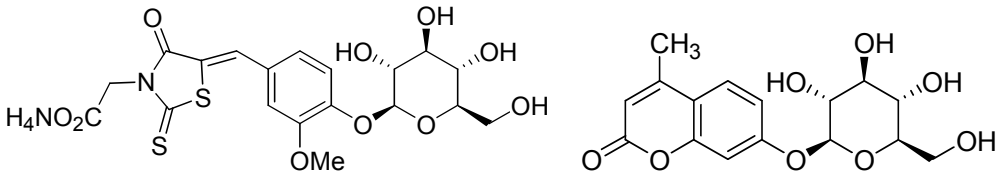


Fig.4.11. Lineweaver-Burk plot for the recombinant Llbglu enzyme with PNP-Galactoside

4.3.6. Determination of aglycone preference of the recombinant enzyme by synthetic glucosides

Two synthetic glucosides one chromomeric substrate, VRA-G (Ammonium 5-[4-(β-D-glucopyronosyl0xy)-3-methoxymethylene]-2-thioxothiazolidin-4-one-3-acetate and another flourogenic substrate, 4-Methyl-umbellifryl-β-D-glucopyranoside were used (structures are shown in Fig.4.12.) for hydrolysis under standard assay conditions.



VRA-G 4MUG

Fig.4.12. Structures of Glucosides

4.3.6.1. Determination of *Km* and *Vmax* for VRA-G

The kinetic rate constants, *Km* and *Vmax* of the VRA-G were determined. Enzyme assays were carried out with varying concentration (0.5 mM-2.0 mM) of VRA-G with fixed dilution of enzyme *i.e* 3.0 µL (0.52mg/mL) at optimum assay conditions *i.e* at 45^{o}C in 100 mM-200 mM Citrate-phosphate buffer, pH 4.8. Released redcoloured

phenol was read at 490 nm. One unit was defined as amount of enzyme required to release 1 µmol VRA per min under standard assay conditions.. The velocity (U/ml/min) was plotted against substrate concentration. *Km* and *Vmax* were determined by Lineweaver-Burk plot (Fig.4.13) drawn using program Origin and *Km* was found to be 3.20 mM and *Vmax* 0.832mM/mg/min under standard assay conditions

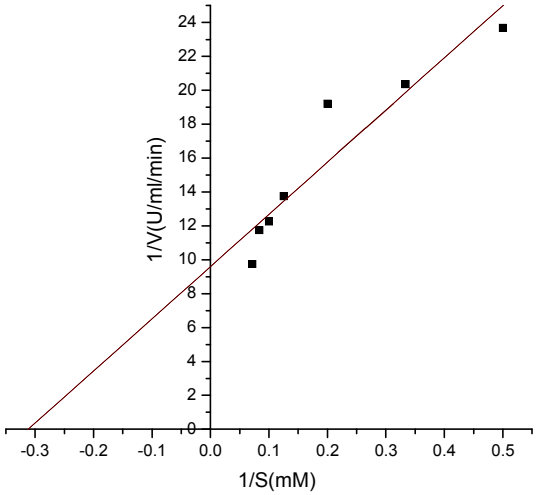


Fig.4.13. Lineweaver-Burk plot for the recombinant Llbglu enzyme with VRA-G

4.3.6.2. Determination of *Km* and *Vmax* for 4MUG

The kinetic rate constants, *Km* and *Vmax* of the 4MUG were determined. Enzyme assays were carried out with varying concentration (10 µM-200 µM) of 4MUG with fixed dilution of enzyme *i.e* 20 µL (0.02 mg/mL) at optimum assay conditions *i.e* at 45^{o}C in 100 mM-200 mM Citrate-phosphate buffer, pH 4.8. The fluorescence was measured at 448 nm. One unit of enzyme activity is the amount of enzyme required to release 1 µmol 4-Methylumbelliferone per minute under standard assay conditions. The velocity (U/ml/min) was plotted against substrate concentration. *Km* and *Vmax* were determined by Lineweaver-Burk plot (Fig.4.14) drawn using program Origin. *Km* was found to be 89.28 µM and *Vmax* 4.17mM/mg/min under standard assay conditions

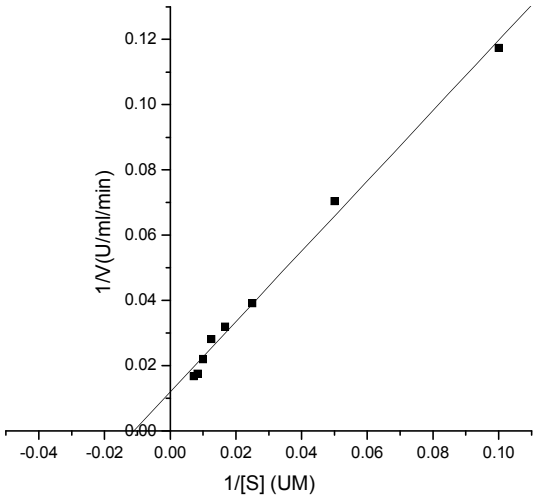


Fig.4.14. Lineweaver-Burk plot for the recombinant Llbglu enzyme with 4MUG

4.3.7. Bioinformatic analysis of *L. leucocephala* β-glucosidase

4.3.7.1. Homology Modeling of *L. leucocephala* β-glucosidase

To find the homologous sequences in Protein Data Bank, the primary sequence of *Leucaena leucocephala* β-glucosidase was searched against PDB using BLASTP program at NCBI server (http://www.ncbi.nlm.nih.gov/blast). Among all the homologs, cyanogenic β-glucosidase from white clover (*T. repens*, PDB ID.1CBG) was the closest to Llbglu1, with 70% identity at the amino acid sequence level. Strictosidine glucosidase (2JF7:A) from *Rauvolfia serpentina* , Dhurrinase (1V02:E) from *Sorghum bicolor*, Myrosinase (1MYR:A) from *Sinapis alba* showed 53%, 49% and 47% identity respectively. The three-dimensional coordinates of white clover β-glucosidase structure (1CBG) were used as template to generate the 3D model of the Llbglu1 using the program Modeller (Marti-Renom *et al*., 2000). Out of 507 residues submitted for homology model, 482 residues were modeled in the structure. 24 N-terminal residues remained unmodeled because they are not having regular secondary structures and coming in loop region. Modeled structure was visualized with program PyMoL (DeLano, 2002). In *Leucaena* β-glucosidase a $(\beta/\alpha)_8$ fold was observed as shown in

figure 4.15, which is a common feature among the Family 1 β-glucosidases belonging to clan GH-A.

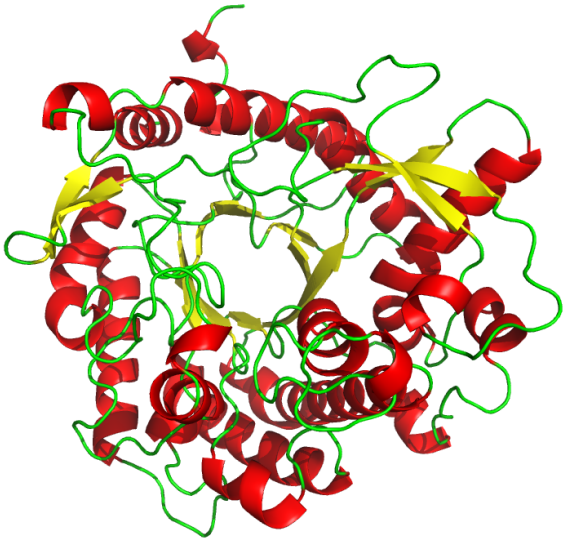


Fig. 4.15. A cartoon representation of the *Leucaena* β-glucosidase (Llbglu1) modeled structure. α-helices, β-sheets and loops are shown in red, yellow and green colour respectively. Top view looking down the TIM barrel axis.

4.3.7.2. Checking the quality of the model with PROCHECK

The quality of the model was examined using the program PROCHECK (Laskowski *et al.*, 1993). The Ramachandran (φ, ψ) (Ramachandran and Sasisekhara, 1968) plot for Llbglu1shown in figure 4.16. 90.3% of residues are in allowed region of The Ramachandran (φ, ψ) plot and 9.3% residues are in additionally allowed regions, 0.5% in the generously allowed region. No residue was present in disallowed region.

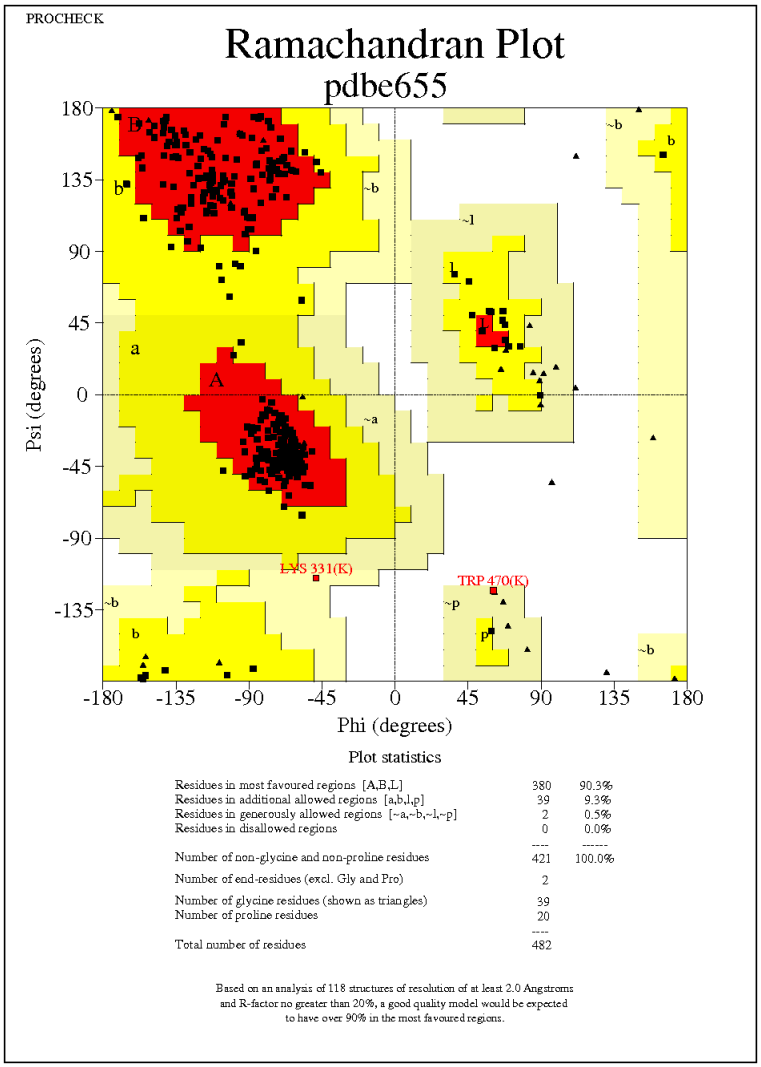


Fig. 4.16. Ramachandran(ϕ,ψ) plot for modeled *Leucaena β-glucosidase* (Llbglu1)

4.3.7.3. Secondary structure calculation and comparison between *L. leucocephala* β-glucosidase and *Trifolium* CBG

In order to compare secondary structures of Llbglu1 and 1CBG, pair wise sequence alignment using ClustalW has been done. Modeled structure of Llbglu1 and the pair wise alignment have been used as input for web based program ESPRIPT (Gouet *et al.*, 1999) to generate comparative picture of secondary structural elements (α helices, β-sheets and turns) of Llbglu1 with the already known 1CBG. The comparative secondary structural elements of *Leucaena* β-glucosidase with that of 1CBG are shown along the amino acid sequences in figure 4.17. As shown in figure it is quite evident that the contents of secondary structural elements in Llbglu1 are more or less similar to 1CBG.

Due to presence of TIM fold, locations of secondary structures in Llbglu1 are conserved and matching exactly with 1CBG.

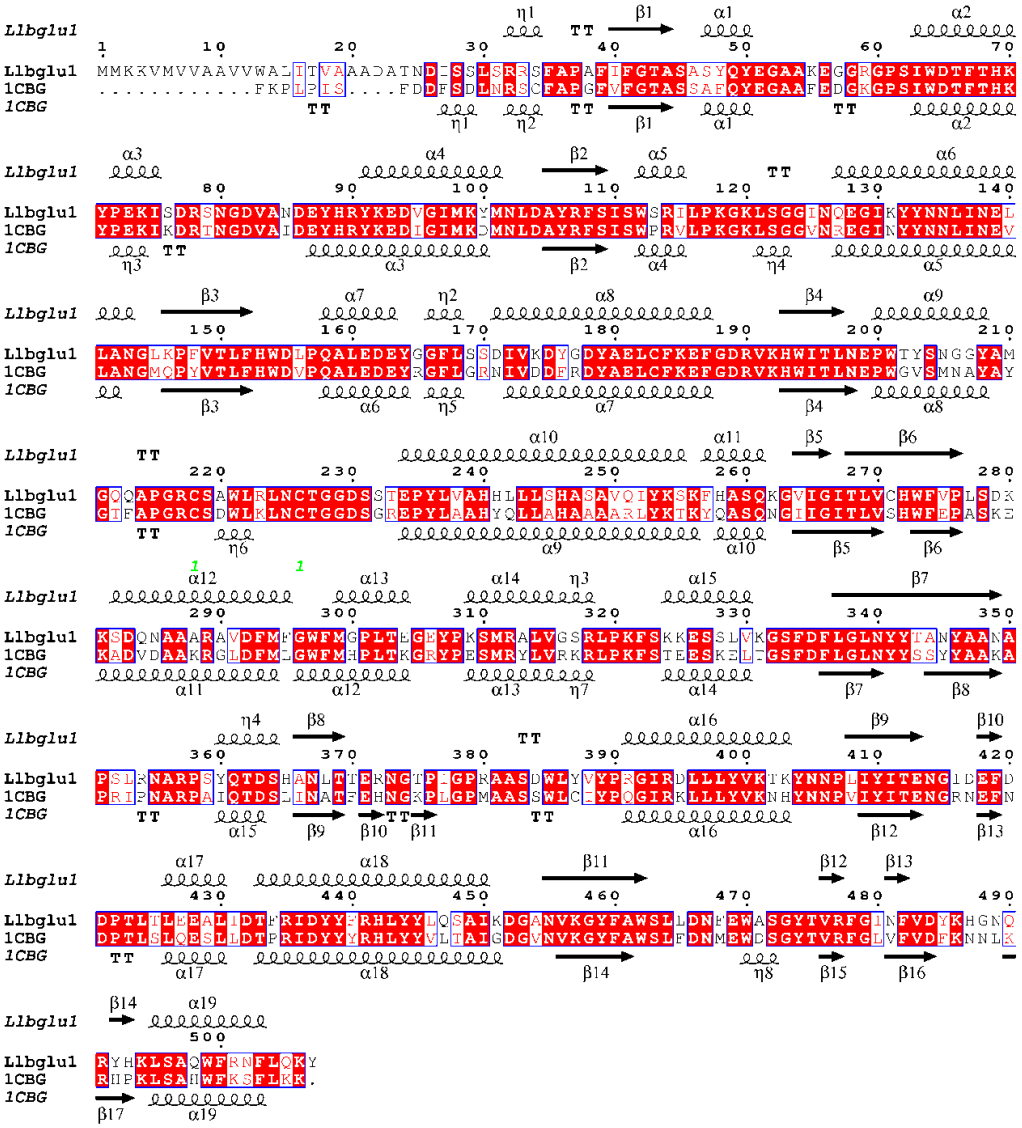


Fig. 4.17. Pair wise sequence alignment of Llbglu1 and 1CBG. The secondary structures of Llbglu and CBG are shown on the top and bottom of sequence alignment respectively. Coils and arrows represent helices and strands respectively. α, β, η and TT correspond to α-helix, β-strand, 3_{10}-helix and β-turn respectively. The Figure was generated using web based program ESPRIPT (Gouet *et al*., 1999).

4.3.7.4. Structural superimposition of *L. leucocephala* β-glucosidase with *Trifolium* CBG

Pair wise structural alignment of modeled *Leucaena* β-glucosidase was done with *Trifolium* CBG using combinatorial extension algorithm at SDSC-CE (http://cl.sdsc.edu/ce.html) (Shindyalov., 1998). Cartoon view of the superposed structures is shown in the figure 4.18 and active site residues are compared in the figure 4.19. There have been a sufficient number of crystal structures of GH1 enzymes gathered to clearly establish the links between active-site amino acid residues and ligand components. *Z. mays* ZmGlu1 β-glucosidase has a slot-like active site, with the catalytic proton donor/base and nucleophile being Glu191 and Glu406, respectively (Czjzek *et al*., 2000). In *Trifolium* cyanogenic β-glucosidase those catalytic residues Glu 183 and Glu 397 corresponds Glu 199 and Glu 413 in Llbglu1 respectively (shown in fig 4.19). Other important residues in the active site of Llbglu1 are His 153 (137), Asn 198 (182), Val 270 (254), Asn 340 (324), Trp 462 (446) (corresponding residues in 1CBG are written in brackets). Most of the active site residues lie on the loops of the TIM barrel fold.

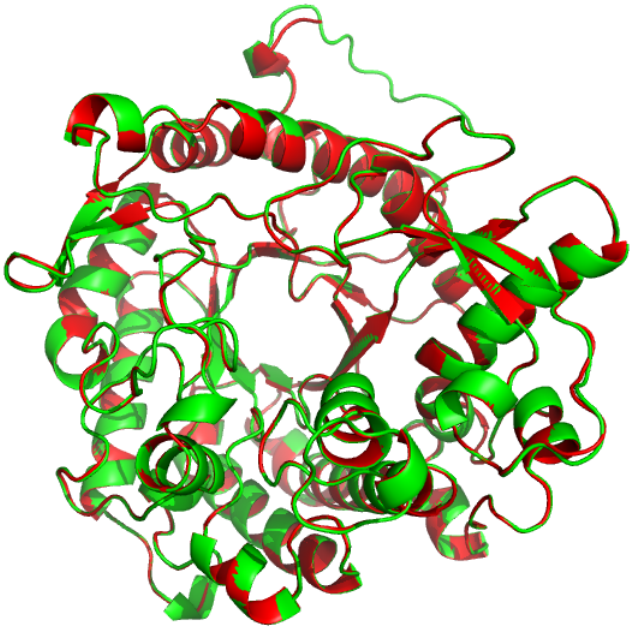


Fig. 4.18. Cartoon view of superimposition of modeled Llbglu1 (red) with 1CBG (green).

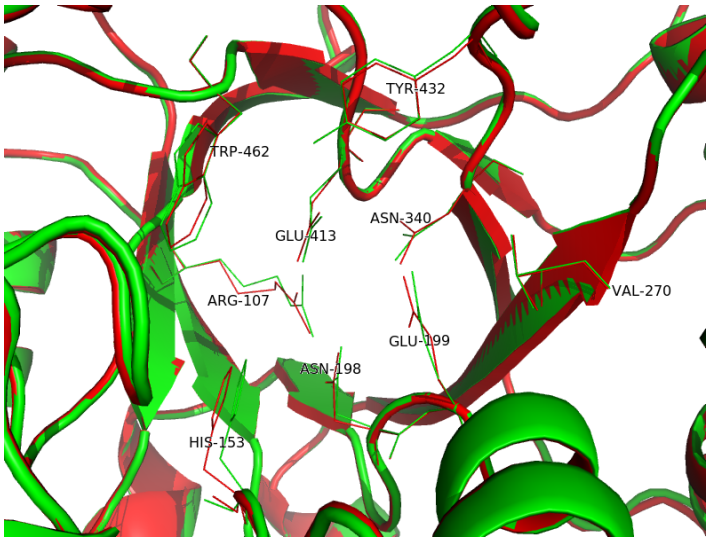


Fig. 4.19. Top view of the barrel showing the superimposition of active site residues of Llbglu1 with *Trifolium* CBG. 1CBG and modeled Llbglu1 are shown in Green and Red colour respectively whereas active site residues are represented in line form. Residue numbers indicated in the figure are of modeled Llbglu1.

4.3.7.5. Analysis of amino acids involved in Glycone binding and catalysis in *Leucaena* β- Glucosidase with other β- glucosidases.

CE-MC server (http://pathway.rit.albany.albany.edu/~cemc/) CE-MC multiple protein structure alignment server provides a web based facility for the alignment of multiple protein structures (known pdbs/user derived pdbs) based on Cα-coordinate distances using combinatorial extension (CE) and Monte Carlo optimization methods (Guda *et al*., 2001, 2004). 10 structures of Family 1 β glucosidases (most of them are of plant origin) were aligned with that of modeled Llbglu1 to compare the important residues involved in glycone binding and catalysis. The catalytic glutamic acid residues in conserved regions (ITLNEP/ITENG) and other conserved residues involved in glycone binding pocket are shown in fig 4.20.

```
                                                                *
  0 1MYR:A      0 -------LNSSSFEADFIFGVASSAYQIEGTI---GRGLNIWDGFTHRYPDKS-GPDHGNGDTTCDSFSY
  1 2E40:A      0 -----------KLPKSFVWGYATAAYQIEGSPDKDGREPSIWDTFCK-APGKI--ADGSSGDVATDSYNR
  2 1V03:A      0 RLSPWEIPRRDWFPPSFLFGAATSAYQIEGAWNEDGKGPSTWDHFCHNFPEWI--VDRSNGDVAADSYHM
  3 1E56:A      0 -LSPSEIPQRDWFPSDFTFGAATSAYQIEGAWNEDGKGESNWDHFCHNHPERI--LDGSNSDIGANSYHM
  4 1CBG:A      0 -------LNRSCFAPGFVFGTASSAFQYEGAAFEDGKGPSIWDTFTHKYPEKI--KDRTNGDVAIDEYHR
  5 1V02:E      0 RLSPWEIPRRDWFPPSFLFGAATSAYQIEGAWNEDGKGPSTWDHFCHNFPEWI--VDRSNGDVAADSYHM
  6 2JFE:X      0 ------------FPAGFGWAAATAAYQVEGGWDADGKGPCVWDTFTHQGGERV--FKNQTGDVACGSYTL
```


7 1WCG:A 0 ----------KFPKDFMFGTSTASYQIEGGWNEDGKGENIWDRLVHTSPEVI--KDGTNGDIACDSYHK
8 1E4M:M 0 -------LNSSSFSSDFIFGVASSAYQIEGTI---GRGLNIWDGFTHRYPNKSG-PDHGNGDTTCDSFSY
9 1V02:A 0 RLSPWEIPRRDWFPPSFLFGAATSAYQIEGAWNEDGKGPSTWDHFCHNFPEWI--VDRSNGDVAADSYHM
10 8LBG:_ 0 -------LSRRSFAPAFIFGTASASYQYEGAAKEGGRGPSIWDTFTHKYPEKI—SDRSNGDVANDEYHR
0 1MYR:A 70 WQKDIDVLDELNATGYRFSIAWSRIIPRGKRSRGVNQKGIDYYHGLIDGLIKKGITPFVTLFHWDLPQTL
1 2E40:A 70 WREDVQLLKSYGVKAYRFSLSWSRIIPKGGRSDPVNGAGIKHYRTLIEELVKEGITPFVTLYHWDLPQAL
2 1V03:A 70 YAEDVRLLKEMGMDAYRFSISWPRILPKGTLAGGINEKGVEYYNKLIDLLLENGIEPYITIFHWDTPQAL
3 1E56:A 70 YKTDVRLLKEMGMDAYRFSISWPRILPKGTKEGGINPDGIKYYRNLINLLLENGIEPYVTIFHWDVPQAL
4 1CBG:A 70 YKEDIGIMKDMNLDAYRFSISWPRVLPKGKLSGGVNREGINYYNNLINEVLANGMQPYVTLFHWDVPQAL
5 1V02:E 70 YAEDVRLLKEMGMDAYRFSISWPRILPKGTLAGGINEKRVEYYNKLIDLLLENGIEPYITIFHWDTPQAL
6 2JFE:X 70 WEEDLKCIKQLGLTHYRFSLSWSRLLPDGTTG-FINQKGIDYYNKIIDDLLKNGVTPIVTLYHFDLPQTL
7 1WCG:A 70 YKEDVAIIKDLNLKFYRFSISWARIAPSGVMN-SLEPKGIAYYNNLINELIKNDIIPLVTMYHWDLPQYL
8 1E4M:M 70 WQKDIDVLDELNATGYRFSIAWSRIIPRGKRSRGVNEKGIDYYHGLISGLIKKGITPFVTLFHWDLPQTL
9 1V02:A 70 YAEDVRLLKEMGMDAYRFSISWPRILPKGTLAGGINEKRVEYYNKLIDLLLENGIEPYITIFHWDTPQAL
10 8LBG:_ 70 YKEDVGIMKYMNLDAYRFSISWSRILPKGKLSGGINQEGIKYYNNLINELLANGLKPFVTLFHWDLPQAL
0 1MYR:A 140 QDEYEGFLDP-----QIIDDFKDYADLCFEEFGDSVKYWLTINQLYSVPTRGYGSALDAPGRCSPTVDPS
1 2E40:A 140 DDRYGGWLNK----EEAIQDFTNYAKLCFESFGDLVQNWITFNEPWVISVMGYGNGIFAPGHV-------
2 1V03:A 140 VEAYGGFLDE-----RIIKDYTDFAKVCFEKFGKTVKNWLTFNDPETFCSVSYGTGVLAPGRCSPGVSCA
3 1E56:A 140 EEKYGGFLD-KSHK-SIVEDYTYFAKVCFDNFGDKVKNWLTFNDPQTFTSFSYGTGVFAPGRCSPGLDCA
4 1CBG:A 140 EDEYRGFLGR-----NIVDDFRDYAELCFKEFGDRVKHWITLNEPWGVSMNAYAYGTFAPGRCSDWLKLN
5 1V02:E 140 VDAYGGFLDE-----RIIKDYTDFAKVCFEKFGKKVKNWLTFNEPETFCSVSYGTGVLAPGRCSPGVSCA
6 2JFE:X 140 ED-QGGWLSE-----AIIESFDKYAQFCFSTFGDRVKQWITINEANVLSVMSYDLGMFPPGIP-------
7 1WCG:A 140 QD-LGGWVNP-----IMSDYFKEYARVLFTYFGDRVKWWITFNEPIAVC-KGYSIKAYAPNLN-------
8 1E4M:M 140 QDEYEGFLDP-----QIIDDFKDYADLCFEEFGDSVKYWLTINQLYSVPTRGYGSALDAPGRCSPTVDPS
9 1V02:A 140 VDAYGGFLDE-----RIIKDYTDFAKVCFEKFGKTVKNWLTFNEPETFCSVSYGTGVLAPGRCSPGVSCA
10 8LBG:_ 140 EDEYGGFLSS-----DIVKDYGDYAELCFKEFGDRVKHWITLNEPWTYSNGGYAMGQQAPGRCSAWLRLN
0 1MYR:A 210 CYAGNSSTEPYIVAHHQLLAHAKVVDLYRKNYTH-----QGGKIGPTMITRWFLPYNDTDRHSIAATERM
1 2E40:A 210 -----SNTEPWIVSHHIILAHAHAVKLYRDEFKE----KQGGQIGITLDSHWLIPYDD-TDASKEATLRA
2 1V03:A 210 VPTGNSLSEPYIVAHNLLRAHAETVDIYNKYHKG-----ADGRIGLALNVFGRVPYTN-TFLDQQAQERS
3 1E56:A 210 YPTGNSLVEPYTAGHNILLAHAEAVDLYNKHYKR----D-DTRIGLAFDVMGRVPYGT-SFLDKQAEERS
4 1CBG:A 210 CTGGDSGREPYLAAHYQLLAHAAAARLYKTKYQA---S-QNGIIGITLVSHWFEPASK-EKADVDAAKRG
5 1V02:E 210 VPTGNSLSEPYIVAHNLLRAHAETVDIYNKYHKG-----ADGRIGLALNVFGRVPYTN-TFLDQQAQERS
6 2JFE:X 210 ----HFGTGGYQAAHNLIKAHARSWHSYDSLFRK--K--QKGMVSLSLFAVWLEPADPNSVSDQEAAKRA
7 1WCG:A 210 ----LKTTGHYLAGHTQLIAHGKAYRLYEEMFKP-T---QNGKISISISGVFFMPKNAESDDDIETAERA
8 1E4M:M 210 CYAGNSSTEPYIVAHHQLLAHAKVVDLYRKNYTH-----QGGKIGPTMITRWFLPYNDTDRHSIAATERM
9 1V02:A 210 VPTGNSLSEPYIVAHNLLRAHAETVDIYNKYHKG-----ADGRIGLALNVFGRVPYTN-TFLDQQAQERS
10 8LBG:_ 210 CTGGDSSTEPYLVAHHLLLSHASAVQIYKSKFHAS----QKGVIGITLVCHWFVPLSD-KKSDQNAAARA
0 1MYR:A 280 KQFFLGWFMGPLT-NGTYPQIMIDTVG--------------------ARLPTFSPEETNLVKGSYDFLGL
1 2E40:A 280 MEFKLGRFANPIY-KGEYPPRIKKILG--------------------DRLPEFTPEEIELVKGSSDFFGL
2 1V03:A 280 MDKCLGWFLEPVV-RGDYPFSMRVSAR--------------------DRVPYFKEKEQEKLVGSYDMIGI
3 1E56:A 280 WDINLGWFLEPVV-RGDYPFSMRSLAR--------------------ERLPFFKDEQKEKLAGSYNMLGL
4 1CBG:A 280 LDFMLGWFMHPLT-KGRYPESMRYLVR--------------------KRLPKFSTEESKELTGSFDFLGL
5 1V02:E 280 MDKCLGWFLEPVV-RGDYPFSMRVSAR--------------------DRVPYFKEKEQEKLVGSYDMIGI
6 2JFE:X 280 ITFHLDLFAKPIFIDGDYPEVVKSQIA----------SMSQKQGYPSSRLPEFTEEEKKMIKGTADFFAV
7 1WCG:A 280 NQFERGWFGHPVY-KGDYPPIMKKWVDQKSKEEGLPW----------SKLPKFTKDEIKLLKGTADFYAL
8 1E4M:M 280 KEFFLGWFMGPLT-NGTYPQIMIDTVG--------------------ERLPSFSPEESNLVKGSYDFLGL
9 1V02:A 280 MDKCLGWFLEPVV-RGDYPFSMRVSAR--------------------DRVPYFKEKEQEKLVGSYDMIGI
10 8LBG:_ 280 VDFMFGWFMGPLT-EGEYPKSMRALVG--------------------SRLPKFSKKESSLVKGSFDFLGL
0 1MYR:A 350 NYYFTQYAQPSP----NPV-------NATNHT-------AMMDAGAKLTYINASGHYIGP------LFES
1 2E40:A 350 NTYTTHLVQDGG---------S--------DE-------LA--GFVKTGHTRADGTQLGT----------
2 1V03:A 350 NYYTSTFSKHIDLSPNNSPV-----------L-------NTDDAYASQETKGPDGNAIGP----PT----
3 1E56:A 350 NYYTSRFSKNIDISPNYSPV-----------L-------NTDDAYASQEVNGPDGKPIGP----------
4 1CBG:A 350 NYYSSYYAAKAPRIPNARP-----------A--------IQTDSLINATFEHN-GKPLGP----------
5 1V02:E 350 NYYTSTFSKHIDLSPNNSPV-----------L-------NTDDAYASQETKGPDGNAIGP----------
6 2JFE:X 350 QYYTTRLIKYQENKK--GEL-----------G-------ILQDAEIEFFPD-----PSWI----------
7 1WCG:A 350 NHYSSRLVTFG---------------------SDPNPNF-NPDASYVTSVD-----EAWLKPNE------
8 1E4M:M 350 NYYFTQYAQPSP----NPV---NSTN----HT-------AMMDAGAKLTYINASGHYIGP----------

```
 9 1V02:A    350 NYYTSTFSKHIDLSPNNSPV-----------L-------NTDDAYASQETKGPDGNAIGP----------
10 8LBG:_    350 NYYTANYAANAPSLRN-A--R----------PS-------YQTDSHANLTTERN-GTPIGP----------
```

```
                                                                                  ↓
 0 1MYR:A    420 D------------------------GGDGSS---NIYYYPKGIYSVMDYFKNKYYNPLIYVTENGISTP
 1 2E40:A    420 -QS-----------------------------DMGWLQTYGPGFRWLLNYLWKAYD-KPVYVTENGFPV-
 2 1V03:A    420 --------------------------------GNAWINMYPKGLHDILMTMKNKYGNPPMYITENGMGDI
 3 1E56:A    420 ---PM----------------------------GNPWIYMYPEGLKDLLMIMKNKYGNPPIYITENGIGDV
 4 1CBG:A    420 -----MA---------------------------ASSWLCIYPQGIRKLLLYVKNHYNNPVIYITENGRNEF
 5 1V02:E    420 -------PT-----------------------GNAWINMYPKGLHDILMTMKNKYGNPPMYITENGMGDI
 6 2JFE:X    420 ---------N----------------------VDWIYVVPWGVCKLLKYIKDTYNNPVIYITENGFPQS
 7 1WCG:A    420 ----------T----------------------P-YIIPVPEGLRKLLIWLKNEYGNPQLLITENGYGDD
 8 1E4M:M    420 -----------LFEKD----KADSTD---------NIYYYPKGIYSVMDYFKNKYYNPLIYVTENGISTP
 9 1V02:A    420 ----------------PT--------------GNAWINMYPKGLHDILMTMKNKYGNPPMYITENGMGDI
10 8LBG:_    420 ------------------RA------------ASDWLYVYPRGIRDLLLYVKTKYNNPLIYITENGIDEF
```

```
                                                                *      **       *
 0 1MYR:A    490 GS----------ENRKESMLDYTRIDYLCSHLCFLNKVIKEKDVNVKGYLAWALGDNYEFNNGFTVRFGL
 1 2E40:A    490 KG-----E-N-DLPVEQAVDDTDRQAYYRDYTEALLQAVTEDGADVRGYFGWSLLDNFEWAEGYKVRFGV
 2 1V03:A    490 DK------G--DLPKPVALEDHTRLDYIQRHLSVLKQSIDL-GADVRGYFAWSLLDNFEWSSGYTERFGI
 3 1E56:A    490 DT----K--ETPLPMEAALNDYKRLDYIQRHIATLKESIDL-GSNVQGYFAWSLLDNFEWFAGFTERYGI
 4 1CBG:A    490 ND---P-----TLSLQESLLDTPRIDYYYRHLYYVLTAIGD-GVNVKGYFAWSLFDNMEWDSGYTVRFGL
 5 1V02:E    490 DK--G------DLPKPVALEDHTRLDYIQRHLSVLKQSIDL-GADVRGYFAWSLLDNFEWSSGYTERFGI
 6 2JFE:X    490 DP---------------APLDDTQRWEYFRQTFQELFKAIQLDKVNLQVYCAWSLLDNFEWNQGYSSRFGL
 7 1WCG:A    490 G----------------QLDDFEKISYLKNYLNATLQAMYEDKCNVIGYTVWSLLDNFEWFYGYSIHFGL
 8 1E4M:M    490 GD----------ENRNQSMLDYTRIDYLCSHLCFLNKVIKEKDVNVKGYLAWALGDNYEFNKGFTVRFGL
 9 1V02:A    490 DK-G-------DLPKPVALEDHTRLDYIQRHLSVLKQSIDL-GADVRGYFAWSLLDNFEWSSGYTERFGI
10 8LBG:_    490 DDP--------TLTLEEALIDTFRIDYYFRHLYYLQSAIKD-GANVKGYFAWSLLDNFEWASGYTVRFGI
```

```
 0 1MYR:A    560 SYINWNN--VTDRDLKKSGQWYQKFIS
 1 2E40:A    560 THVDYET---QKRTPKKSAEFLSR---
 2 1V03:A    560 VYVDREN--GCERTMKRSARWLQEFNG
 3 1E56:A    560 VYVDRNN--NCTRYMKESAKWLKEFNT
 4 1CBG:A    560 VFVDFKN--NLKRHPKLSAHWFKSFLK
 5 1V02:E    560 VYVDREN--GCERTMKRSARWLQEFNG
 6 2JFE:X    560 FHVDFED-PARPRVPYTSAKEYAKII-
 7 1WCG:A    560 VKIDFNDP-QRTRTKRESYTYFKNVVS
 8 1E4M:M    560 SYIDWNN--VTDRDLKKSGQWYQSFIS
 9 1V02:A    560 VYVDREN--GCERTMKRSARWLQEFNG
10 8LBG:_    560 NFVDYKH--GNQRYHKLSAQWFRNFLQ
```

Fig. 4.20. Multiple structural alignment in sequence form. *Sinapis alba* myrosinase (1MYR:A), *Sinapis alba* myrosinase (1E4M:M), *Sorghum bicolor* dhurrinase1 (1V03:A), *Sorghum bicolor* dhurrinase1 (1V02:E), *Sorghum bicolor* dhurrinase1 (1V02:A), Basidiomycete β-glucosidase bgl1a (2E40:A), *Maize* β-glucosidase zmglu1 (1E56:A), white clover *Trifolium repens* β-glucosidase (1CBG:A), human cytosolic β-glucosidase (2JFE:X), Aphid myrosinase (1WCG:A), *Leucaena leucocephala* β-glucosidase (8LBG:_). Catalytic glutamic acid residues in conserved regions (ITLNEP/ITENG) are marked with arrow. Conserved residues involved in glycone binding pocket are shown in red.

4.3.7.6. Docking of Rhodiocyanosides to the *Leucaena* β-glucosidase

Glycosyl hydrolase Family 1 involved in defense hydrolyses of defensive glycosides on injury. The defensive glycosides Benzoxazinoid glucosides are present in Poaceae and the glucosinolates are present in capparales. In leguminaceae the defensive glycosides are either isoflavonoid glucosides (Naoumkina *et al*., 2007) or hydroxynitrile glucosides (Barrett *et al*., 1995, Morant *et al*., 2008). Phylogenetic analysis of *Leucaena* β-glucosidase shows that it is closely related to Lotus β-glucosidases and shows 80-90 % sequence identity. Lotus β-glucosidases hydrolyse Rhodiocyanosides (Morant *et al*., 2008). However so far, there are no hydroxynitrile glycosides are reported from *Leucaena*. So, docking studies were carried out with Rhodiocyanoside A as ligand, into the modeled *Leucaena leucocephala* β-glucosidase in an attempt to find its probable natural substrate in vivo. The best docked pose of Enzyme-ligand complex using Flexible ligand fit docking program in DS Modeling 1.2-SBD Docking Module by Accelrys™ Software (Accelrys Software Inc., 2003) is shown in figure 4.21.

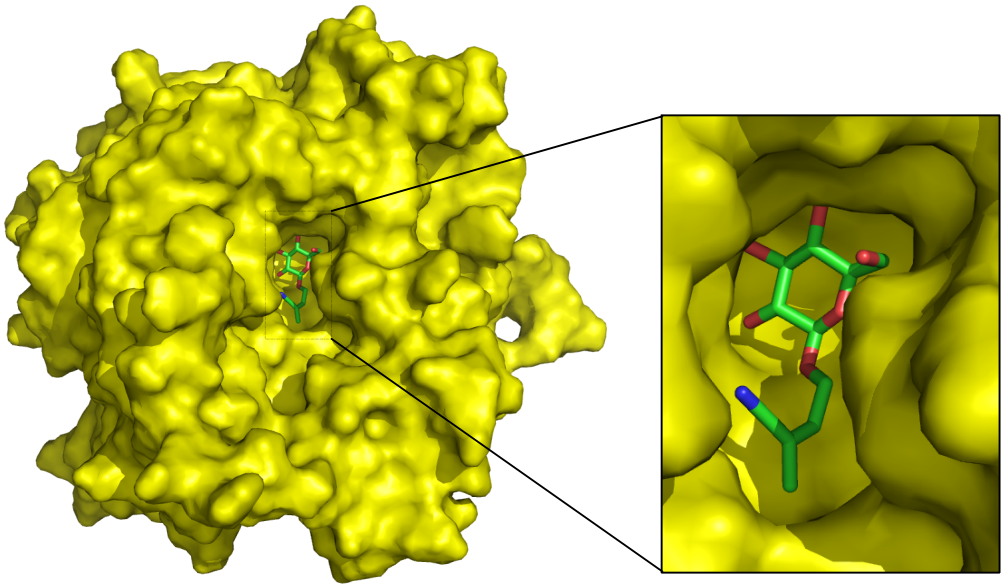
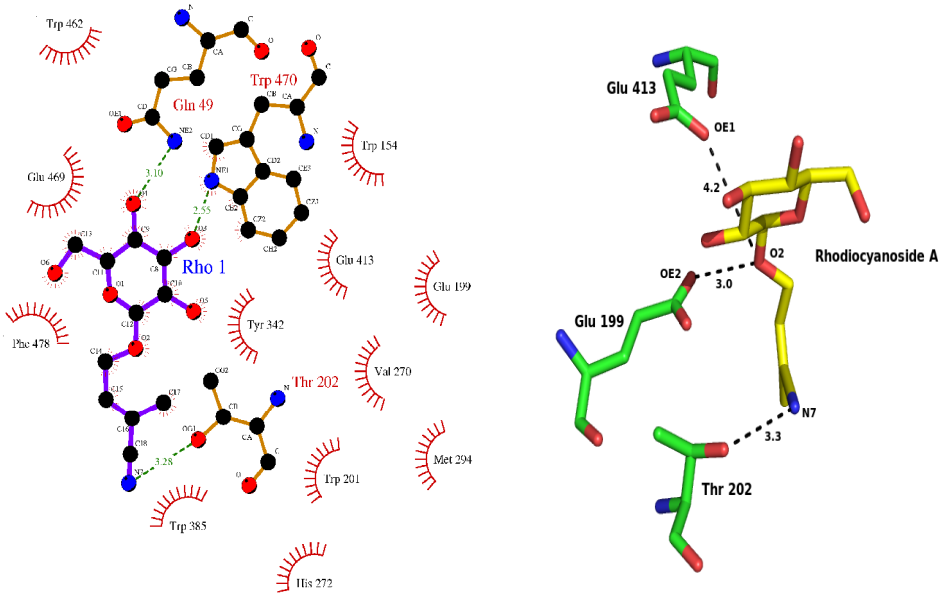


Figure 4.21. The surface structure of the Llbglu1 with the docked Rhodiocyanoside A (Carbon - green, Oxygen - red, Nitrogen - blue).

Theoretically, the free-energy change (ΔG) of the best pose of enzyme-ligand complex was calculated as -5.517 kcal/mol. Many variations of amino acid residues occur with different GH1 members, but their very similar active site structures ensure that their analogous residues will have most of the same interactions. In general, those glycon glycosyl ligands that are free to take up different ring conformations on binding in active sites of GH members are found as relaxed 4C1 conformers (Hill *et al*., 2008). Deeper in the cleft it has glycon-binding region, with Glu199 and Glu413 interacting with the O2 atom of the glycon glucosyl residue. The distances of O2 were found 3.0 Å and 4.2 Å from the Glu199 and Glu413 side chain terminal oxygens respectively (shown in figure 4.22B). These distances clearly show that one water molecule can come and hydrolyse the glycosidic bond by following the well known acid base catalysis mechanism.

Receptor ligand interactions are shown in Ligplot which was generated through PDBSum on ebi server (http://www.ebi.ac.uk/pdbsum). Three hydrogen bonds are clearly visible. O3 and O4 of ligand are acting as hydrogen bond donor whereas NE1(Trp 470) and NE2(Gln 49) act as hydrogen bond acceptor. Aglucon moiety of the ligand (N7, H-bond acceptor) forms hydrogen bond with Thr 202 (OG1, H-bond donor). Because of this hydrogen bonding, Rhodiocyanoside A could be a potential natural substrate for the enzyme, even sufficient ΔG value (-5.517 kcal/mol) of enzyme-ligand complex also supports this fact.

Ligplot (Wallace *et al*., 1995), a schematic diagram of Llbglu1-Rhodiocyanoside A (protein-ligand) interactions is shown in figure 4.22A. The interactions shown are those mediated by hydrogen bonds and by hydrophobic contacts. Hydrogen bonds are indicated by dashed lines between the atoms involved, while hydrophobic contacts are represented by an arc with spokes radiating towards the ligand atoms they contact. The contacted atoms are shown with spokes radiating back.

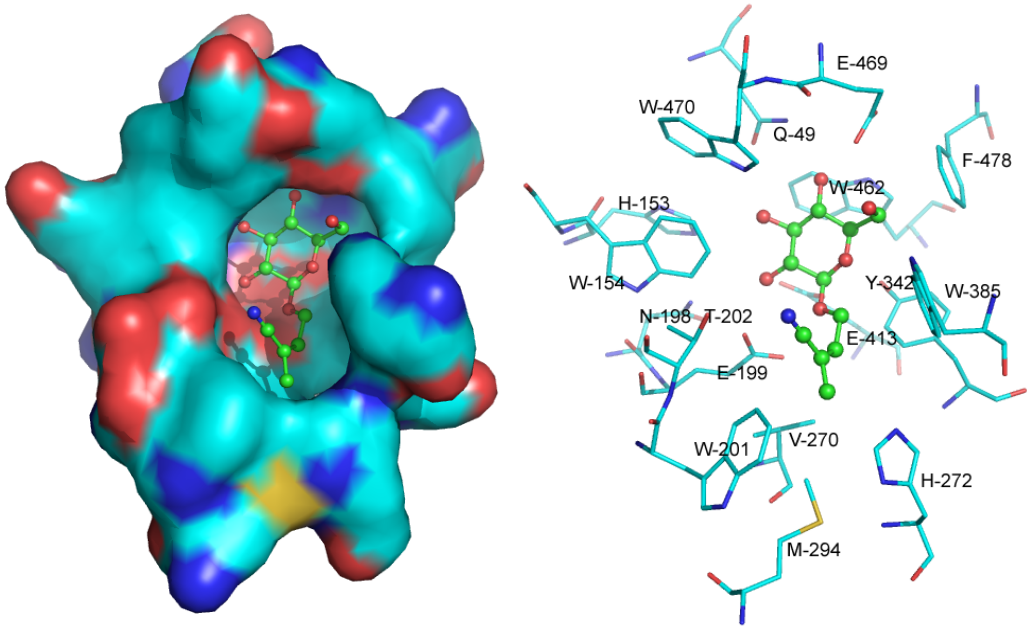


A B

Figure 4.22. Interactions of catalytic residues of Llbglu1 with Rhodiocyanoside A in Enzyme-Substrate Complex obtained through Ligand flexible fit docking. **A)**. Ligplot: schematic diagrams of Llbglu1-Rhodiocyanoside A (protein-ligand) interactions. The interactions shown are those mediated by hydrogen bonds and by hydrophobic contacts. Hydrogen bonds are indicated by dashed lines between the atoms involved, while hydrophobic contacts are represented by an arc with spokes radiating towards the ligand atoms they contact. The contacted atoms are shown with spokes radiating back. Ligplot has been generated through PDBSum on ebi server. **B)**. Three dimensional orientations of acid/base catalytic residues Glu 199 and Glu 413 (both shown in green colour) in binding site of Llbglu1 along with the substrate Rhodiocyanoside A (shown in yellow colour). The distances of glycosidic oxygen of Rhodiocyanoside A with sidechain oxygen of catalytic Glu-199 and Glu-413 are 3.0 Å and 4.2 Å. Aglucon nitrogen in Rhodiocyanoside A (N_7) forms hydrogen bonding with Thr 202(green color).

Total 17 residues were found in 5 Å vicinity of the docked ligand in Llbglu1 pocket (Fig 4.23). These residues are lining the full active site pocket. These are Glu 49, His 153,

Trp 154, Asn 198, Glu 199, Trp 201, Thr 202, Val 270, His 272, Met 294, Tyr 342, Trp 385, Glu 413, Trp 462, Glu 469, Trp 470 and Phe 478. Out of 17 residues, 9 are aromatic ring containing amino acids (W-5, H-2, F-1, Y-1). Definitely these aromatic ring containing amino acids have important role to attract and hold the substrate till the end of the hydrolysis reaction. Tryptophan residue, Trp-385 (of Llbglu1) is conserved within the enzyme family 1, and its role in substrate recognition has been described previously (Czjzek *et al*., 2000). The Trp 385 in all known plant glucosidases shows its side chain torsion angle $\chi \sim 60^{o}$. The residues Trp 462 and Glu 469 in Llbglu1 are also conserved in the active pocket of other GH 1 members (Barleben *et al*., 2008).



(A) (B)

Fig.4.23. A). Molecular surface structure of the residues lining the active site pocket of the Llbglu1 enzyme with Rhodiocyanoside A (ball & stick representation) positioned in the bindingcleft. B). Locations of all the 17 residues (stick representation) forming the binding pocket containing docked Rhodiocyanoside A ligand.

4.4. Conclusions

Naturally occurring glucosidic substrates contain a broad range of aglycone groups, including cyanogenic glucosides (Barette *et al*., 1995; Eksittukul and Chulavatanatol 1988), cellobiose (Ferriera and Terrra, 1983), phenolic glucosides (Podstolski and

Levak, 1970), thioglucosides (Durham and Poulton, 1989), and isoflavonoid glucosides (Svasti *et al*., 1999). The identification of β-glucosidases, their substrates, and the nature of their interactions will not only shed light on the structure and function of the enzymes, but also help define their biological significance in vivo. To investigate the molecular basis of enzymatic catalysis and to engineer it to achieve desirable catalytic properties, a reliable expression and purification system was developed in this study.

A 23 amino acid long signal sequence was predicted in Llbglu 1 by signalP program at C-terminus. The Llbglu1 was cloned without signal sequence, into pET 28 a(+) vector at *EcoR I* and *Not I* sites. The plasmid (pET-Llbglu1-His) was mobilized in to *E.coli* BL 21(DE3). Individual colonies were screened for protein expression. The positive colonies expressing the recombinant protein were grown in LB media for protein expression by 1mM IPTG induction for 4-7 h at 37^{o}C and cells were harvested and lysed, both soluble and in-soluble fractions were analysed on SDS-PAGE. Gradual accumulation of recombinant protein (~55 kD) was observed in insoluble fraction from 1-7 h (Fig.4.3) at the above conditions.

Growth conditions of the culture like temperature, IPTG concentration and period of incubation were optimized so that expressed protein was in soluble form (active form). Activity was monitored in crude lysate by assaying for β-glucosidase activity with PNPG as substrate. Temperature 18^{o}C, 0.05mM IPTG and 9 h of IPTG induction were optimum for the production of active enzyme. The recombinant protein was purified by Immobilized Metal Affinity chromatography, using Ni-NTA agarose beads (Qiagen, USA). The purified Protein concentration was determined by Lowry's methods using BSA as standard. The purified protein was analyzed on SDS-PAGE and it was ~55kD .

Enzyme assay conditions for the recombinant *Leucaena* β-glucosidase were optimized. pH 4.8 (100-200mM Citrate-Phophate buffer), 45^{o}C were found to be optimum. Linearity in time drive observed upto 40 min of incubation.

The purified recombinant Llbglu1 hydrolase was incubated at its optimum pH (pH 4.8) with potential substrates provided at 2 mM final concentration. Enzyme activity was determined by measuring the rate of PNP (or ONP) production spectrophotometrically at 420 nm with subsequent use of standard curves. The enzyme hydrolysed PNP-

glucoside and PNP-galactoside. To find the glucone specificity, kinetic constants like Km and *Vmax* were determined for these substrates *Km* and *Vmax* was found to be 40 μM and 0.8278 μM/mg/min respectively for PNP-glycoside. *Km* and *Vmax* was found to be 1.75mM and 0.1008 μM/mg/min for PNP-galactoside.To find the aglycone specificity two synthetic glucosides VRA-G and 4MUG were used. The *Km* for these substrates was found to be 3.20 mM and 89.28 μM respectively.

Bioinformatics analysis of the *Leucaena* β-glucosidase(Llbglu1) were carried out for further characterization**.** Secondary structure prediction of *L. leucocephala* β-glucosidase was compared with Trifolium cyanogenic β-glucosidase(1CBG). 1CBG was found to be closest to Llbglu1, with 70% identity at the amino acid sequence level. The structure of white clover β-glucosidase was used as template to generate the 3D model of the Llbglu1 using the program Modeler (Marti-Renom *et al*., 2000). The modeled structure of Llbglu1 was visualized using program PyMoL (DeLano ,WL; 2002). In *Leucaena* β-glucosidase a $(\beta/\alpha)_8$ fold was observed, which is a common feature among the Family 1 β-glucosidases. The quality of the model was examined using the program PROCHECK (Laskowski *et al*., 1993). Pairwise structural superimposition of modeled Leucaena β-glucosidase was done with trifolium cbg using Combinatorial extension algorithm at SDSC-CE (http://cl.sdsc.edu/ce.html)

In family 1 GH s Glutanic acid serves as catalytic proton donor/base and nucleophile. In Trifolium cyanogenic β-glucosidase those catalytic residues Glu 183 and Glu 397 corresponds Glu 199 and Glu 413 in Llbglu1 respectvely. Analysis of amino acids involved in Glycone binding and catalysis in *Leucaena* β-Glucosidase with other β-glucosidases was done by structural alignment at CEMC server (http://pathway.rit.albany.albany.edu/~cemc/. A natural hydroxynitrle glucoside, Rhodiocyanoside A was docked using program Accelrys, in to the active pocket of Llbglu1.

5.1. Introduction

Quantitative Real Time PCR (QPCR) is a powerful tool for gene expression analysis and was first demonstrated by Higuchi *et al*., (1992, 1993). QPCR quantitates the initial amount of the template most specifically, sensitively and reproducibly, and is a preferable alternative to other forms of quantitative reverse transcriptase PCR that detect the amount of final amplified product at the end-point (Freeman *et al*., 1999; Raeymaekers, 2000). QPCR monitors the fluorescence emitted during the reaction as an indicator of amplicon production during each PCR cycle (i.e., in real time) as opposed to the endpoint detection (Higuchi *et al*., 1992, 1993). The QPCR system is based on the detection and quantitation of a fluorescent reporter (Lee *et al*., 1993; Livak *et al*., 1995). Quantitative reverse transcription real-time PCR (qRT-PCR) has been rapidly adopted in gene expression studies that require highly sensitive and accurate quantification of mRNA levels from minimal tissue sample. qRT-PCR can detect differences in transcript abundance or template copy number between samples, and is amenable for high-throughput analysis (Gentle *et al*., 2001). Generally qRT-PCR assays use one of two standard procedures; relative or absolute quantification. Either procedure can be applied for gene expression studies that require knowledge of how specific transcript abundance varies within a sample set.

Relative quantification compares expression of the target gene to that of one or more reference genes within the same sample. Reference genes should be consistently expressed across the samples being surveyed (e.g. developmental series, tissue set, experimental regime). Housekeeping genes such as 18S rRNA, β-actin, elongation factor-1α, and GAPDH are commonly used as reference genes in qRT-PCR experiments (Bustin, 2000; Olsvik *et al*., 2005). However, many recent studies have shown that expression levels of housekeeping genes can vary considerably within a sample set, making them unsuitable for use as reference genes for data normalization in qRT-PCR (Dhar *et al*., 2002; Tricarico *et al*., 2002; Vandesompele *et al*., 2002; Perez-Novo *et al*., 2005; de la Vega, 2006). Selection of suitable reference genes is not a simple process, and the use of three or more reference genes is becoming more widely adopted in order to achieve accurate quantification (Vandesompele *et al*., 2002; Meller *et al*., 2005). Reference genes for qRT-PCR can be reliably selected from microarray expression data

(Woods *et al*., 2004). Microarrays allow rapid screening of a large number of genes and can be used to determine a set of genes that are transcribed at a consistent level throughout a sample set. Unfortunately, for many organisms such a large dataset is not available, precluding the identification of genes whose transcripts are maintained at stable levels across the samples being surveyed.

Absolute quantification requires a suitable standard of known concentration (such as plasmid DNA containing a clone of the target gene) to make serial dilutions for establishing a standard curve. The standard curve gives a linear relationship between cycle threshold (Ct) values and the initial amount of starting template, allowing determination of the concentration of unknown samples based on their Ct values (Bustin, 2000; Wong and Medrano, 2005). This method can accurately quantify the number of template copies in a known amount of starting sample.

The choice of quantification procedure is dependent on the experimental design and availability of either characterized reference gene sequences or suitable template for construction of a standard curve. Absolute quantification is considered to be more labour-intensive than relative quantification due to the necessity to create a reliable set of standards (Pfaffl *et al*., 2004).

In the present study spatial and temporal expression of Llbglu1 gene is studied with gene specific primers and probe using quantitative PCR in developing seedlings of *L. leucocephala*. Initial transcripts of Llbglu1 in different tissues are determined by absolute quantification method.

5.2 Materials and Methods

5.2. 1. Isolation of total RNA

Leucaena leucocephala seedlings grown *in vitro* under culture room conditions were chosen for isolation of total RNA. For RNA isolation, root and shoot can be obtained from 5 day old seedlings, however leaves can be obtained only after 15 day old seedlings. The total RNA was isolated from root shoot and leaves samples of 20 day old

seedlings by TRIzol reagent (Sigma, USA) method. The isolation of total RNA was described previously (Chapter 2: Section 2.8.4.5) . Purity was analysed by measuring the OD at 260/280 and quantified spectrophotometrically.

5.2.2. cDNA first strand synthesis by Reverse Transcription

Total RNA was isolated from shoot, root, leaf of 0, 5, 10, 15 day old and 30 day old seedlings of *L. leucocephala* plants. An aliquot of total RNA normalized for uniform amplification of the 18S rRNA was used for synthesis of cDNA first strand which was used as template for quantitative PCR. cDNA first strand was synthesized using ImPromII™ Reverse Transcription System (Promega, USA) as per the manufacturer's guidelines. Detailed procedure was described in Chapter 2, Section 2.8.4.8.

5.2.3. Designing of primers and probes for QPCR Product Detection in Real Time

Primers and probe were synthesized from (Eurogentec S.A, Searing Belgium) from cDNA sequence of *Leucaena* β-glucosidase (Acc no. EU328158). Taqman probe methodology was used for measuring the fluorescence in real time. The Taqman reaction requires a hybridization probe labeled with two different fluorescent dyes. In the present study the reporter dye is a (FAM) and quenching dye is (BHQ).

5.2.4. Optimization of PCR Reaction conditions

Efficiency of target amplification was optimized by optimizing the concentration of Primers and Probe. ROX was used as reference dye and its concentration and cycling conditions were optimized.

5.2.5. Quantitative measurement of initial target sequences by CT Values

Computer software examines the fluorescence intensity of both the reporter dye (FAM) and the quenching dye (BHQ). The fluorescence intensity of the quenching dye, BHQ, changes very little over the course of the PCR amplification, therefore, the intensity of BHQ dye emission serves as an internal standard with which to normalize the reporter dye (FAM) emission variations. The software calculates a value termed dRn (or dRQ)

using the following equation: dRn = (Rn+)-(Rn-), where Rn + = emission intensity of reporter/emission intensity of quencher at any given time in a reaction tube, and Rn- = emission intensitity of re- porter/emission intensity of quencher measured prior to PCR amplification in that same reaction tube. For the purpose of quantitation, the last three data points (dRns) collected during the extension step for each PCR cycle were analyzed. Most PCR amplifications reach a plateau phase of reporter fluorescence emission if the reaction is carried out to high cycle numbers. The amplification plot is examined early in the reaction, at a point that represents the log phase of product accumulation. This is done by assigning an arbitrary threshold that is based on the variability of the base-line data. Once the threshold is chosen, the point at which the amplification plot crosses the threshold is defined as Ct. Ct is reported as the cycle number at this point and the value is predictive of the quantity of input target.

5.2.6. Spatial and temporal expression of Llbglu1 gene in developing seedlings of *L. leucocephala*

Leucaena seedlings were grown *in vitro* and cultured for different period of time. Total RNA was isolated from arial parts and roots of 0 day , 5 day, 10 day,15 day and 30 day old seedlings. Equal amoumt of RNA was used for synthesis of cDNA from all the samples. The relative level of expression of Llbglu1 was analysed by qRT-PCR. 18S RNA was used as internal control. Plasmid DNA of the clone cLlbglu1 containing the full-length insert of the Leucaena β-glucosidase was diluted serially and used for standard curve for absolute quantification. The initial transcript abundance of the Llbglu1 gene in arial parts and roots was determined by corresponding Ct values.

5.2.6.1. Spatial Expression profile of Llbglu1 gene

cDNA of *L. leucocephala* from root, shoot and leaves of 15 day old seedlings was used for qRT-PCR. 18S RNA was used as internal control for normalization of all cDNA samples.

5.2.6.2. Absolute quantification of Llbglu1 in different tissues of *L .leucocephala*

For absolute quantification, serial dilutions of the plasmid DNA of the clone cLlbglu (with *Leucaena* β-glucosidase insert) was done. The copy number and Ct values were

plotted and served as standard curve. The Ct values for root, shoot and leaves cDNA were extrapolated to determine the initial target sequences (Llbglu1) in those samples.

5.2.6.3. Temporal expression analysis of Llbglu1 gene in aerial parts.

cDNA was prepared from arial parts of 0 day, 5 day, 10 day ,15 day and 30 day old L. leucocephala seedling and used for qRT-PCR analysis. Expression pattern of the Llbglu1 gene was studied by determining the fluorescence (Ct values).

5.2.6.4. Temporal expression analysis of Llbglu1 gene in root

cDNA was prepared from roots of 0 day, 5 day, 10 day ,15 day and 30 day old L. leucocephala seedlings and used for qRT-PCR analysis. Expression pattern of the Llbglu gene was studied by determining the fluorescence (Ct values).

5.3. Results and discussions

5.3. 1. Isolation of total RNA

Total RNA was isolated from arial parts and roots of seedlings of *Leucaena* of different day old. Purity of isolated total RNA was confirmed by measuring OD at 260/280 nm and also by visualization on 1.5% TAE Agarose gel. Sample dilution was adjusted and absorbance of 1.0 at 260 nm corresponds to 50 µg RNA/ mL.

5.3.2. cDNA first strand synthesis by Reverse Transcription

Equal amount of RNA was used from all the samples for cDNA synthesis. 1µg of total RNA was used for the synthesis of cDNA from all tissue samples. (Described in chapter 2, Section 2.8.4.8)

5.3.3. Designing of primers and probes for PCR Product Detection in Real Time

Primers and probes for qRT-PCR are degined from Eurogentec, Belgium. The assay uses fluorescent Taqman methodology and an instrument capable of measuring fluorescence in real time (Stratagene Mx 3000P). The Taqman reaction requires a hybridization probe labeled with two different fluorescent dyes. One dye is a reporter

dye (FAM), the other is a quenching dye (BHQ). When the probe is intact, fluorescence energy transfer occurs and the reporter dye fluorescent emission is absorbed by the quenching dye (BHQ). During the extension phase of the PCR cycle, the fluorescent hybridization probe is cleaved by the 5'-3' nucleolytic activity of the DNA polymerase. On cleavage of the probe, the reporter dye emission is no longer transferred efficiently to the quenching dye, resulting in an increase of the reporter dye fluorescence emission spectra. Nucleotide regions chosen for primer and probe are shown in Fig 5.1.

```
  1 ATGATGAAGA AGGTGATGGT AGTAGCCGCC GTCGTGTGGG CTCTCATAAC
 51 TGTTGCTGCA GCTGATGCAA CAAATGATAT TTCCTCTCTC AGTCGCCGGA
101 GTTTCGCCCC GGCCTTCATT TTCGGCACCG CCTCCGCCTC CTACCAGTAC
151 GAAGGTGCAG CAAAGGAAGG CGGCAGAGGA CCCAGCATAT GGGATACCTT
201 CACCCACAAA TACCCAGAAA AAATATCGGA TCGCAGCAAC GGAGATGTAG
251 CTAACGACGA ATATCATCGG TATAAGGAGG ATGTTGGAAT CATGAAGTAT
```

Fig. 5.1.Region of nucleotide sequences used for primer and probe designing for qRT-PCR. Primers shown with arrow and probe with double arrow.

5.3.4 . qRT-PCR cycling conditions

Gene specific Primers and probe are used at a concentration of 100 nM each. The reference dye ROX was used at 30 nM concentration. Reactions were carried out with a Mx 3000P Real-Time PCR (Stratagene.USA). The amplification protocol is as 95 °C for 3 min, 40 cycles of 95 °C for 30 s, 65 °C for 45 s, 78 °C for 20 s, then 95 °C for 1 min and 55 °C for 1 min. Dissociation curves were obtained to confirm that each reaction yielded single, specific products.

5.3.5. Spatial and temporal expression analysis of Llbglu1 gene in developing seedlings of *L. leucocephala*

Leucaena seedlings were grown *in vitro* and cultured for different period of time. The expression pattern of Llbglu1 gene is studied by qRT-PCR, in arial parts and roots of 0 day, 5 day, 10 day,15 day and 30 day old seedlings. The relative abundance of initial transcript level of Llbglu1 is determined by the Ct values. 18S RNA was used as internal control. Plasmid DNA of the clone cLlbglu1 containing the full-length insert of the *Leucaena* β-glucosidase was diluted serially and used for standard curve for absolute quantification.

5.3.5.1. Spatial expression profile of Llbglu gene in seedlings of ***L. leucocephala*.** cDNA of *L. leucocephala* from root, shoot and leaves of 15 day old seedlings was used for qRT-PCR. 18S RNA was used as internal control for normalization of all cDNA samples. PCR resulted into single amplification plot for the respective cDNA sample (Shown in Fig 5.2)

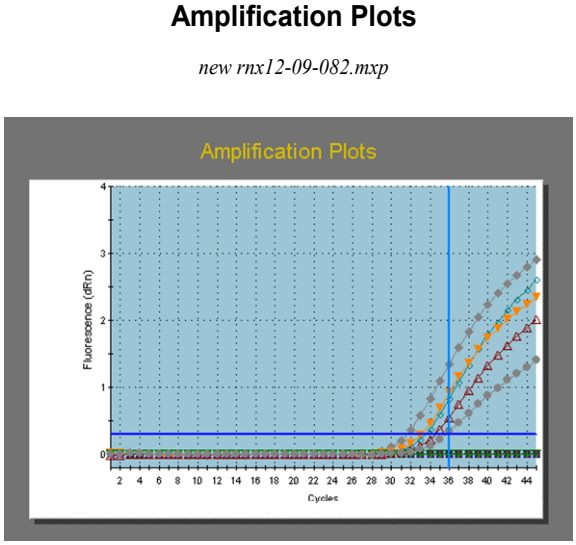


Fig. 5.2. qRT-PCR Amplification plots Llbglu1 gene with cDNA of root, shoot and leaves of 15 day seedlings of *L. leucocephala*.

The reference dye ROX was used for normalization. The baseline threshold fluorescence was set as 0.3053. Ct values for different cDNA samples are shown in table 5.1.

Table 5.1 qRT-PCR data for Llbglu1 gene expression in root, shoot and leaves.

Assay	Well Type	Threshold (dRn)	Ct (dRn)
ROX	15 day root	Reference	Reference
FAM	15 day root	0.3053	31.52
ROX	15 day root	Reference	Reference
FAM	15 day root	0.3053	31.67
ROX	15 day shoot	Reference	Reference
FAM	15 day shoot	0.3053	**33.59**

ROX	15 day shoot	Reference	Reference
FAM	15 day shoot	0.3053	33.18
ROX	15 day leaves	Reference	Reference
FAM	15 day leaves	0.3053	34.61
ROX	15 day leaves	Reference	Reference
FAM	15 day leaves	0.3053	35.72

The organ-specific expression of Llbglu1 gene in *L .leucocephala* seedlings was examined by quantitative real-time (QRT)-PCR analysis. As shown in Fig. 5.3, the Llbglu1 was expressed in all organs under the natural conditions. Maximum transcripts expressed in roots, while relatively less accumulation was detected in stems and followed by leaves (shown in Fig.5.3).

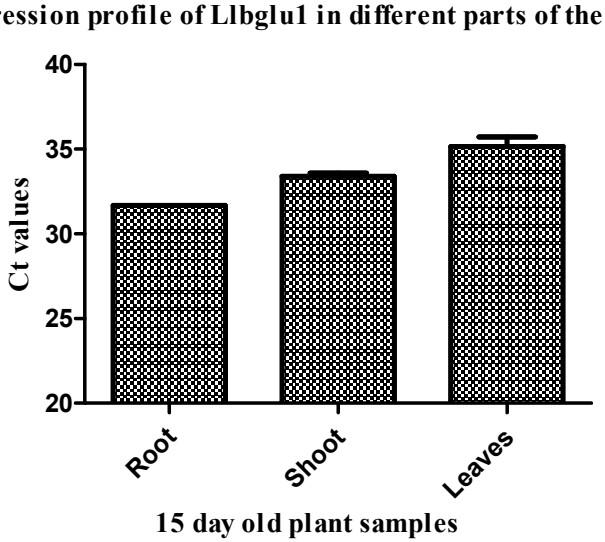


Fig. 5.3 Expression pattern of Llbglu1 gene in different organs of the 15 day old seedlings of *L. leucocephala*

5.3.5.2.Absolute quantification of Llbglu1 in different tissues of *L.leucocephala*.

For absolute quantification of the Llbglu gene transcripts in different cDNA samples, a standard curve is prepared using the plasmid DNA of the clone cLlbglu (with *Leucaena* β-glucosidase insert). The copy number and Ct values were plotted and served as standard curve. The Ct values for root, shoot and leaves cDNA were extrapolated to determine the initial target sequences (Llbglu1). The transcript copy number in root found to be maximum (i.e 6125) shown in Fig 5.4.

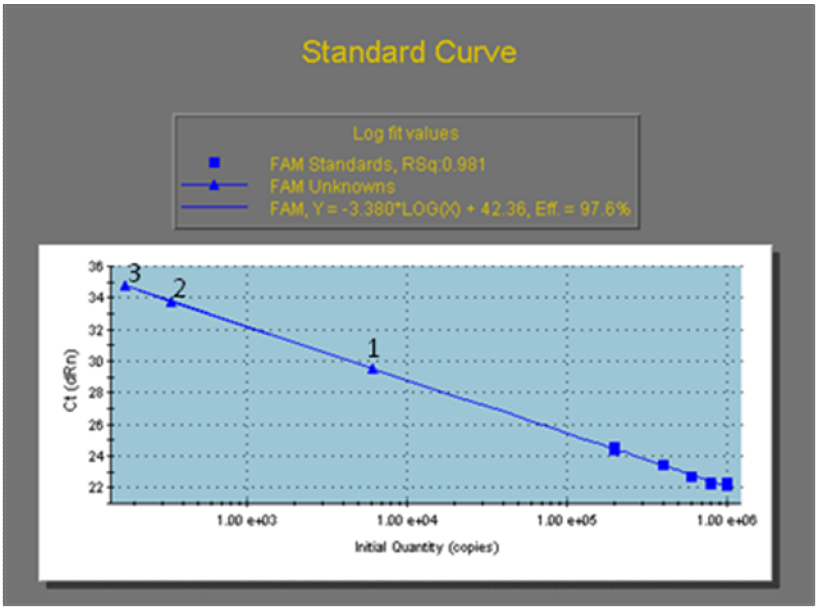


Fig 5.4. Standard curve showing the initial template copies of the clone cLlbglu1 and Ct values. Initial copies of LLbglu1 in root, shoot and leaves samples are indicated by ▲ symbol with numbers 1, 2 and 3 respectively.

Table 5.2. Initial template copies of the clone cLlbglu1 and Ct values used for standard and initial copies of Llbglu1 in root, shoot and leaves.

	Initial plasmid Quantity (copies)	Ct (dRn)	Tissue	
FAM Standards, RSq:0.981				
	200000	24.52919		
	200000	24.41062		
	400000	23.47164		
	600000	22.69722		
	800000	22.2481		
	800000	22.28556		
	1000000	22.1543		
	1000000	22.31793		
FAM Unknowns				
	6125.094	29.56037	Root	
	342.5135	33.79416	Shoot	
	177.0872	34.76262	Leaves	
FAM, Y = -3.380*LOG(X) + 42.36, Eff. = 97.6%				

From the graph it can be inferred that the expression of Llbglu1 is maximum in roots followed by shoot and leaves. In *Lotus japonicus* three β-glucosidases namely, LjBGD2, LjBGD4 and LjBGD7 are found. They are involved in defense, and there was almost 98 % sequence identity among the sequences. Based on the RT-PCR analysis, LjBGD2 and LjBGD4 are expressed in aerial parts of *L. japonicus*, with an apparent highest expression level in young leaves, and LjBGD7 is expressed exclusively in roots (Morant *et al*., 2008). LjBGD7 has 77% identity with that of Llbglu1.

5.3.5.3. Temporal expression analysis of Llbglu1 gene in aerial parts.

cDNA of L .leucocephala was prepared from arial parts of 0 day, 5 day, 10 day ,15 day and 30 day old seedlings and used for qRT-PCR. Baseline fluorescence was set as 0.3445. Ct values for different day old shoot is shown in Table 5.3.

Table 5.3. qRT-PCR data for Llbglu1 gene expression in arial parts

Assay	Well Type	Threshold (dRn)	Ct (dRn)
ROX	0 day shoot	Reference	Reference
FAM	0 day shoot	0.3445	32.54
ROX	0 day shoot	Reference	Reference
FAM	0 day shoot	0.3445	32.34
ROX	5 day shoot	Reference	Reference
FAM	5 day shoot	0.3445	31.09
ROX	5 day shoot	Reference	Reference
FAM	5 day shoot	0.3445	31.86
ROX	10 day shoot	Reference	Reference
FAM	10 day shoot	0.3445	31.75
ROX	10 day shoot	Reference	Reference
FAM	10 day shoot	0.3445	31.22
ROX	15 day shoot	Reference	Reference
FAM	15 day shoot	0.3445	33.81
ROX	15 day shoot	Reference	Reference
FAM	15 day shoot	0.3445	33.43
ROX	30 day shoot	Reference	Reference
FAM	30 day shoot	0.3445	33.17
ROX	30 day shoot	Reference	Reference
FAM	30 day shoot	0.3445	33.15

The expression pattern of Llbglu1 gene in arial parts remained almost the same under nomal conditions for the first 10 days and later decreased upto 30 days (shown in Fig.5.5).

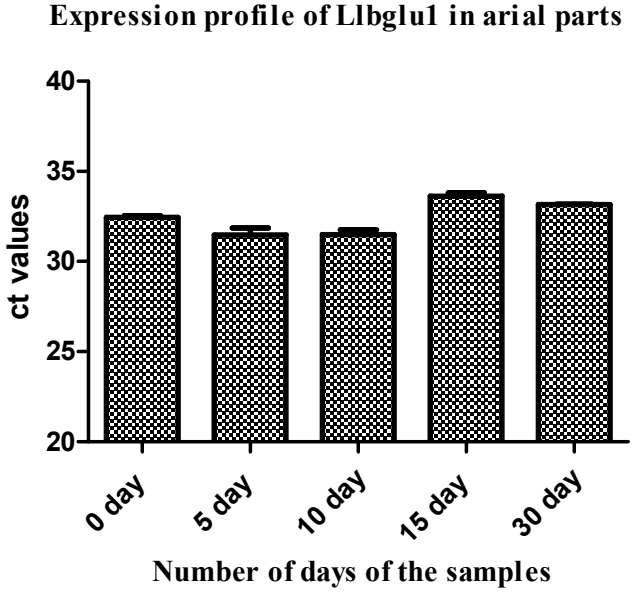


Fig 5.5 Expression profile of Llbglu gene in arial parts of developing seedlins of *Leucaena leucocephala*.

5.3.5.4. Temporal expression analysis of Llbglu gene in root

Assay	Well Type	Threshold (dRn)	Ct (dRn)
ROX	0 day root	Reference	Reference
FAM	0 day root	0.3487	34.65
ROX	0 day root	Reference	Reference
FAM	0 day root	0.3487	34.51
ROX	5 day root	Reference	Reference
FAM	5 day root	0.3487	33.33
ROX	5 day root	Reference	Reference
FAM	5 day root	0.3487	33.03
ROX	10 day root	Reference	Reference
FAM	10 day root	0.3487	31.53
ROX	10 day root	Reference	Reference
FAM	10 day root	0.3487	31.38
ROX	15 day root	Reference	Reference
FAM	15 day root	0.3487	31.62

ROX	15 day root	Reference	Reference
FAM	15 day root	0.3487	31.91
ROX	30 day root	Reference	Reference
FAM	30 day root	0.3487	35.66
ROX	30 day root	Reference	Reference
FAM	30 day root	0.3487	34.38

cDNA of *L. leucocephala* was prepared from roots of 0 day, 5 day, 10 day ,15 day and 30 day old seedlings and used for qRT-PCR. Baseline fluorescence was set as 0.3487. Ct values for different day old root are shown in Table 5.4.

Table 5.4. qRT-PCR data for Llbglu1 gene expression in roots.

The expression pattern of Llbglu1 gene in different day old root is shown in Fig. 5.6. Under normal conditions the expression of the gene increased gradually from day 0 day to day 10, later its expression decreased from 15day to 30 day old root.

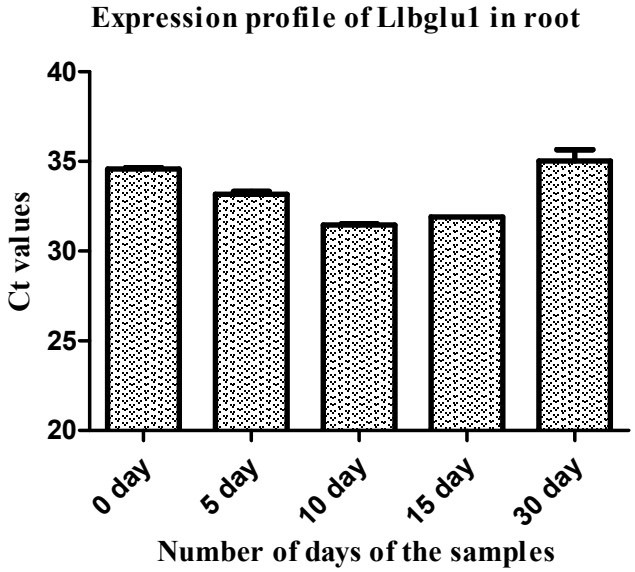


Fig.5.6. Expression profile Llbglu1 gene in roots of developing seedlings of *L. leucocephala.*

The preliminary quantitative PCR data shows that under normal conditions the Llbglu1 gene is expressed in both arial parts and underground parts in developing seedlings. However, when relative level of expression of Llbglu1 is analyzed, maximum number

of transcripts of the gene are observed in roots than arial tissues. However in support of this view, further studies like northern blot and enzyme activity from these tissues are needed.

5.4. Conclusions

In the present study spatial and temporal expression of Llbglu1 gene is studied with gene specific primers and probe using quantitative PCR in developing seedlings of *L .leucocephala*. Initial transcripts of Llbglu1 in different tissues are determined by absolute quantification method.

Leucaena leucocephala seedlings grown *in vitro* under culture room conditions were chosen for isolation of total RNA. For RNA isolation, root and shoot can be obtained from 5 day old seedlings, however leaves can be obtained only after 15 day old seedlings. The total RNA was isolated from root, shoot and leaves samples of 20 day old seedlings by TRIzol reagent (Sigma, USA) method. The isolation of Total RNA was described previously (Chapter 2: Section 2.8.4.5) . Purity was analysed by measuring the OD at 260/280 and quantified spectrophotometrically.

Total RNA was isolated from shoot, root, leaf of 0, 5, 10, 15 day old and 30 day old seedlings of *L. leucocephala* plants. An aliquot of total RNA normalized for uniform amplification of the 18S rRNA was used for synthesis of cDNA first strand which was used as template for quantitative PCR. cDNA first strand was synthesized using ImPromII™ Reverse Transcription System (Promega, USA) as per the manufacturer's guidelines.

Primers and probe were synthesized from (Eurogentec S.A, Searing Belgium) from cDNA sequence of *Leucaena* β-glucosidase (Acc no EU328158). Taqman probe methodology was used for measuring the fluorescence in real time. The Taqman reaction requires a hybridization probe labeled with two different fluorescent dyes. In the present study the reporter dye is a (FAM) and quenching dye is (BHQ).

Leucaena seedling were grown *in vitro* and cultured for different period of time. The expression pattern of Llbglu1 gene is studied by qRT-PCR, in arial parts and roots of

0 day , 5 day, 10 day,15 day and 30 day old seedlings. The relative abundance of initial transcript level of Llbglu1 is determined by the Ct values.18S RNA was used as internal control. Plasmid DNA of the clone cLlbglu1 containing the full-length insert of the *Leucaena* β-glucosidase was diluted serially and used for standard curve for absolute quantification.

The organ-specific expression of Llbglu1 gene in *L .leucocephala* seedlings was examined by quantitative real-time (QRT)-PCR analysis. As shown in Fig. 5.3, the Llbglu1 was expressed in all organs under the natural conditions. Maximum transcripts expressed in roots, while relatively less accumulation was detected in stems and followed by leaves. In *Lotus japonicus* three β-glucosidases namely, LjBGD2, LjBGD4 and LjBGD7 are found. They are involved in defense, and there was almost 98 % sequence identity among the sequences. Based on the RT-PCR analysis, LjBGD2 and LjBGD4 are expressed in aerial parts of *L. japonicus*, with an apparent highest expression level in young leaves, and LjBGD7 is expressed exclusively in roots (Morant *et al*., 2008). LjBGD7 has 77% identity with that of Llbglu1. The expression pattern of Llbglu1 gene in arial parts remained almost the same under nomal conditions for the first 10 days and later decreased upto 30 days. The preliminary quantitative PCR data shows that under normal conditions the Llbglu1 gene is expressed in both arial parts and underground parts in developing seedlings. However when relative level of expression of Llbglu1 is analyzed, maximum number of transcripts of the gene are observed in roots than arial tissues.

Summary

β-Glycosidases that belong to the family 1 glycoside hydrolases catalyze hydrolysis of the β-glycosidic bond in β-glycosides consisting of two carbohydrate moieties or a carbohydrate moiety linked to an aryl or alkyl aglucone. In plants, β-glycosidases serve a number of diverse and important functions, including bioactivation of defense compounds, cell wall degradation in endosperm during germination, activation of phytohormones, and lignification . Plants produce myriad of secondary metabolites involved in defense against pathogens and herbivores. These defense compounds are often stored as β-glycosides and bioactivated by specific β-glycosidases. Glycosylation serves to protect the plant against the toxic effects of its own chemical defense system, to increase solubility, and to facilitate storage. Examples of two-component plant defense systems wherein β-glycosidases act as the bioactivator include the α-hydroxynitrile glycosides (cyanogenic glycosides) that are found in numerous different plant, benzoxazinoid glycosides graminae, avenacosides in Avena sativa , isoflavonoid glycosides in legumes and glucosinolates in brassicales.

In this thesis a cDNA clone encoding a Family 1 Glycosyl hydrolase has been isolated from *Leucaena leucocephala* and characterized. The enzyme was heterologously expressed and purified. The recombinant enzyme was characterized by both biochemical and computational methods. The expression profile of the β-glucosidase gene was studied spatio-temporally in developing seedlings by quantitative RT-PCR.

Isolation, Cloning and Characterization of a β-glucosidase gene from *Leucaena leucocephala*

Primers were designed from the selected Family 1 glycosyl hydrolases that are involved in defense, from the NCBI database. *Leucaena* gDNA and cDNA was used to amplify partial sequences 390 bp and 141 bp of the β-glucosidase.The sequences were submitted to NCBI genbank database under accession number DQ417200 and DQ883447 respectively. Rapid Amplification of cDNA Ends (RACE, 3' RACE and 5' RACE) was done to obtain the fulllength cDNA clone of the β-glucosidase. 3'RACE has resulted into amplification of ~1.2 kb fragment, it has a coding sequence of 1183 bp

followed by stop codon TAA. The sequence has a 3'untranslated region(3'UTR) of 119 bp followed by a poly A tail. In 5'RACE an amplicon of approximately 1.0 kb was amplified and cloned. A 18 bp 5'UTR was observed followed by start codon ATG. The sequence was submitted to NCBI GenBank database (Accession no. EF061245**).** Gene specific forward and reverse primers were designed from the start and stop codon of 5'and 3'RACE sequences respectively. PCR was done using cDNA as template. An expected size of approximately 1.5 kb amplicon was amplified. The amplicon was cut, eluted from gel and ligated into pGEMT-Easy vector. It was confirmed by restriction digestion with *EcoRI* which released 1.5 kb fragment (Fig.3.18). The clone now referred as cLlbglu1 was sequenced. The sequence showed 77% identity with *Lotus* β-glucosidase (EU710846) and 78% identity with Rose β-glucosidase (AB426520). The sequence was submitted to NCBI GenBank database (EU328158). Slot blot hybridization showed the presence of at least two copy number of the gene. Multiple sequence alignment of nucleotide and deduced amino acid sequence of Llbglu1 with other Family1 GHs involved in defence was done. Phylogenetic and functional relationship of *Leucaena leucocephala* glycosyl hydrolase with other Family 1 glycosyl hydrolases revealed that it is a β-glucosidase involved in defense. It clustred with *Lotus japonicas* β-glucosidases which hydrolyses the hydroxynitrile glucosidases.

Cloning, expression and purification of *Leucaena leucocephala* β-glucosidase in prokaryotic system

The Llbglu1 was cloned without signal sequence, into pET 28 a(+) vector at *EcoR I* and *Not I* sites. The plasmid (pET-Llbglu1-His) was mobilized in to *E.coli* BL 21(DE3). Individual colonies were screened for protein expression. Positive colonies were selected for recombinant protein expression. Growth conditions of the culture like temperature, IPTG concentration and period of incubation were optimized so that expressed protein was in soluble form (active form). Enzyme activity was monitored in crude lysate by assaying for β-glucosidase activity with PNPG as substrate. Temperatures 18 ^{o}C, 0.05 mM IPTG and 9 h of IPTG induction were optimum for the production of active enzyme. The recombinant enzyme was purified by Immobilized Metal Affinity chromatography, using Ni-NTA agarose beads (Qiagen, USA). The purified enzyme concentration was determined by Lowry's methods using BSA as

standard. The purified protein was analyzed on SDS-PAGE and it has a molecular weight of ~55kD. Enzyme assay conditions for the recombinant *Leucaena* β-glucosidase were optimized. pH 4.8 (100-200 mM Citrate-Phophate buffer), 45^{o}C were found to be optimum.. The purified recombinant Llbglu1 hydrolase was incubated at its optimum pH (pH 4.8) with potential substrates provided at 2 mM final concentration. Enzyme activity was determined by measuring the rate of PNP (or ONP) production spectrophotometrically at 420 nm with subsequent use of standard curves. The recombinant enzyme hydrolysed both PNP-glucoside and PNP-galactoside. To find the glucone specificity, kinetic constants like *Km* and *Vmax* were determined for these substrates. *Km* and *Vmax* was found to be 40 μM and Vmax 0.8278 μM/mg/min respectively for PNP-glycoside. Km and Vmax was found to be 1.75mM and Vmax 0.1008 μM/mg/min for PNP-galactoside. To find the aglycone specificity two synthetic glucosides VRA-G and 4MUG were used.the km was was found to be 3.20 mM and 89.28 μM respectively.

Bioinformatics analysis of the *Leucaena* β-glucosidase was carried out for further characterization**.** Secondary structure prediction of L. leucocephala β-glucosidase was compared with *Trifolium* cyanogenic β-glucosidase (1CBG). The structure of white clover β-glucosidase was used as template to generate the 3D model of the Llbglu1 using the program Modeler (Marti-Renom *et al*., 2000). The modeled structure of Llbglu1 was visualized using program PyMoL (DeLano ,WL; 2002) . In Leucaena β-glucosidase a $(\alpha/\beta)_8$ fold was observed, which is a common feature among the Family 1 β-glucosidases. The quality of the modeled structure was checked with PROCHECK program and Ramachandran plot. Pairwise structural superimposition of modeled *Leucaena* β-glucosidase was done with *Trifolium* cbg using Combinatorial extension algorithm at SDSC-CE (http://cl.sdsc.edu/ce.html). In family 1 GH s Glutanic acid serves as catalytic proton donor/base and nucleophile. In *Trifolium* cyanogenic β-glucosidase those catalytic residues Glu 183 and Glu 397 corresponds Glu 199 and Glu 413 in Llbglu1 respectvely. Analysis of amino acids involved in Glucose binding and catalysis in *Leucaena* β-Glucosidase with other β- glucosidases was done by structural alignment at CEMC server (http://pathway.rit.albany.albany.edu/~cemc/. A natural hydroxynitrle glucoside, Rhodiocyanoside A was docked using program Accelrys, in to

the active pocket of Llbglu1, to find the residues invoved in substrate binding and catalysis.

Spatial expression analysis of the β-glucosidase gene in different tissues of *Leucaena leucocephala*

In the present study spatial and temporal expression of Llbglu1 gene is studied with gene specific primers and probe using quantitative PCR in developing seedlings of *L. leucocephala*. Initial transcripts of Llbglu1 in different tissues are determined by absolute quantification method. *Leucaena leucocephala* seedlings grown *in vitro* under culture room conditions were chosen for isolation of total RNA. *Leucaena* seedlings were grown *in vitro* and cultured for different period of time. The expression pattern of Llbglu1 gene is studied by qRT-PCR, in arial parts and roots of 0 day, 5 day, 10 day, 15 day and 30 day old seedlings. The relative abundance of initial transcript level of Llbglu1 is determined by the Ct values.18S RNA was used as internal control. Plasmid DNA of the clone cLlbglu1 containing the full-length insert of the *Leucaena* β-glucosidase was diluted serially and used for standard curve for absolute quantification.The organ-specific expression of Llbglu1 gene in *L. leucocephala* seedlings was examined by quantitative real-time (QRT)-PCR analysis. The Llbglu1 was expressed in all organs under the natural conditions. Maximum transcripts expressed in roots, while relatively less accumulation was detected in stems and followed by leaves

Future prospects

1) Isolation, cloning and characterization of this Family 1 glycosyl hydrolase from *Leucaena* will help in understanding the biosynthesis and metabolism of defense β-glucosides in this plant and plant based resistance to insect or pathogen attack.
2) Manipulation of regulation of this gene (overexpression/suppression) will help in developing transgenic *Leucaena* with improved resistance to insect and pathogen.
3) The gene can also be transferred to other compatible crop plants for improved resistance to insect or pathogen attack.
4) Crytal structure studies of this Family 1 glycosyl hydrolase enzyme will help in understanding the amino acid residues involved in aglucone binding and specificity.

References

Accelrys, 2003. *Structure-Based Drug Design with Discovery Studio, Accelrys, version 0406*, Accelrys Software Inc., San Diego

Agrawal, A.A., Kurashige, N.S., 2003. A role for isothiocyanates in plant resistance again the specialist herbivore Pieris rapae. J. Chem. Ecol. 29, 1403–1415.

Aharoni, A., Giri, A.P., Deuerlein, S., Griepink, F., de Kogel, W.J., Verstappen, F.W.A., Verhoeven, H.A., Jongsma, M.A., Schwab, W., Bouwmeester, H.J., 2003. Terpenoid metabolism in wild-type and transgenic Arabidopsis plants. Plant Cell. 15, 2866–2884.

Akiyama, T., Kaku, H., Shibuya, N., 1998. A cell wall-bound β-glucosidase from germinated rice: purification and properties. *Phytochemistry*. 48, 49-54.

Allison, M.J., Hammond, A.C., Jones, R.J., 1990. Detection of ruminal bacteria that degrade toxic dihydroxypyridine compounds produced from mimosine. Appl Environ Microbiol. 56, 590-594.

Arnheim, N., Erlich, H., 1992. Polymerase Chain Reaction Strategy. Annu Rev Biochem. 61, 131-156.

Austin, M.T., Early, R.J., Brewbaker, J.L., Sun, W., 1997. Yield, psyllid resistance, and phenolic concentration of Leucaena in two environments in Hawaii. Agronomy Journal. 89, 507–515.

Babcock, G.W., Esen, A., 1994. Substrate specificity of maize β-glucosidase. *Plant Sci*. 101,31-39.

Baird, S.D., Hefford, M.A., Johnson, D.A., Sung, W.L., Yaguchi,M. Seligy, V.L., 1990. The Glu residue in the conserved Asn-Glu-Pro sequence of two highly divergent endo _-1,4-glucanasesis essential for enzyme activity. Biochem. Biophys. Res. Comm.169, 1035–1039.

Bak, S., Paquette, S., Morant, M., Morant, A.V., Saito, S., Bjarnholt, N., Zagrobelny, M., Jørgensen, K., Osmani, S., Hamann, T., Simonsen, H.T., Perez, R.S., van Heeswijck, T.B., Jørgensen, B., Møller, B.L., 2006. Cyanogenic glucosides: a case study for evolution and application of cytochromes P450. Phytochem. Rev. 5, 309–329.

Barleben, L., Panjikar, S., Ruppert, M., Koepke, J., Stockigta J., 2007. Molecular Architecture of Strictosidine Glucosidase: The Gateway to the Biosynthesis of the MonoterpenoidIndole Alkaloid Family. The Plant Cell. 19, 2886–2897.

Barrett,T., Suresh,C.G., S.P. Tolley,S.P., Dodson,E.J., Hughes,M.A., 1995. The crystal structure of a cyanogenic β-glucosidase (linamarase) from white clover (*Trifolium repens* L.), a family 1 glycosyl hydrolase.Structure. 3, 951–960.

Baumeler, A., Hesse, M., Werner,C., 2000. Benzoxazinoids-cyclic hydroxamic acids, lactams and their corresponding glucosides in the genus Aphelandra (Acanthaceae). Phytochemistry 53, 213–222.

Be´ guin, P. 1990. Molecular biology of cellulose degradation. Annu. Rev. Microbiol. 44, 219–248.

Beale, M.H., Birkett,M.A., Bruce, T.J.A., Chamberlain, K., Field, L.M., Huttly, A.K., Martin, J.L., Parker, R., Phillips, A.L., Pickett, J.A., 2006. Aphid alarm pheromone produced by transgenic plants affects aphid and parasitoid behavior. Proc Natl Acad Sci USA. 103, 10509–10513.

Berrin, J.G., Czjzek, M., Kroon, P.A., McLauchlan, W.R., Puigserver, A., Williamson, G. and Juge, N. 2003. Substrate (aglycone) specificity of human cytosolic β-glucosidase. Biochem. J. 373, 41–48.

Bjarnholt, N., Rook, F., Motawia, M.S., Cornett, C., Jørgensen, C., Olsen, C.E., Jaroszewski, J.W., Bak, S., Møller, B.L., 2008. Diversification of an ancient theme: hydroxynitrile glucosides. Phytochemistry. 69, 1507–1516.

Blumberg, D.D., 1987. Creating a ribonuclease-free environment. Meth. Enzymol. 152, 20-24.

Brewbaker, J.L., Sorenson, C.T., 1990. Leucaena: New tree crops from interspecific hybrids. In: Janick. J and Simon J. (Eds). Advances in new crops Timber Press, Oregon. pp 238

Brown, P.D., Morra, M.J., 1995. Glucosinolate-containing plant tissues as bioherbicides. J. Agric. Food Chem. 43, 3070–3074.

Burgos, N.R., Talbert, R.E., 2000. Differential activity of allelochemicalsfrom Secalecereale in seedling bioassays. Weed Sci. 48, 302–310.

Burmeister, W.P., Cottaz, S., Driguez, H., Iori, R., Palmieri, S., Henrissat, B., 1997. The crystal structure of Sinapis alba myrosinase and a covalent glycosyl-

enzyme intermediate provide insights into the substrate recognition and activesite machinery of an S-glycosidase. Structure. 5, 663–675.

Bustin, S.A., 2000. Absolute quantification of mRNA using real-time reverse transcription polymerase chain reaction assays. J. Mol. Endocrinol. 25, 169–193.

Cairns, J.R.K., Champattanachai, V., Srisomsap, C., Wittman-Liebold, B., Thiede, B., Svasti, J., 2000. Sequence and expression of Thai rosewood β-glucosidase/β-fucosidase, a family 1 glycosyl hydrolase glycoprotein. J Biochem. 128, 999–1008.

Cambier, V., Hance, T., De Hoffmann, E., 2001. Effects of 1,4-benzoxazin-3-one derivatives from maize on survival and fecundity of *Metopolophium dirhodum* (Walker) on artificial diet. J. Chem. Ecol. 27, 359–370.

Chuankhayan, P., Hua, Y., Svasti, J., Sakdarat, S., Sullivan, P.A., KetudatCairns,J.R., 2005 Purification of an isoflavonoid 7-O-β-apiosyl-glucoside β-glucosidase and its substrates from *Dalbergia nigrescens* Kurz. *Phytochemistry*. 66, 1880-1889.

Chuankhayan, P., Rimlumduan, T., Svasti, J., Cairns, J.R.K., 2007a. Hydrolysis of soybean isoflavonoid glycosides by Dalbergia b-glucosidases. J Agric Food Chem. 55,2407–2412.

Chuankhayan, P., Rimlumduan, T., Tantanuch,W., Kongsaeree, P.T.,Metheenukul, P., Svasti, J., Jensen, O.N., Cairns, J.R.K., 2007b. Functional and structural differences between isoflavonoid b glycosidases fromDalbergia sp. Arch Biochem Biophys. 468, 205–216.

Chumakov, K.M., 1994. Reverse transcriptase can inhibit PCR and stimulate primer dimerformation. PCR Methods Appl. 4, 62-64.

Conn, E.E., 1980. Cyanogenic Compounds. Annu. Rev. Plant Physiol. Plant Mol. Biol. 31, 433–451.

Conn,E.E., 1993. β -Glucosidases in plants: substrate speciWcity, in: A. Esen (Ed.), β-Glucosidases Biochemistry and Molecular Biology, American Chemical Society, Washington, DC, , pp. 15–26.

Coutinho, P.M., Henrissat, B. 1999. Carbohydrate-Active Enzymes server at URL: http://afmb.cnrs-mrs.fr/_cazy/ CAZY/index.html.

Czjzek, M., Cicek, M., Zamboni, V., Bevan, D.R., Henrissat, B., Esen, A., 2000. The mechanism of substrate (aglycone) specificity in b-glucosidases is revealed by crystal structures of mutant maize b-glucosidase-DIMBOA, -DIMBOAGlc, and -dhurrin complexes. Proc. Natl. Acad. Sci. USA. 97, 13555–13560.

Davies, G., Henrissat, B., 1995. Structures and mechanisms of glycosyl hydrolases. Structure. 3, 853–859.

de la Vega, E., 2006. A molecular approach to study the interaction between environmental stress, immune response and disease in the black tiger prawn (*Penaeus monodon*). Ph.D. Thesis. University of Queensland, Queensland, Australia.

DeLano ,W.L., 2002.The pymol user's Manual DeLano scientific ,San Carlos,CA, USA.

Dhar, A.K., Roux, M.M., Klimpel, K.R., 2002. Quantitative assay for measuring the Taura syndrome virus and yellow head virus load in shrimp by real-time RT-PCR using SYBR Green chemistry. J. Virol. Methods 104, 69–82.

Dharmawardhana, D.P., Ellis, B.E., Carlson, J.E., 1995. A β-glucosidase from lodgepole pine xylem specific for the lignin precursor coniferin. Plant Physiol. 107, 331–339.

Dharmawardhana, D.P., Ellis, B.E., Carlson, J.E., 1999. cDNA cloning and heterologous expression of coniferin β-glucosidase. Plant Mol. Biol. 40, 365–372.

Durham, P.L., J.E. Poulton, J.E., 1989.Effect of castanospermine and related polyhydroxyalkaloids on puriWed myrosinase from *Lepidium sativaum* seedlings, Plant Physiol. 90, 48.

Duroux, L., Delmotte, F.M., Lancelin, J-M., Keravis, G., Jay-Alleand, C., 1998. Insight into naphthoquinone metabolism: β-glucosidase-catalysed hydrolysis of hydrojuglone β-D-glucopyranoside. *Biochem J*. 333, 275-283.

Eckert, K.T., Kunkel, T.A., 1990. High fidelity DNA synthesis by the Thermus aquaticus DNA polymerase. Nucleic Acids Res. 18,3739- 3744.

Eksittikul,Y., Chulavatanatol,M., 1988. Characterization of cyanogenic β-glucosidase (linamarase) from cassava (*Manihot esculenta* Crantz), Arch. Biochem. Biophys. 266, 263–269.

Escamilla-Trevino, L.L., Chen,W., Card, M.L., Shih,M.C., Cheng, C.L., Poulton, J.E., 2006. Arabidopsis thaliana β-glucosidases BGLU45 and BGLU46 hydrolyse monolignol glucosides. Phytochemistry. 67, 1651–1660.

Esen, A., 1993. β-glucosidases overview. In β-*glucosidases: Biochemistry and Molecular Biology* Edited by: Esen A. Washington DC: American Chemical Society.1-14. ACS Symposium Series 533.

Fahey, J.W., Zhang, Y.S., Talalay, P., 1997. Broccoli sprouts: an exceptionally rich source of inducers of enzymes that protect against chemical carcinogens. Proc. Natl. Acad. Sci. USA. 94, 10367–10372.

Falk, A., Rask, L., 1995. Expression of a zeatin-O-glucosidedegrading β-glucosidase in Brassica napus. Plant Physiol. 108: 1369–1377.

Feinbeng, A.P., Vogelstein, B., 1983. A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. Anal Biochem. 132, 6-13.

Feinbeng, A.P., Vogelstein, B., 1984. A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. Addendum. Anal Biochem. 137, 266-267.

Ferreira,C., Terra, W.R., 1983. Physical and kinetic properties of a plasmamembrane bound β-glucosidase (cellobiase) from the midgut cells of an insect (*Rhynchostiara americana* larva), Biochem. J. 213, 43–51.

Fia, G., Giovani, G., Rosi, I., 2005. Study of β-glucosidase production by wine-related yeasts during alcoholic fermentation: a new rapid fluorimetric method to determine enzymatic activity. J Appl Microbiol. 99, 509–517.

Fowler, T., 1993. Deletion of the Trichoderma reesei β-glucosidase gene, bgl1. In: A. Esen (Ed.) β-Glucosidases: Biochemistry and Molecular Biology, ACS Symposium Series 533. American Chemical Society, Washington, DC, pp. 56–65.

Freeman, W.M., Walker, S.J., Vrana, K.E., 1999. Quantitative RT-PCR: pitfalls and potential. Biotechniques. 26, 112-122, 124-125.

Frey, M., Chomet, P., Glawischnig, E., Stettner, C., Grun, S., Winklmair, A., Eisenreich, W., Bacher, A., Meeley, R.B., Briggs, S.P., Simcox, K., Gierl, A., 1997. Analysis of a chemical plant defense mechanism in grasses. Science. 277, 696–699.

Frohman, M.A., Dush, M.K., and Martin, G.R., 1988. Rapid production of full-length cDNAs from rare transcripts: amplification using a single gene specific oligonucleotide primer. In: *Proc. Natl. Acad. Sci. USA*. **85**, pp. 8998–9002 .

Frost, C.J., Mescher, M.C., Carlson, J.E., Moraes, C.M.D., 2008. Plant Defense Priming against Herbivores: Getting Ready for a Different Battle. Plant Physiology. 146, 818–824

Gabrys, B., Tjallingii, W.F., 2002. The role of sinigrin in host plant recognition by aphids during initial plant penetration. Entomol. Exp. Appl. 104, 89–93.

Geerlings, A., Matinez-Lozano., Ibañez, M., Memelink, J., van der Heijden, R., Verpoorte R., 2000. Molecular cloning and analysis of strictosidine β-D-glucosidase, an enzyme in terpenoid indole alkaloid biosynthesis in *Catharanthus roseus*. *J Biol Chem*. 275, 3051-3056.

Geiger, C.A., Napompeth B., Van Den Beldt R., 1995. An update on the status of leucaena psyllid in Southeast Asia. In: Shelton H.M., Piggin C.M. and Brewbaker J.L. _eds_, *Leucaena* – Opportunities and Limitations. Proceedings of an International Workshop held in Bogor, Indonesia, 24-29 January 1994.

Gentle, A., Anastasopoulos, F., McBrien, N.A., 2001. High resolution semi-quantitiative real-time PCR without the use of a standard curve. BioTechniques. 31, 502–508.

Givovich, A., Niemeyer, H.M., 1995. Comparison of the effect of hydroxamic acids from wheat on 5 species of cereal aphids. Entomol. Exp. Appl. 74, 115–119.

Givovich, A., Sandstrom, J., Niemeyer, H.M., Pettersson, J., 1994. Presence of a hydroxamic acid glucoside in wheat phloem sap, and its consequences for performance of Rhopalosiphum padi (L) (Homoptera, Aphididae). J. Chem. Ecol. 20, 1923–1930.

Goodenow, M., Huet, T., Saurin, W., Kwok, S., Sninsky, J., Wain-Hobson, S., 1989. HIV-1 isolates are rapidly evolving quasi species: evidence for viral mixtures and preferred nucleotide substitutions. J Acquired Immunol Defic Syndr. 2, 344-352.

Goute, P., Courcelle, E.,Stuart, D.I.,Metoz, F.,1999.ESPript:analysis of multiple sequence alignments in PostScript.Bioinformatics.15,305-308.

Guda, C., Lu, S., Sheeff, E.D., Bourne, P.E, Shindyalov, I.N., 2004. CE-MC: A multiple protein structure alignment server, *Nucleic Acids Res.*32:W100-W103 .

Guda, C., Scheeff, E.D., Bourne, P.E., Shindyalov,I.N., 2001. A new algorithm for the alignment of multiple protein structures using Monte Carlo optimization. Proceedings of the Pacific Symposium on Biocomputing 6: 275-286

Gupta, H.K., Atreja, P.P., 1999. Influence of feeding increasing levels of *Leucaena* leaf meal on the performance of milch goats and metabolism of mimosine and 3, 4- DHP. Anim Feed Sci and Technol. 78, 159.

Halkier, B.A., Gershenzon, J., 2006. Biology and biochemistry of glucosinolates. Annu. Rev. Plant Biol. 57, 303–333.

Hammond, A.C., Allison, M.J., Williams, M.J., 1989b. Persistence of DHP-degrading bacteriabetween growing seasons in subtropical Florida. Leucaena Research Report. 10, 66.

Hammond, A.C., Allison, M.J., Williams, M.J., Prine, G.M., Bates, D.B., 1989a. Prevention of Leucaena toxicosis of cattle in Florida by ruminal inoculation with 3, 4-DHPdegrading bacteria. Am J Vet Res. 50, 2176.

Heid, C.A., Stevens, J., Livak, K.J., Williams, P.M., 1996. Real time quantitative PCR. Genome Res. 6, 986-994.

Henrissat, B., 1991. A classification of glycosyl hydrolases based on amino acid sequence similarities. *Biochem J*. 280, 309-316.

Henrissat, B., Davies, G.J., 1997. Structural and sequencebased classification of glycosyl hydrolases. Curr. Opin. Struct. Biol. 7, 637–644.

Higuchi, R., Dollinger, G., Walsh, P.S., Griffith, R., 1992. Simultaneous amplification and detection of specific DNA sequences. Biotechnology. 10, 413–417.

Higuchi, R., Fockler, C., Dollinger, G., Watson, R., 1993. Kinetic PCR analysis: real-time monitoring of DNA amplification reactions. Biotechnology. 11, 1026–1030.

Higuchi, R., Fockler, C., Dollinger, G., Watson, R., 1993. Kinetic PCR analysis: real-time monitoring of DNA amplification reactions. Biotechnology. 11, 1026–1030.

Hughes, M.A., 1993. Molecular genetics of plant cyanogenic b-glucosidases. In A Esen, ed, b-Glucosidases: Biochemistry and Molecular Biology. American Chemical Society, Washington, DC, pp 153–169

Jones, D.A., 1998. Why are so many food plants cyanogenic? Phytochemistry. 47, 155–162.

Jones, P., Vogt, T., 2001. Glycosyltransferases in secondary plant metabolism: tranquilizers and stimulant controllers. Planta. 213, 164–174.

Jones, R.J., 1994. Management of anti-nutritive factors with special reference to Leucaena.In: Gutteridge RC, Shelton HM (Eds). Forage tree legumes in tropical agriculture. CAB International. Wallingford, UK. pp 216

Karban, R., Agrawal, A,A., Mangel, M., 1997. The benefits of induced defenses against herbivores. *Ecology.* 78 (5), 1351–1355.

Kim, Y.S., Uefuji, H., Ogita, S., Sano, H., 2006. Transgenic tobacco plants producing caffeine: a potential new strategy for insect pest control. Transgenic Res. 15, 667–672.

Kim, Y.W., Kang, K.S., Kim, S.Y., Kim, I.S., 2000.Formation of fibrillar multimers of oat β-glucosidase isoenzymes is mediated by the As-Glu1 monomer. J Mol Biol. 303, 831–842.

Kimmel, A.R., Berger, S.L., 1987. Preparation of cDNA and the generation of cDNA libraries: Overview. Meth Enzymol. 152, 307-316.

Kliebenstein, D., Pedersen, D., Barker, B., Mitchell-Olds, T., 2002. Comparative analysis of quantitative trait loci controlling glucosinolates, myrosinase and insect resistance in Arabidopsis thaliana. Genetics. 161, 325–332.

Kristensen, C., Morant, M., Olsen, C.E., Ekstrom, C.T., Galbraith, D.W., Moller, B.L., Bak S., 2005. Metabolic engineering of dhurrin in transgenic Arabidopsis plantswith marginal inadvertent effects on themetabolome and transcriptome. Proc Natl Acad Sci USA. 102, 1779–1784.

Kristoffersen, P., Brzobohaty, B., Hohfeld, I., Bako, L., Melkonian,M., Palme, K., 2000. Developmental regulation of themaize Zm-g60.1 gene encoding a β-glucosidase located to plastids. Planta. 210, 407–415.

Kwok, S., Kellogg, D.E., McKinney, N., Spasic, D., Goda L., 1990. Effects of primer-template mismatches on the polymerase chain reaction: Human immunodeficiency virus type 1 model studies. Nucleic Acids Res. 18, 999-1005.

Lambrix, V., Reichelt, M., Mitchell-Olds, T., Kliebenstein, D.J., Gershenzon, J., 2001. The Arabidopsis epithiospecifier protein promotes the hydrolysis of glucosinolates to nitriles and influences Trichoplusia ni herbivory. Plant Cell. 13, 2793–2807.

Laskowski, R. A., 2001. PDBsum: summaries and analyses of PDB structures. *Nucleic Acids Res.*, **29**, 221-222

Laskowski, R.A., McArthur, M.W.,Moss, D.S.,Thornton, J.M.,1993. PROCHECK: aprogram to check the stereo-chemical quality of protein structures.J Applied Cryst.26,283-291.

Lazzeri, L., Curto, G., Leoni, O., Dallavalle, E., 2004. Effects of glucosinolates and their enzymatic hydrolysis products via myrosinase on the root-knot nematode *Meloidogyne incognita* (Kofoid et White) Chitw. J. Agric. Food Chem. 52, 6703–6707.

Leah, R., Kigel, J., Svedsen, I., Mundy, J., 1995. Biochemical and molecular characterization of a barley seed β-glucosidase. *J Biol Chem*. 270, 15789-15797.

Lee, K.H., Piao, H.L., Kim, H.Y., Choi, S.M., Jiang, F., Hartung, W., Hwang, I., Kwak, J.M., Lee, I.J., Hwang, I., 2006. Activation of glucosidase via stress-induced polymerization rapidly increases active pools of abscisic acid. Cell. 126, 1109–1120.

Lee, L.G., Connell, C.R., Bloch, W., 1993. Allelic discrimination by nick-translation PCR with fluorogenic probes. Nucleic Acids Res. 21, 3761-3766.

Livak, K.J., Flood, S.J., Marmaro, J., Giusti, W., Deetz, K., 1995. Oligonucleotides with fluorescent dyes at opposite ends provide a quenched probe system useful for detecting PCR product and nucleic acid hybridization. PCR Methods Appl. 4, 357-362.

Lodhi, M.A., Guang-Ning,Ye., Norman, F.W., Bruce, I.R., 1994. A simple and efficient method for DNA extraction from grapevine cultivars, *Vitis* species and *Ampelopsis*. Plant Mol Biol Reporter. 12, 6-13.

Maicas, S., Mateo, J.J., 2005.Hydrolysis of terpenyl glycosides in grape juice and other fruit juices: a review. Appl Microbiol Biotechnol. 67, 322–335.

Marti-Renom,M.A., Stuart, A.C., Fiser, A., Sanchez, R., Melo, F., Sali, A., 2000. *Annu Rev Biophys Biomol Struct.* 29, 291–325.

Maruyama, K and S. Sugano, S, 1994. Oligo-capping: a simple method to replace the cap structure of eukaryotic mRNAs with oligoribonucleotides, *Gene***.** 138, pp. 171–174.

McCarter, J.D., Withers, S.G.,1994. Mechanisms of enzymatic glycoside hydrolysis. *Curr. Opin. Struct. Biol.* 4, 885–892.

Meller, M., Vadachkoria, S., Luthy, D.A., Williams, M.A., 2005. Evaluation of house-keeping genes in placental comparative expression studies. Placenta .26, 601–607.

Melony J. S., Tony V., Lisa A. L., Greg J. C.,Bernard M. D., Nigel P. P.*,* 2007. Real-time RT-PCR quantification of Kuruma shrimp transcripts:A comparison of relative and absolute quantification procedures. Journal of Biotechnology. 129. 391–399.

Mewis, I.Z., Ulrich, C., Schnitzler, W.H., 2002. The role of glucosinolates and their hydrolysis products in oviposition and host-plant finding by cabbage webworm, Hellula undalis. Entomol. Exp. Appl. 105, 129–139.

Mizutani, M., Nakanishi, H., Ema, J., Ma, S.J., Noguchi, E., Inohara-Ochiai, M., Fukuchi-Mizutani, M., Nakao, M., Sakata, K., 2002. Cloning of βprimeve-

rosidase from tea leaves, a key enzyme in tea aroma formation. Plant Physiol. 130, 2164–2176.

Møller, B.L., Seigler, D.S., 1999. Biosynthesis of cyanogenic glucosides, cyanolipids and related compounds. In: Singh, B.J. (Ed.), Plant Amino Acids, Biochemistry and Biotechnology. Marcel Dekker Inc., New York, pp. 563–609.

Morant, A.V., Bjarnholt, N., Kragh. M.E., Kjaergaard, C.H., Jørgensen, K., Paquette, S.M., Piotrowski, M., Imberty, A., Olsen, C.E., Møller, B.L., Bak, S.2008. The β-glucosidases responsible for bioactivation of hydroxynitrile glucosides in Lotus japonicus. Plant Physiol.147(3):1072-91.

Morant, A.V., Jørgensen, K., Jørgensen, B., Dam, W., Olsen, C.E., Møller, B.L., Bak, S., 2007 Lessons learned from metabolic engineering of cyanogenic glucosides. Metabolomics. 3, 383–398.

Morant, A.V., Jørgensen, K., Jørgensen, C., Paquette, S.M., Sa´nchez-Pe´rez R., Møller, B.L., Bak, S., 2008. β-Glucosidases as detonators of plant chemical defense. Phytochemistry. 69, 1795–1813.

Morant, M., Bak, S., Møller, B.L., Werck-Reichhart, D., 2003. Plant cytochromes P450:tools for pharmacology, plant protection and phytoremediation. Curr.Opin.Biotechnol. 14, 151–162.

Morrissey, J.P., Osbourn, A.E., 1999. Fungal resistance to plant antibiotics as a mechanism of pathogenesis. Microbiol. Mol. Biol. Rev. 63, 708.

Müller, C., Agerbirk, N., Olsen, C.E., Boeve, J.L., Schaffner, U., Brakefield, P.M., 2001. Sequestration of host plant glucosinolates in the defensive hemolymph of the sawfly Athalia rosae. J. Chem. Ecol. 27, 2505–2516.

Müller, C., Boeve, J.L., Brakefield, P., 2002. Host plant derived feeding deterrence towards ants in the turnip sawfly Athalia rosae. Entomol. Exp. Appl. 104, 153–157.

Mullis, K.B., 1990. The unusual origin of the polymerase chain reaction. Sci Am. 262, 56–61.

Mullis, K.B., Faloona, F.A., 1987. Specific synthesis of DNA in vitro via a polymerase catalysed reaction. Methods Enzymol. 255,335-350.

Murashige, T., Skoog, F., 1962. A revised medium for rapid growth and bioassays with tobacco tissue cultures. Physiol Plant. 15, 473-497.

Napompeth, B., 1994. *Leucaena* psyllid in the Asia-Pacific region: Implications for its management in Africa. FAO, Bangkok,Thailand, pp. 27.

Nelson, L., 2006. Acute cyanide toxicity: mechanisms and manifestations. J. Emerg. Nurs. 32, S8–S11

Nestle, M., 1997. Broccoli sprouts as inducers of carcinogen-detoxifying enzyme systems: clinical, dietary, and policy implications. Proc. Natl. Acad. Sci. USA. 94, 11149–11151..

Niemeyer, H.M., 1988. Hydroxamic acids (4-hydroxy-1,4-Benzoxazin-3-Ones), defense chemicals in the gramineae. Phytochemistry. 27, 3349–3358.

Nisius, A., 1988. The stromacenter in avena plastids – an aggregation of βglucosidase responsible for the activation of oat leaf saponins. Planta. 173, 474–481.

Norton, B.W., Poppi, D.P.,1995. Composition and nutritional attributes of pasture legumes. In: D'Mello JPF, Devendra C (Eds). Tropical legumes in animal nutrition. CAB International. Wallingford. Oxon, UK. pp 23

Olsvik, P.A., Lie, K.K., Jordal, A.O., Nilsen, T.O., Hordvik, I., 2005. Evaluation of potential reference genes in real-time RT-PCR studies of Atlantic salmon. BMC Mol. Biol. 6, 21–29.

Opassiri, R., Hua, Y., Wara-Aswapati, O., Akiyama, T., Svasti, J., Esen, A., Ketudat Cairns, J.R.,2004.β-Glucosidase, exo-α-glucanase and pyridoxine transglucosylase activities of rice BGlu1. *Biochem J*. 379,125-131.

Opassiri,R., Pomthong,B., Onkoksoong,T.,Akiyama,T., Esen,A., and Cairns, J.R.K., 2006.Analysis of rice glycosyl hydrolase family 1 and expression of Os4bglu12 β-glucosidase. *BMC Plant Biology*. 6, 33.

Osbourn, A., 1996. Saponins and plant defence – a soap story. Trends Plant Sci. 1, 4–9.

Osbourn, A.E., 2003. Saponins in cereals. Phytochemistry 62, 1–4.

Papadopoulou, K., Melton, R.E., Leggett, M., Daniels, M.J., Osbourn, A.E., 1999. Compromised disease resistance in saponin-deficient plants. Proc. Natl. Acad. Sci. USA. 96, 12923–12928.

Perez-Novo, C.A., Claeys, C., Speleman, F., Cauwenberge, P.V., Bachert, C., Vandesompele, J., 2005. Impact ofRNAquality on reference gene expression stability. BioTechniques 39, 52–56.

Pfaffl, M.W., Tichopad, A., Prgomet, C., Neuvians, T.P., 2004. Determination of stable housekeeping genes, differentially regulated target genes and sample integrity: BestKeeper Excel-based tool using pair-wise correlations. Biotechnol. Lett. 26, 509–515.

Podstolski,A., Lewak,S., 1970. Specific phloridzin glucosidases from seeds and leaves of apple tree, Phytochemistry. 9, 289–296.

Pottinger, A.J., Hughes, C.E.,1995. A review of wood quality in *Leucaena*. In H.M. Shelton,C.M. Piggin & J.L. Brewbaker (Eds.) Leucaena opportunities and limitations Proceedings of a workshop held in Bogor, Indonesia, January 1994. ACIAR Proceedings Canberra. 57: 98

Poulton, J.E., 1990. Cyanogenesis in Plants. Plant Physiol. 94, 401–405.

Raeymaekers, L., 2000. Basic principles of quantitative PCR. Mol Biotechnol. 15, 115-122.

Ramachandran, G.N., Sasisekhara, V.,1968.Conformation of polypeptides and proteins.Advances in Protein Chemistry.23,283-438.

Rask, L., Andre´ asson, E., Ekbom, B., Eriksson, S., Pontoppidan, B. and Meijer, J. 2000. Myrosinase: gene family evolution and herbivore defense in Brassicaceae. Plant Mol. Biol. 42: 93–113.

Reischl, U., Wittwer, C.T., Cockerill, F., 2002. Rapid Cycle Real-time PCR: Methods and Applications; Microbiology and Food Analysis. New York: Springer-Verlag

Saha, B.K., Tian, B., Bucy, R.P., 2001. Quantitation of HIV-1 by real-time PCR with a unique fluorogenic probe. J Virol Methods. 93, 33-42.

Saiki, R., Scharf, S., Faloona, F., Mullis, K.B., Horn, G.T., Erlich, H.A., Arnheim, N., 1985. Enzymatic amplification of β-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. Science. 230, 1350-1354.

Saiki, R.K., Gelfand, D.H., Stoffel, S., Scharf, S., Higuchi, R., Horn, G.T., Mullis, K.B., Erlich, H.A., 1988. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. Science. 239, 487-491.

Sambrook, J., Fritsch, E.F., Maniatis, T., 1989. Molecular cloning; A Laboratory Manual, 2nd ed.,New York: Cold Spring Harbor Laboratory Press, Cold Spring Harbor.

Scharf, S.J., Horn, G.T., Erlich, H.A., 1986. Direct cloning and sequence analysis of enzymatically amplified genomic sequences. Science. 233, 1076-1078.

Schnee, C., Kollner, T.G., Held, M., Turlings, T.C.J., Gershenzon, J., Degenhardt, J., 2006. The products of a single maize sesquiterpene synthase form a volatile defense signal that attracts natural enemies of maize herbivores. Proc Natl Acad Sci USA. 103, 1129–1134.

Shatsky, M., Nussinov, R., Wolfson, H.J., 2004.A method for simultaneous alignment of multiple protein structures. Proteins: Structure, Function, and Bioinformatics. 1;56(1):143-56.

Shelton, M., 1995. *Leucaena* forage production and quality in South East Queensland. *Leucaena* News. 2: 6 SO-ITFSM- 52. New Orleans: USDA Forest Service, Southern Forest Experiment Station. pp 8

Shindyalov, I.N., Bourne, P.E., 1998. Protein structure alignment by incremental combinatorial extension (CE) of the optimal path protein engineering. 11(9) 739-747.

Sinnott, M.L., 1990. Catalytic mechanisms of enzymic glycosyl transfer. *Chem. Rev.* 90, 1171–1202.

Sue ,M., Ishihara, A., Iwamura, H., 2000a. Purification and characterization of a β-glucosidase from rye (*Secale cereale* L.) seedlings. Plant Sci. 155, 67–74.

Sue, M., Ishihara, A., Iwamura, H., 2000b. Purification and characterization of a hydroxamic acid glucoside β-glucosidase from wheat (Triticum aestivum L.) seedlings. Planta. 210, 432–438.

Suzuki, H., Takahashi, S., Watanabe, R., Fukushima, Y., Fujita, N., Noguchi, A., Yokoyama, R., Nishitani, K., Nishino, T., Nakayama, T., 2006. An isoflavone conjugate-hydrolyzing β-glucosidase from the roots of soybean (Glycine max) seedlings: purification, gene cloning, phylogenetics, and cellular localization. J Biol Chem. 281, 30251–30259.

Svasti, J, Srisomsap, C., Techasakul, S., Surarit, R.,1999. Dalcochinin-8'-O-β D-glucoside and its β-glucosidase enzyme from *Dalbergia cochinchinensis*. *Phytochemistry*. 50, 739-743.

Tan, W., Wang, K., Drake, T.J., 2004. Molecular beacons. Curr Opin Chem Biol. 8, 547-553.

Tangendjaja, B., Willis, R.B.H., 1980. Analysis of mimosine and 3, 4-DHP by High Performance Liquid Chromatography. J Chromatog. 202, 317.

Tattersall, D.B., Bak, S., Jones, P.R., Olsen, C.E., Nielsen, J.K., Hansen, M.L., Hoj, P.B., Moller, B.L., 2001. Resistance to an herbivore through engineered cyanogenic glucoside synthesis. Science. 293, 1826–1828.

Terry, C.F., Shanahan, D.J., Ballam, L.D., Harris, N., McDowell, D.G., Parkes HC., 2002. Real-time detection of genetically modified soya using Lightcycler and ABI 7700 platforms with TaqMan, Scorpion, and SYBR Green I chemistries. J AOAC Int. 85, 938-944.

Todd, J.A., Bell, J.I., McDevitt, H.O., 1987. HLA-DQ bgene contributes to susceptibility and resistance to insulin-dependent diabetes mellitus. Nature. 329, 599-604.

Traw, B.M., Dawson, T.E., 2002. Differential induction of trichomes by three herbivores of black mustard . *Oecologia*. 131 (4), 526–532.

Tricarico, C., Pinzani, P., Bianchi, S., Paglierani, M., Distante, V., Pazzagli, M., Bustin, S., Orlando, C., 2002. Quantitative real-time reverse transcription polymerase chain reaction: normalization to rRNA or single housekeeping genes is inappropriate for human tissue biopsies. Anal. Biochem. 309, 293–300.

Trimbur, D.E., Warren, R.A.J., Withers, S.G., 1992. Regiondirected mutagenesis of residue surrounding the active site nucleophile in _-glucosidase from *Agrobacterium faecalis*. J.Biol. Chem. 267, 10248–10251.

Van der Velden, V.H., Hochhaus, A., Cazzaniga, G., Szczepanski, T., Gabert, J., van Dongen, J.J., 2003. Detection of minimal residual disease in hematologic malignancies by realtime quantitative PCR: principles, approaches, and laboratory aspects. Leukemia. 17, 1013-1034.

Vandesompele, J., De Preter, K., Pattyn, F., Poppe, B., Van Roy, N., De Paepe, A., Speleman, F., 2002. Accurate normalization of real-time quantitative RTPCRdata by geometric averaging of multiple internal control genes. Genome Biol. 3 (7), 1–12.

Vanetten, H.D., Mansfield, J.W., Bailey, J.A., Farmer, E.E., 1994. 2 classes of plant antibiotics – phytoalexins versus phytoanticipins. Plant Cell. 6, 1191– 1192.

Verdoucq, L., Moriniere, J., Bevan, D.R., Esen, A., Vasella, A., Henrissat, B., Czjzek, M., 2004. Structural determinants of substrate specificity in family 1 βglucosidases – novel insights from the crystal structure of sorghum dhurrinase-1, a plant β-glucosidase with strict specificity, in complex with its natural substrate. J. Biol. Chem. 279, 31796–31803.

Vet, J.A., Marras, S.A., 2005. Design and optimization of molecular beacon real-time polymerase chain reaction assays. Methods Mol Biol. 288, 273-290.

Volloch, V, Schweitzer, and S. Rits, S,1994. Ligation-mediated amplification of RNA from murine erythroid cells reveals a novel class of β globin mRNA with an extended 5′- untranslated region. *Nucleic Acids Res.* 22 , pp. 2507–2511.

Wallace, A. C., Laskowski, R. A.,and Thornton, J. M., 1995. LIGPLOT: A program to generate schematic diagrams of protein-ligand interactions. *Prot. Eng.*, **8**, 127-134

Wang, E., Wang, R., DeParasis, J., Loughrin, J.H., Gan, S., Wagner, G.J., 2001. Suppression of a P450 hydroxylase gene in plant trichome glands enhances natural-product-based aphid resistance. Nat Biotechnol. 19, 371–374.

Warzecha, H., Gerasimenko, I., Kutchan, T.M., Stockigt, J., 2000. Molecular cloning and functional bacterial expression of a plant glucosidase specifically invoved in alkaloid biosynthesis. *Phytochemistry.* 54, 657-666.

Weissenberg, M., Levy, A., Svoboda, J.A., Ishaaya, I., 1998. The effect of some Solanum steroidal alkaloids and glycoalkaloids on larvae of the red flour beetle Tribolium castaneum, and the tobacco hornworm, Manduca sexta. Phytochemistry 47, 203–209.

Wheeler, R.A., Brewbaker, J.L., 1990. An evaluation of results from the *Leucaena* psyllid trial network. Leucaena Research Reports. 11: 23–31.

Wink, M, 2004. Phytochemical diversity of secondary metabolites in: Encyclopedia of Plant and Crop Science, Marcel Dekker Inc. 915-919.

Withers, S.G., Warren, R.A.J., Street, I.P., Rupitz, J.B., Kempton, J.B., Aebersold, R., 1990. Unequivocal demonstration of the involvement of a glutamate residue as a nucleophile in the mechanism of a 'retaining' glycosidase. J. Am. Chem. Soc. 112, 5887–5889.

Wittwer, C.T., Herrmann, M.G., Moss, A.A., Rasmussen, R.P., 1997. Continuous fluorescence monitoring of rapid cycle DNA amplification. Biotechniques. 22,130-131, 134-138.

Wittwer, C.T., Ririe, K.M., Andrew, R.V., David, D.A., Gundry, R.A., Balis, U.J., 1997. The LightCycler: a microvolume multisample fluorimeter with rapid temperature control. Biotechniques. 22, 176-81.

Wong, M.L., Medrano, J.F., 2005. Real-time PCR for mRNA quantification. BioTechniques. 39, 75–85.

Woods, R.G., Roper, K.E., Gauthier, M., Bebell, L.M., Sung, K., 2004. Gene expression during early ascidian metamorphosis requires signaling by Hemps, and EGF-like protein. Development 131, 2921–2933.

Xu, Z.W., Escamilla-Trevino, L.L., Zeng, L.H., Lalgondar, M., Bevan, D.R., Winkel, B.S.J., Mohamed, A., Cheng, C.L., Shih, M.C., Poulton, J.E., Esen, A., 2004. Functional genomic analysis of Arabidopsis thaliana glycoside hydrolase family 1. Plant Mol. Biol. 55, 343–367.

Zasada, I.A., Ferris, H., 2004. Nematode suppression with brassicaceous amendments: application based upon glucosinolate profiles. Soil Biol. Biochem. 36, 1017–1024.

Abstracts published:

1. S. Srivastava, M. Arha, S. K. Gupta, **N. M. Shaik**, A. K. Yadav, P. Kulkarni, O. U. Abhilash, B. M. Khan and S. K. Rawal (2004). *Assessment of Leucaena Biodiversity Using Molecular Marker.* In Abstract volume of National Science Day, NCL, Pune, India.

2. M. Arha, S. K. Gupta, **N. M. Shaik,** S. Srivastava, A. K. Yadav, P. Kulkarni, O. U. Abhilash, B. M. Khan and S. K. Rawal (2006). *Metabolic engineering of Leucaena leucocephala for eco-friendly paper and pulp industry.* In Abstract volume of National Science Day, NCL, Pune, India.

3. M. Arha, S. K. Gupta, **N. M. Shaik,** S. Srivastava, A. K. Yadav, P. Kulkarni, O. U.Abhilash, B. M. Khan and S. K. Rawal (2006). *An approach for genetic modification of Leucaena leucocephala for eco-friendly pulp and paper production.* In Proceedings International Symposium on Frontiers of Genetic Engineering and Biotechnology: Retrospect and Prospect, Osmania University, Hyderabad, Andhra Pradesh, India.

4. O. U. Abhilash, M. Arha, S. K. Gupta, **N. M. Shaik**, S. Srivastava, A. K. Yadav, P. Kulkarni, B. M. Khan and S. K. Rawal (2006). *Cell wall bound peroxidase from Leucaena leucocephala- A candidate enzyme involved in lignin polymerisation.* In Proceedings of International Symposium on Frontiers of Genetic Engineering and Biotechnology: Retrospect and Prospect, Osmania University, Hyderabad, Andhra Pradesh, India.

Research papers published/under review/under preparation:

1. V. L. Sirisha, S. Prashant, D. Ranadheer, P. Ramprasad, N. M. Shaik, Manish Arha, S. K. Gupta, **Sameer Srivastava**, A. K. Yadav, P. S. Kulkarni, O. U. Abhilash, B. M. Khan, Shuban K. Rawal and P. B. Kavi Kishor (2008). *Direct shoot organogenesis and plant regeneration from hypocotyl explants in selected genotypes of Leucaena leucocephala—A leguminous pulpwood tree*. Indian Journal of Biotechnology, Vol 7, pp 388-393.

2. **N. M. Shaik**, M. Arha, S. K. Gupta, **S. Srivastava**, A. K. Yadav, P. S. Kulkarni, O. U. Abhilash, B. M. Khan and S. K. Rawal (2008). *High Frequency Regeneration of Leucaena Leucocephala - A Leguminous Pulpwood Tree Species* (communicated).

3. **N. M. Shaik**, S. Srivastava, R. Tatkare, M. Arha, S. K. Gupta, A. K. Yadav, P. S. Kulkarni, O. U. Abhilash, N. Akula, B. M. Khan and S. K. Rawal (2008). *Genetic transformation of Leucaena leucocephala with anti-sense constructs of Lignin biosynthesis pathway genes* (Manuscript under preparation).

4. **N. M. Shaik**, M. Arha, S. K. Gupta, S. Srivastava, A. K. Yadav, P. S. Kulkarni, O. U. Abhilash, B. M. Khan and S. K. Rawal (2008). *Molecular cloning and characterization of a Family 1 Glycosyl hydrolase from Leucaena leucocephala* (Manuscript under preparation).

5. Sameer Srivastava, R. K. Gupta, M. Arha, S. K. Gupta, **N. M. Shaik,** S. Prashant, A. K. Yadav, P. Kulkarni, O. U. Abhilash, R. J. Santosh, R. K. Vishwakarma, V. L. Sirisha, S. K. Rawal, P. B. Kavi Kishor, B. M. Khan (2008). *Molecular cloning, characterization and expression analysis of Cinnamoyl-Co A Reductase (CCR) gene in developing seedlings of Leucaena leucocephala, a paper and pulp yielding tree species* (communicated)

6. M. Arha, S. K. Gupta, **N. M. Shaik,** S. Srivastava, A. K. Yadav, P. Kulkarni, O. U. Abhilash, B. M. Khan and S. K. Rawal (2005). *Isolation, cloning and characterization of Caffeoyl CoA 3-O- methyl transferase (CCoAOMT) from Leucaena leucocephala*. In Proceedings of National Symposium on Plant Biotechnology: New Frontiers, CIMAP, Lucknow, Uttar Pradesh, India.

7. **N. M. Shaik**, M. Arha, S. K. Gupta, , S. Srivastava, A. K. Yadav, P. Kulkarni, O. U. Abhilash, B. M. Khan and S. K. Rawal (2005). *High frequency regeneration and transformation of Leucaena leucocephala*. In Proceedings of National Symposium on Plant Biotechnology: New Frontiers, CIMAP, Lucknow, Uttar Pradesh, India.

In Books:

1. **N. M. Shaik**, M. Arha, S. K. Gupta, S. Srivastava, A. K. Yadav, P. S. Kulkarni, O. U. Abhilash, B. M. Khan and S. K. Rawal (2007). *Multiple shoot regeneration from the cotyledonary nodes of Leucaena leucocephala*. In Recent trends in Horticultural Biotechnolgy, Vol. I, pp 337-340, Eds. Raghunath Keshavachandran *et al.*, New India Publishing Agency, New Delhi.