Molecular Studies of Lignin Metabolism in *Leucaena leucocephala*

A THESIS

SUBMITTED TO THE

UNIVERSITY OF PUNE

FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY
IN
BIOTECHNOLOGY

BY

ARUN KUMAR YADAV

UNDER THE GUIDANCE OF
Dr. B. M. KHAN
HEAD
PLANT TISSUE CULTURE DIVISION

NATIONAL CHEMICAL LABORATORY
PUNE - 411008
INDIA

OCTOBER 2009

Molecular Studies of Lignin Metabolism in *Leucaena leucocephala*

A THESIS

SUBMITTED TO THE

UNIVERSITY OF PUNE

FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY
IN
BIOTECHNOLOGY

BY

ARUN KUMAR YADAV

UNDER THE GUIDANCE OF
Dr. B. M. KHAN
HEAD
PLANT TISSUE CULTURE DIVISION

NATIONAL CHEMICAL LABORATORY
PUNE-411008
INDIA

OCTOBER 2009

Dedicated to my family

CERTIFICATE

This is to certify that the work incorporated in the thesis entitled **"Molecular Studies of Lignin Metabolism in *Leucaena leucocephala*"** submitted by **Arun Kumar Yadav** 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 entitled **"Molecular studies of Lignin Metabolism in *Leucaena leucocephala*"** 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.

October, 2009

Arun Kumar Yadav
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 an exquisite 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 (former HOD, PTC division, NCL, Pune) 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 acknowledge the contributions of Dr. C.G.Suresh, Senior Scientist and his student Mr. Manas Sule, Biochemical Division, NCL, Pune, to carry out the Molecular modeling studies in his lab and his support.

I would like to duly acknowledge Prof. David R. Nelson, Associate professor Dept. of Molecular sciences, University of Tennessee, USA, for authentic categorization of LlCAld5H protein as a CYP84A16.

I would also like to acknowledge the contributions of Mrs. M.V.Shirgurkar and Dr. D.K.Kulkarni Scientist, Plant Tissue Culture Division, NCL, Pune, to carry out the histochemical studies and allowing me to use the microscopy facilities.

I wish to thank Dr. D. C. Agrawal for allowing me to use gel documentation and particle gun facilities.

I duly acknowledge my teachers and HOD, M.S. University of Baroda, Prof. G. Naresh Kumar and my project mentor Dr. (Mrs.) Jayashree pohnerkar and Dr. Shashikant Acharya for constant encouragement and significant suggestions that had been very helpful to carry out this research.

My heartfelt gratitude goes to Dr.Sucheta.N.Karande for making me feel homely in Pune during the course of this work and also providing me a helping hand in tissue culture studies.

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

Special thanks are due to Mr. Rashesh Parikh, Mr., Mr and Mrs. Murawala, Mr. Sachin (from NCCS) for helping me time and again during my experimentation.

Mr. Santosh Jadhav, Mr. Rishi and Mr. Kannan deserves the special thanks for their help and support to carry out, Southern hybridization, ELISA studies and formatting/editing of the thesis respectively.

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

I thank trainees Deepchand, Mahendra, Mrinal, Nishi and Kiran for their help during my experimental work.

I take this opportunity to thank my friends Dr. Bhushan, Dr. Manish Pathak Dr. Nagraj, Bhuban, Shashi, JP, Sagar, Manje, Dr. Elan, Dr. Santosh, Dr. Chetan, Dr. Atul F., Dr. Satya, Ambrish, SuryaPrasad, Sachin, Dr. Kishore bhaiya, Asutosh, Avinash, Parth, Krunal, Shweta, Swapna, and Venkatesh for their help, support, best wishes and lighter moments shared together throughout my work period.

I thank my friends at MSU Vadodara, Keyur, Hemendra, Chirayu, Hemendra, Heena, for their kind cooperation in this work.

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

I owe my deepest gratitude to my grandparents, My Mom and dad, brother (Arjun) and sister (Usha) who have constantly encouraged me to excel. I would like to thank my wife (Bineeta) for her motivation and support during this work.

I would like to heartily acknowledge my little angel (daughter) Ananya, as her smileful welcome whenever I returned home from the lab inculcates brings quanta of energy to carry out further work.

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

Arun Kumar Yadav

Abbreviations

ALA	δ-amino levulenic 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
ELISA	Enzyme linked immuno sorbent assay
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
*K*m	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
δ	Delta
%	Percentage
°C	degree Celsius
µg	microgram
µg/L	Micrograms per liter
µL	microlitre
µm	micrometer
µM	micromolar
4CL	4-Coumarate coenzyme A ligase

Chapter 1 Introduction

1.0 Paper 1
1.1 International Trades 2
1.2 Consumption 2
1.3 Wood 3
1.4 Lignin 6
1.4.1 Evolution of the lignification of xylem cells 9
1.4.2 Developmental control of lignin composition 10
1.5 Pulping process 11
1.5.1 Biological Pulps 12
1.6 Lignification 13
1.6.1 Overview of the monolignol biosynthetic pathway 13
1.6.2 Transport and Storage of monolignols 22
1.6.3 Polymerization 23
1.7 Why *Leucaena*? 24
1.8 *Leucaena* 25
1.8.1 *Leucaena leucocephala* species growth forms 26
1.9 Wood 26
1.10 Why CAld5H from *Leucaena*? 26
1.11 CAld5H is a Cytochrome P450 protein (CYP84A1) 27
1.11.1 Characteristic Structural Features 28

Chapter 2 Material and Methods

2.1 Plant Material 31
2.2 Glassware 32
2.3 Plastic ware 32
2.4 Chemicals 32
2.5 Equipment 33
2.6 Buffers and Solutions 34

2.6.9 Different media used for Studies 40
2.7 Host Cells 47
2.8 Methods 48
2.8.1 Bacterial culture conditions 48
2.8.2 Bacterial transformation 48
2.8.3 Colony Screening by PCR 49
2.8.4 Isolation of Nucleic Acids and Polymerase Chain Reaction (PCR) 50
2.8.4.1 Preparation of competent cells using $CaCl_2$ 50
2.8.4.2 *E. coli* transformation 50
2.8.4.3 Restriction digestion of DNA 52
2.8.4.4 Extraction and purification of DNA from agarose gels 52
2.8.4.5 RNA Extraction 53
2.8.4.6 mRNA Purification 53
2.8.4.7 Spectrophotometric determination of nucleic acid concentration 54
2.8.4.8 First strand cDNA synthesis by Reverse Transcription 54
2.8.4.9 Polymerase Chain Reaction (PCR) 56
2.8.4.10 Rapid amplification of cDNA ends (RACE) 57
2.8.5 Quantitative real time PCR (QRT PCR) 63
2.8.6 Nucleic Acids Hybridization 67
2.8.6.1 Southern Hybridization/ Slot blot 68
2.8.7 Hybridization 69
2.8.8 Expression and Purification of Recombinant Protein 70
2.8.8.3 Enzyme Activity 72
2.8.8.4 Polyacrylamide gel electrophoresis (PAGE) 73
2.8.8.5 Preparation of the Separating Gel 73
2.8.8.6 Preparation of the Stacking Gel 73
2.8.8.7 Preparation of the sample 73
2.8.8.8 Loading and Running the polyacryl-amide Gel 73
2.8.9 Histology and Immunocytolocalization 74
2.8.9.1 Histochemical Staining 75
2.8.9.2 GUS histochemical assay 75

2.8.10 *Agrobacterium* mediated tobacco transformation 76
2.8.11 ELISA (Enzyme-Linked Immunosorbent Assay) 77

Chapter 3 Isolation, cloning and characterization of CAld5H gene from Leucaena leucocephala

3.1 Introduction 78
3.2 Materials and Methods 79
3.3 Results and Discussion 83
3.3.1 Multiple sequence alignment of amino acid sequences of reported CAld5H protein. 83
3.3.2 Multiple sequence alignment of nucleotide sequences of the coniferaldehyde 5-hydroxylase/ F5H/CYP450 from different plants. 84
3.3.3 Genomic DNA extraction and PCR 88
3.3.3.1 RNA isolation, cDNA synthesis and PCR. 89
3.3.3.2 PCR amplification of partial cDNA fragments of *Leucaena* CAld5H. 90
3.3.4 Slot Blot Analysis 92
3.3.4.1 Southern hybridization 93
3.3.5 Rapid Amplification of cDNA Ends (RACE) 94
3.3.5.1 5’ RACE 95
3.3.5.2 3’ RACE 99
3.3.6 Isolation of Full-length cDNA of the CAld5H from *Leucaena leucocephala.* 103
3.3.7 Characterization of LlCAld5H cDNA sequence 104
3.3.7.1 Restriction map of LlCAld5H 106
3.3.7.2 Multiple sequence alignment of nucleotide sequences of selected plant CPP450s. 109
3.3.7.3 Characterization of amino acid sequence 115
3.3.7.4 LlCAld5H belongs to the CYP450s family and classified as CYP84A40 116
3.3.7.5 Analysis of amino acid sequence of the *Leucaena* coniferaldehyde 5-hydroxylase for conserved domain 116

3.3.7.6 Multiple sequence alignment of *Leucaena leucocephala* CAld5H amino acids with plant CYP450s proteins. 117
3.3.7.7 Theoretical molecular weight (MW) and Isoelectric Point of the coniferaldehyde 5-hydroxylase. 120
3.3.7.8 Amino acid composition of the *Leucaena* coniferaldehyde 5-hydroxylase 120
3.3.7.9 Hydropathy index of the LlCAld5H amino acids 121
3.3.8 Phylogenetic relationship of *L. leucocephala* coniferaldehyde 5-hydroxylase protein with plant P450s families 122
3.3.8.1 Phylogenetic relationship of *L. leucocephala* coniferaldehyde 5-hydroxylase gene with plant P450s gene families. 125
3.3.9 Discussions 126
3.3.10 Conclusion 127

Chapter 4A Heterologous expression of CAld5H, its purification and characterization

4A.1. Introduction 128
4A.2 Material and Methods 131
4A.3 Results and Discussion 138
4A.3.1 Construction of LlCAld5H protein expression vector 138
4A.3.2 Heterologous expression of LlCALd5H in *E. coli* BL21 (DE) 140
4A.3.3 Generation of polyclonal antibody and IgG purification 142
4A.3.4 Western blot analysis 143
4A.3.5 Optimizing expression of active LlCAld5H protein 144
4A.3.6 Molecular modeling of LlCAld5H 145
4A.3.7 Discussion 151
4A.4 Conclusion 154

Chapter 4B Spatial and temporal expression studies of CAld5H gene and immunocytolocalization

4B.1 Introduction 155

4B.1.1 Introduction of Real-Time quantitative PCR 156
4B.2 Materials and method 156
4B.3 Results and Discussion 159
4B.3.1 Optimization of molar concentration of primers 159
4B.3.2 Spatio-temporal expression studies of CAld5H gene in *Leucaena* 160
4B.3.3 Relative quantification of CAld5H transcripts of *Leucaena* grown on salt stress 161
4B.3.4 Histology and Immunocytolocalization 163
4B.3.5 Discussion 167
4B.4. Conclusion 169

Chapter 5 Genetic transformation in *Leucaena* and tobacco

5.1. Introduction 170
5.2 An introduction of genetic transformation of grasses and trees 173
5.3 Materials and methods 178
5.3.1 Bacterial strain and plasmid 178
5.3.2 Plant Material 178
5.3.3 A. *tumefaciens* culture conditions, transformation and selection 178
5.3.4 Transformation of *Leucaena leucocephala* 179
5.4 Results and Discussion 182
5.4.1 Cloning and construction of antisense vector 182
5.4.2 Tobacco transformation 186
5.4.3 Analysis of tobacco transformants 188
5.4.3.1 PCR based screening using hygromycin gene specific primers 188
5.4.4 Construction of plant transformation vector containing *LlCAld5H* gene in sense orientation 189
5.4.5 Transformation of *Leucaena* and selection 191
5.4.6 Analysis of *Leucaena* transformants 192
5.4.6.1. PCR based screening using hygromycin resistance gene specific primers 195
5.4.6.2 PCR based screening using *LlCAld5H* gene specific primers 195

5.4.6.3 ELISA 197
5.4.7. Discussions 200
5.5. Conclusion 201
Summary 202
Future prospects 205
References 206
Publication 227
Appendix 229

Chapter 1

General Introduction

1.0 Paper

Over the centuries, paper has been one of the unsung drivers of world development, providing the means for people to record and communicate ideas, news and works of art. Even with the advent of the modern world of plastic bags, electronic communication and paperless offices, paper remains an essential ingredient of homes, factories, offices and schools. Although the price of pulp is highly cyclical, the long term trend is ever upwards as demand continues to grow, cheap sources of raw materials diminish and the capital cost of mills increases. In principle, the process of making paper is simple. Wood chips are converted to pulp which in turn is processed to create paper, cardboard and a range of other products. In reality, the processes are a complex mixture of chemical and physical reactions which take place in both batch and continuous mode. The capital cost of the plant is high, skilled operators are required and the potential for environmental degradation throughout the value chain is huge. According to the Jaakko Pöyry 1999 paper market study, the paper industry is today a 300 million-ton plus industry valued at over US$ one hundred billion per year. Industrial grades accounted for 171 million tons, or 58% of the world paper consumption in 1997. Consumption of graphical papers, such as newsprint, printing and writing papers, totaled 126 million tons in 1997, with uncoated woodfrees, at 44 million tons, the single largest grade group in this area. World demand for paper and paperboard is forecast to grow from the current 300 million tons to over 420 million tons by the year 2010 or an average growth rate of 2.8% per annum. The world's major pulp and paper producers are: North America (USA and Canada), South America (Brazil, Chile and Argentina), Europe and Asia. The combined share of North America and Western Europe of global paper and paperboard production has declined since 1980 from about 67% to 62% while the combined production of Latin America and Asia (excluding Japan) has increased from about 11% to 22%. This trend is expected to continue, with the focus of production growth gradually shifting closer to the regions of faster paper consumption growth.

1.1 International Trades

The economic turmoil in Asia had a strong negative impact on the growth of pulp and paper demand in that region, while the depreciated local currencies have translated into a significant cost advantage for Asian companies vis-à-vis their western competitors. Consequently, export-oriented Asian companies have been able and willing to increase their sales onto the Western markets through aggressive pricing. The leading exporters are:

Table 1. World's leading exporters of forest products, pulp and paper.

Forest Products	Europe, Canada, Asia and USA
Pulp	Canada, USA, Sweden, Brazil, Chile and Finland.
Newsprint	Canada, Sweden, Finland And USA.

1.2 Consumption

Paper remains the dominant and essential vehicle of modern communications. In addition, far from ushering in the paperless office, the advent of computers and other electronic equipment has fueled paper demand. By one estimate, personal computers alone account for 115 billion sheets of paper per year worldwide. Communications, however, makes up less than half of the world's paper use; a bigger share is now taken by the booming packaging industry. In the developing world, paper consumption is growing quite rapidly by more than 7% annually between 1980 and 1994 but average per capita consumption remains low, at about 15 kilograms per year. This is well below the 30 to 40 kilograms of paper per capita per year considered the minimum level to meet basic needs for communication and literacy. In the industrialized countries, consumption is vastly greater over 200 kilograms per capita per year in the United States and around 160 kilograms per capita per year in Western Europe. Unlike consumption trends in other mature commodity sectors, paper consumption shows little sign of decoupling from

economic growth. Paper consumption is projected to grow by about 50% by 2010 from its level in 2000. The biggest increase over 80% is expected to occur in developing countries in Asia, where demand is being driven by rapid growth of both incomes and population. Asia, however, despite having the world's fastest increases in local wood production, is likely to experience shortfalls in the supply of all wood products, but especially pulp and paper, soon after the turn of the century.

India: Consumption of paper in India is set to double from the current 7 million tonnes per annum in the next eight years, according to Indian Paper Manufacturers Association (IPMA). The industry growth, pegged now at 7-8 percent annually, will double, thanks to socio-economic development, increasing literacy rates and increased government spending on education. The growth in paper is inevitable and is likely to contribute significantly to the government's target of achieving an overall growth of 12 percent in manufacturing. Unfortunately, the rising cost of raw materials, primarily wood and wood products, is a big challenge facing the industry.

1.3 Wood

Wood characteristics vary in different types of plant. For instance, conifers (gymnosperms) produce softwood whereas angiosperms produce hardwoods. Softwoods are mainly composed of three cell types, tracheids (which play a role both in rigidity and conduction), and axial and ray parenchyma cells. Hardwoods are mainly made of fibers, vessels, and axial and ray parenchyma cells. Vessels transport water and solutes through the vascular system while fibers provide rigidity, and ray cells facilitate centripetal nutrition (Higuchi, 1997). Tracheids, vessels and fibers vary in shape and size (Table 2). The dimensions and chemical composition of the different cell types of wood depend on genetic, developmental and environmental factors (Vallette and de Choudens, 1992).

Table 2 Dimensions of the different cell types of Softwoods and Hardwoods (Fengel and Wegener, 1984).

Cell Type	Character	Cell Dimensions
Soft wood (***Picea abies***)		
Tracheids	Diameter	20 – 40 µm
	Wall thickness	2.1 – 4.3 µm
	Length	1.7 – 3.7 mm
Hard wood (***Fagus sp.***)		
Fibers	Diameter	15 – 20 µm
	Wall thickness	5.0 µm
	Length	0.6 – 1.3 mm
Vessels	Diameter	5.0 – 100 µm
	Wall thickness	1.0 µm
	Length	0.3 – 0.7 mm

Softwoods and hardwoods differ in their pulping characteristics. Indeed, the individual cell types within wood differ in their chemical characteristics, reflecting the underlying differences in biochemistry and molecular biology that are only beginning to be appreciated. The three major components of wood cell walls are cellulose, hemicellulose and lignin. Long molecules of cellulose provide the skeleton of the walls. Linear cellulose chains are aligned together in structures known as 'elementary fibrils' or 'protofibrils' that, in turn, associate into more complex structures called microfibrils. Microfibrils are highly organized and form distinct fibrillar cell wall layers (Delmer and Amor, 1995). Hemicelluloses and other carbohydrates provide the matrix of the cell wall. Lignin, a heterogeneous hydrophobic phenolic polymer, encrusts the other wall components to waterproof and strengthen the wall. In a transverse plane, the parietal structure of wood cells is made of a primary and a secondary wall, the latter consisting of two or three layers, designated S1, S2, and S3 (Fig.1.1). The various cell wall layers differ in chemical composition (Mellerowicz *et al*., 2001). Lignin deposition is one of the final stages of xylem cell differentiation and mainly takes place during secondary thickening of the cell wall (Donaldson, 2001). Lignin deposition proceeds in different phases, each preceded by the deposition of carbohydrates, and starts at the cell corners in the region of the middle lamella and the primary wall when S1 formation has initiated. When the formation of the polysaccharide matrix of the S2 layer is completed, lignification proceeds through the secondary wall. The bulk of lignin is deposited after

cellulose and hemicellulose deposition in the S3 layer. Generally, lignin concentration is higher in the middle lamella and cell corners than in the S2 secondary wall (Baucher *et al*., 1998).The three monolignols (S, G and H) are incorporated at different stages of cell wall formation. Typically, H units are deposited first, followed by G units, and S units still later in angiosperms (Donaldson, 2001). Lignin in vessels is generally enriched in G units, whereas lignin in fibers is typically enriched in S units (Saka and Goring, 1985). A large proportion of S unit is also found in secondary walls of ray parenchyma (Fergus and Goring, 1970).

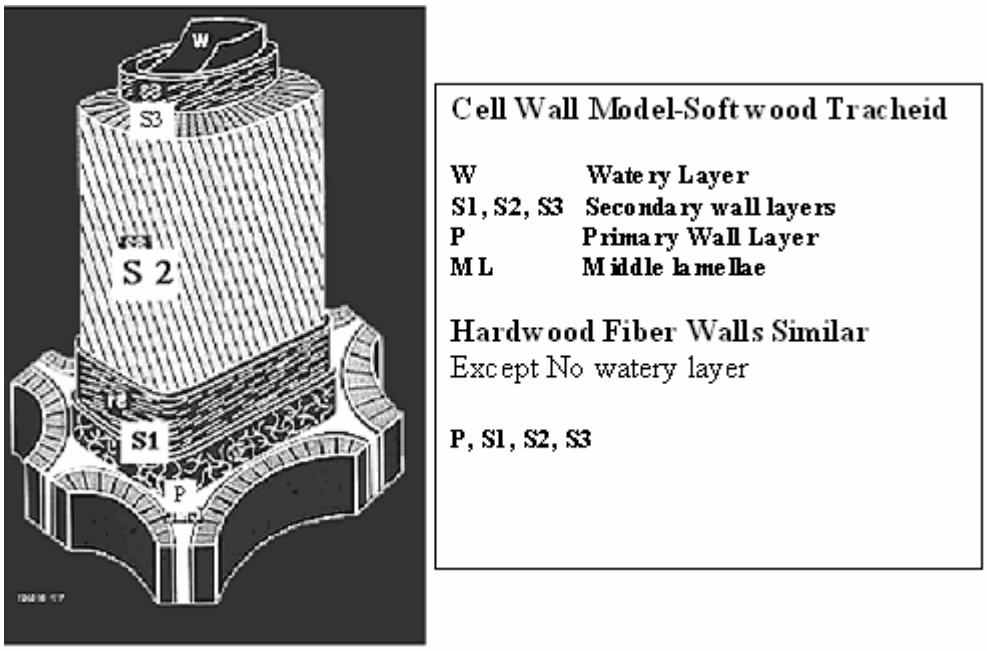


Figure c 1.1. Plant cell wall model

In gymnosperms, the lignin deposited in compression wood is enriched in H units (Timell, 1986). The difference in timing of monolignol deposition is associated with variations in lignin condensation in the individual cell wall layers, as shown by immunocytochemistry with antibodies raised against pure H, pure G, or mixed GS synthetic lignins (Chabannes *et al*., 2001; Joseleau and Ruel, 1997). Lignin deposition and the relative incorporation of the different monolignols into the polymer are spatially and temporally regulated. The mechanisms controlling this process are not yet fully resolved but are likely governed by the interplay between the spatio-temporal expression

of monolignol biosynthetic genes, the kinetics of monolignol delivery to the cell wall and the chemistry of monolignol coupling to the growing polymer in the complex macromolecular environment of the cell wall.

1.4 Lignin

The word lignin is derived from a Latin word *lignum* that means wood. Lignin is a complex three-dimensional, branched heteropolymer of bonded hydroxycinnamyl alcohols that crosslink carbohydrate polymers together, thereby rigidifying plant cell walls (Boerjan W *et al*; 2003, Anterola AM, 2002). The importance of lignin ranges from its fundamental roles in the evolution of land plants, global carbon cycling, and plant growth and development, and its role in the biotic and abiotic stress resistance of plants, to the practical importance of lignin in agriculture and the utilization of plant materials. The adaptive significance of the lignification of xylem cells during the evolution of land plants (Raven JA; 1977) is evidenced by the proliferation of land plants, such that the mass of lignin in the biosphere is second only to the mass of cellulose. It forms an integral cell wall component of all vascular plants, representing on an average of 25% of the terrestrial plant biomass. Lignin content is higher in softwoods (27-33%) than in hardwood (18-25%) and grasses (17-24%). The highest amount of lignin (35–40 %) occurs in compression wood on the lower part of branches and leaning stems (Fengel and Wegner, 1984; Sarkanen *et al.,* 1971). Lignin does not occur in algae, lichens or mosses (Nimz and Tutschek, 1977), whereas the "lignins" of bark differ in their structure from typical wood lignin (Zimmermann *et al*., 1985). Lignins are complex racemic aromatic heteropolymers synthesized from the dehydrogenative polymerization of monolignols, namely coumaryl, coniferyl and sinapyl alcohol monomers differing in their degree of methoxylation (Freudenberg and Neish, 1968) (Fig.1.2). These monolignols produce, respectively, *p*-hydroxyphenyl (**H)**, guaiacyl (**G)**, and syringyl (**S)** phenylpropanoid units when incorporated into the lignin polymer. The amount and composition of lignins vary among taxa, cell types and individual cell wall layers, and are influenced by developmental and environmental cues (Campbell and Sederoff, 1996). Angiosperm

(hardwood) lignins consist of **G** and **S** units and traces of **H** units, whereas gymnosperm (softwood) lignins are composed mostly of **G** units with low levels of **H** units.

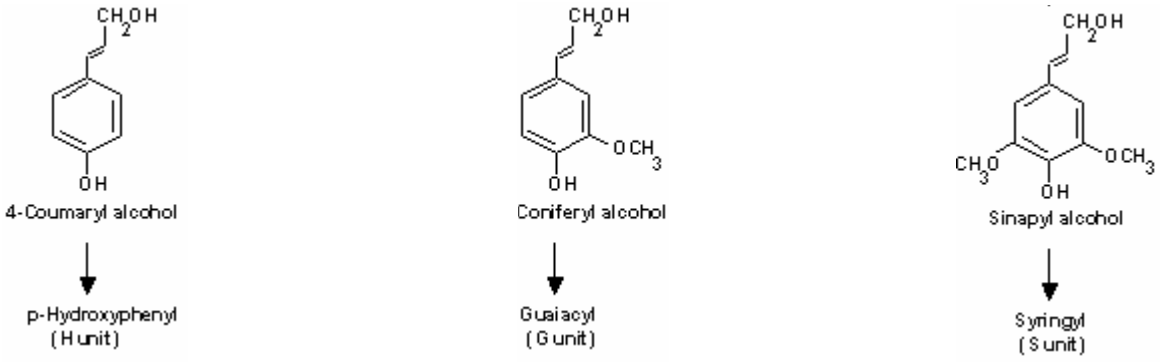


Figure 1.2. Lignin monomers

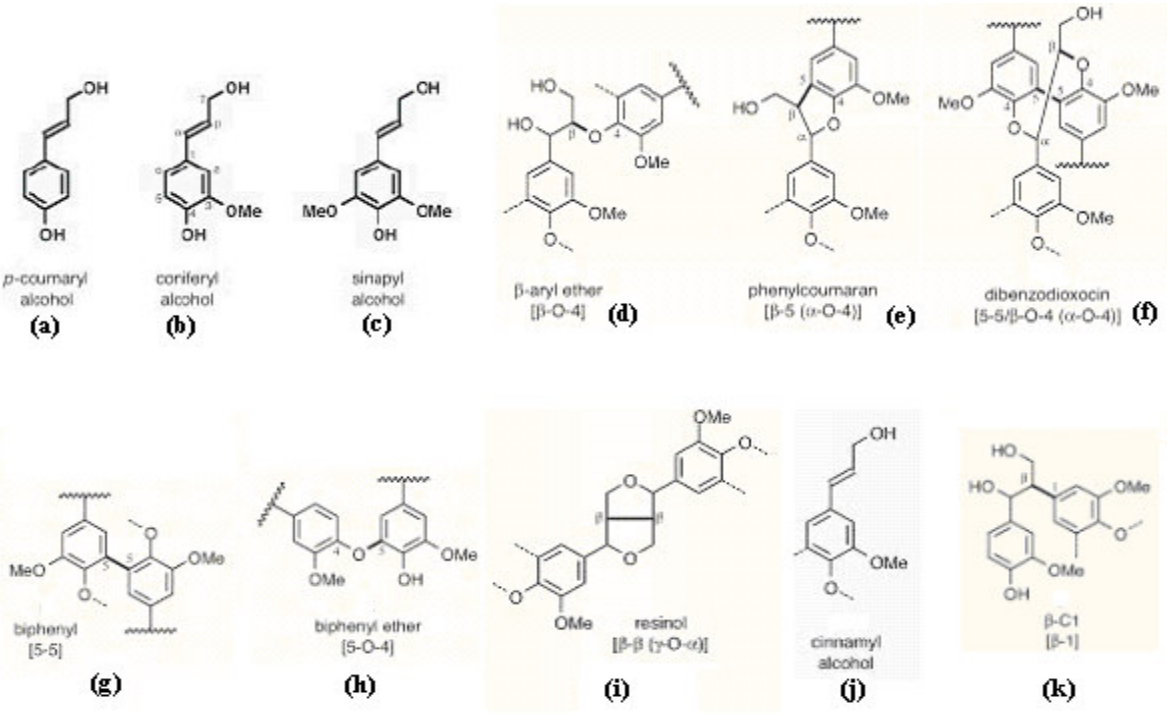


Figure 1.3. Monolignols *p*-coumaryl (a), coniferyl (b) and sinapyl alcohol (c). Different structures formed due to "end wise" coupling (d-f). Coupling between preformed lignin

oligomers results in units linked 5–5 and 5–O–4 (g, h). Coupling of two monolignols form resinol β–β units (i) or cinnamyl alcohol (j) end groups (Boerjan *et al*., 2003).

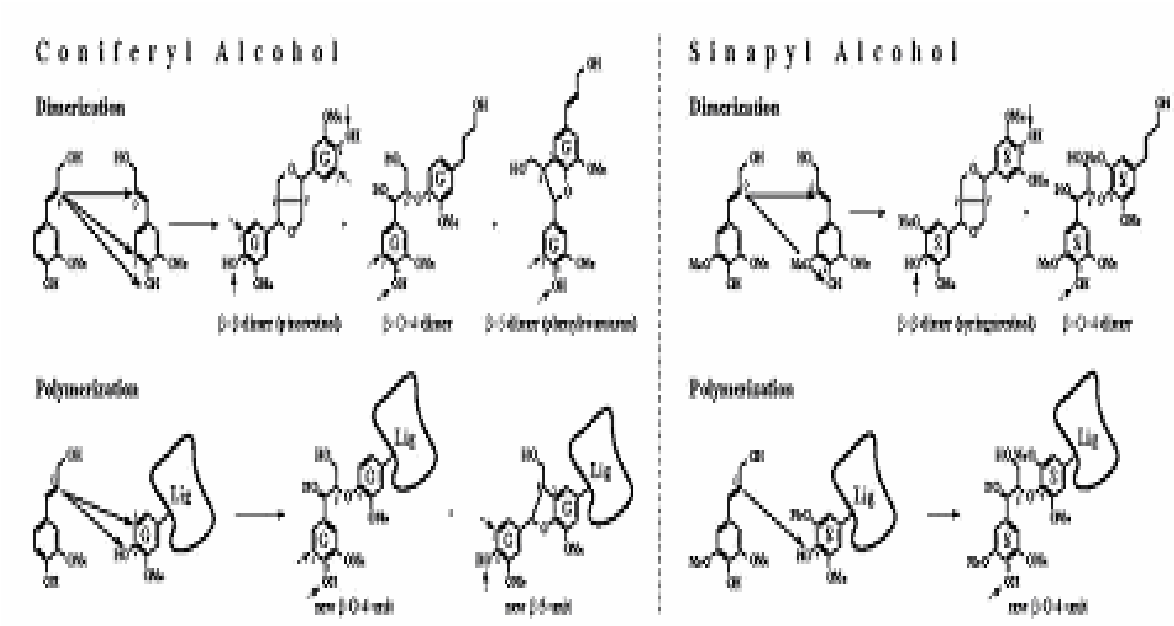


Figure 1.4. Radical coupling of monolignols and differences between dimerization and lignin polymerization. During lignification, dimerization reactions are rare and the major linkages come from coupling of a monolignol radical with a radical of the growing lignin oligomer/polymer. (Boerjan *et al*., 2003.

Lignification is the process by which H, G and S units are linked together via radical coupling reactions (Sarkanen and Ludwig, 1971; Freudenberg and Neish, 1968). The main "end-wise" reaction couples a new monomer (usually a monolignol and usually at its β position) to the growing polymer, giving rise to different structures (Fig.1.3d-f) all of which are β-linked. Coupling between preformed lignin oligomers result in units linked 5–5 and 5– O–4 (Fig.1.3g, h). The coupling of two monolignols is a minor event, with resinol (β–β) units (Fig.1.3 i) or cinnamyl alcohol end groups (Fig.1.3 j) as the

outcome. Monolignol dimerization and lignin polymerization are substantially different processes (Adler, 1997), explaining why lignification produces frequencies of the various units that are different from those produced by dimerization or bulk polymerization in vitro (Fig.1.4).

1.4.1 Evolution of the lignification of xylem cells

The ability to synthesize lignin has been essential in the evolutionary adaptation of plants from an aquatic environment to land and provides crucial structural integrity to the cell wall and stiffness and strength of the stem (Chabannes *et al.,* 2001). The function of lignin in conductive xylem cells is twofold; the hydrophobic nature of lignin makes the secondary walls impermeable to water, thereby facilitating hydraulic conductance; and the lignin crosslink rigidify the thickened wall, allowing it to resist collapse from the high negative hydrostatic pressures in xylem (Sperry J., 2003). The cross linking of cellulose in the xylem secondary wall is also crucial for the mechanical support of the whole plant (Sperry J., 2003, Raven JA., 1977). In gymnosperms, xylem tracheids function in both mechanical support and water transport (Esau K, 1977). Tracheid walls are rich in guaiacyl lignin (G lignin) and do not contain syringyl (S) lignin, apparently because they lack the enzymes for sinapyl alcohol synthesis (Humphreys *et al*; 1999). By contrast, the secondary xylem of angiosperm trees contains two more specialized cell types: the vessel elements, which conduct water, and the fibre cells, which provide mechanical support (Esau K, 1977). Separation of these functions provides a more efficient and economical architecture for the secondary xylem. This efficiency is evident in the fewer but larger vessel elements and the reduced lignin in the woody stems of angiosperm (20% of dry matter) compared with gymnosperm (30%) trees. A lower lignin and higher carbohydrate content requires significantly less energy and carbon for growth (Amthor JS, 2003). Interestingly, in angiosperms, the water-conductive xylem vessel elements of secondary xylem and the primary xylem cells are rich in or contain only G lignin, like the tracheids of the more ancient gymnosperms, whereas the nonconductive xylem fiber cells are rich in S lignin. The fact that water-conducting cells in both gymnosperms and angiosperms

are principally comprised of G lignin suggests a strong selective pressure to conserve the pathway for and the regulation of G lignin biosynthesis in the water conducting cells of xylem during land plant evolution. The discovery and analysis of the three genes that are specifically involved in S lignin synthesis,

those encoding ferulate 5- hydroxylase /coniferaldehyde 5- hydroxylase (F5H/CAld5H), caffeic acid/5-hydroferulic acid O-methyltransferase (COMT) and sinapyl alcohol dehydrogenase (SAD), shows that these genes evolved more recently than the G lignin genes (Humphreys *et al*., 1999; Li *et al*., 2003).

1.4.2 Developmental control of lignin composition

In hardwoods, monolignols are first incorporated into the middle lamella of the primary cell wall during the S_1 and S_2 developmental stages of the secondary wall, whereas incorporation into the secondary wall takes place mainly after the S_3 stage (Terashima *et al*., 1993). Lignification of the vessel wall, cell corner and middle lamella starts in the early stages of xylem differentiation, whereas lignification of fiber secondary walls proceeds most actively in the middle and late stages.

High-resolution microautoradiography of developing hardwood xylem has revealed that the three monolignol units are incorporated at different stages of cell wall formation. H (*p*-hydroxyphenyl) units are incorporated in the earliest stages mainly in the cell corner and middle lamella during the period of formation of the S_1 layer. The deposition of G lignin continues through the early to late stages. S lignin is deposited mainly in the middle and late developmental stages, mostly in the secondary fiber wall, when formation of the inner layer of the secondary wall has already commenced (Terashima and Fukushima; 1993). A somewhat similar situation appears to occur during development of vascular tissues of monocots (Lam and He; 1991).

The fact that different monolignols are incorporated into the cell wall at different stages of vascular differentiation, and in different regions of the wall, strongly argues for tightly regulated pathways to G and S lignins that can function either independently or together. Independent pathways, rather than a metabolic grid, provide the greatest potential for such tight developmental control.

1.5 Pulping process

Lignins as they occur in nature (protolignins) have been grouped into several types, characteristic of hardwoods, softwoods and grasses. Within each type there is a lot of variation: lignins differ from species to species and from one tissue to the next in the same plant even within different parts of the same cell. The process of removing them from the plant changes their form and chemical makeup to a greater or lesser extent, which makes then hard to study and way account for the large and growing number of analytical techniques in use. No one method is ideal for all cases, and the limitations of each method have to be borne in mind when results are interpreted. Two major categories of processes exist for the production of paper pulp: chemical and mechanical. The chemical process uses chemicals to remove lignin from fiber cell walls to obtain long and flexible fibers that consist of polysaccharides only (Gierer, 1985), whereas mechanical pulping process focuses on the mechanical separation of fibers without the removal of lignin. Mechanical pulping processes may be based on any one of the following processes: stone ground wood (SGW) process, the pressure ground wood (PGW) process, the refiner mechanical pulp (RMP) procedure, the thermo-mechanical process (TMP) process, the chemimechanical (CMP) and the chemi-thermomechanical (CTMP) processes, and bleaching of pulp. Though the pulp yield is highest in mechanical pulping, the pulp has limited bleachability and the paper becomes yellow as it ages. In contrast, chemical pulping yields individual intact fibers free from lignin that can interact with other fibers via hydrogen bonds, making very strong paper. Nowadays, the Kraft process is the most widely used chemical procedure for the production of paper. However, this process is gradually being replaced by thermomechanical (TMP) and

chemithermomechanical (CTMP) pulping methods that give higher pulp yields and consume less water. In 2000, the world wood pulp production has been estimated at 1, 35, 852 thousands of metric tons, from which 6.3% are of mechanical pulp, 15.6% TMP pulp, 3.6% semi-mechanical pulp (includes mechanical pulps mildly treated with sodium sulfite (Na_2SO_3) and sodium carbonate (Na_2CO_3) to partially remove lignin and hemicelluloses and followed by bleaching), and 74.5% chemical pulp (Food and Agriculture Organization, 2001).

1.5.1 Biological Pulps

Biopulping is achieved by treatment of wood chips with lignin degrading fungi prior to pulping. Such a pretreatment reduces energy requirements and improves the paper strength. In addition to the economical benefits, the bio-pulping process is not harmful to the environment, because only benign materials are used and no additional waste streams are generated (Reid, 1991; Akhtar *et al*., 1998a; Highley and Dashek, 1998). Biomechanical pulping is on the way towards being commercialized (Akhtar *et al*., 1998b; Breen and Singleton, 1999). A fungal pretreatment in chemical pulping has also been shown to be beneficial for subsequent pulping processes because part of the lignin is removed or modified (Akhtar *et al*., 1998a).

Pulp intended for use in fine papers, on the other hand, is prepared by cooking to remove the lignin, as well as to separate the fibers. The process ordinarily has to be stopped before all of the lignin is removed, however, because the lignin gets more resistant to removal as pulping proceeds, and the cellulose becomes more vulnerable to the chemicals used. If pulping proceeded until all the lignin was gone, most of the cellulose would be gone too. Most of the remaining lignin is removed by bleaching. Because of the flap about dioxin, mills are using as little chlorine in bleaching as possible, and some are using other bleaching chemicals instead. Some of the residual lignin that used to be removed by bleaching is now being removed in some pulp mills by extended delignification, which Smook defines as “Kraft pulping” modification within a

continuous digester which utilizes a two stage, partially counter-current cooking sequence to achieve lower kappa number and improved overall washing efficiency. This can bring the percent of lignin down to about 1% on the weight of the paper. Bleaching is still necessary, whether or not extended delignification has been used. Products from Kraft pulping are loosely referred to as **"kraft lignins"** and products resulting from chlorine bleaching as "chlorolignins," but we know much less about them than we do about model compounds and lignin preparations. Lignosulfonates, byproducts of sulfite pulping, which can also be produced from Kraft black liquor, have commercial value because they can be used for so many different purposes such as the additive to concrete mixtures, textile dyes, brick clay and animal feed.

1.6 Lignification

Lignification process occurs in three steps: Monolignol biosynthesis, Transport and Storage of monolignols and Polymerization of monolignols.

1.6.1 Overview of the monolignol biosynthetic pathway

The general outline of the pathway that gives rise to coniferyl alcohol and other monolignols was proposed ~45 years ago (Higuchi and Brown, 1963; Freudenberg, 1965; Freudenberg and Neish, 1968). Doubts still remain about the degree to which the pathway follows the same sequence in different plant species or even different cell types within a species. Lignin monomers, or monolignols, are produced intracellularly, then exported to the cell wall, and subsequently polymerized. The monolignols are products of the phenylpropanoid pathway, starting from phenylalanine which in turn is produced by shikimate pathway (Fig.1.5) along with the tyrosine and tryptophan. Phosphoenol-pyruvate (PEP) and erythrose 4-phosphate are the precursors of shikimate pathway. In this pathway photosynthetically fixed carbohydrates in plants are being channelised into formation of aromatic amino acids, which act as precursors of various secondary metabolites and some of the plant hormone like auxin as well. Coniferyl alcohol and other monolignols are derived from phenylalanine in a multistep process (Fig.1.6).

Several other major classes of plant products in addition to lignin are derived from phenylalanine, including flavonoids, coumarins, stilbenes, and benzoic acid derivatives (Dixon and Paiva, 1995, Holton and Cornish, 1995). The initial steps in the biosynthesis of all these compounds are shared through the general phenylpropanoid pathway. Phenylpropanoid compounds are so named because of the basic structure of a three-carbon side chain on an aromatic ring, which is derived from L-phenylalanine.

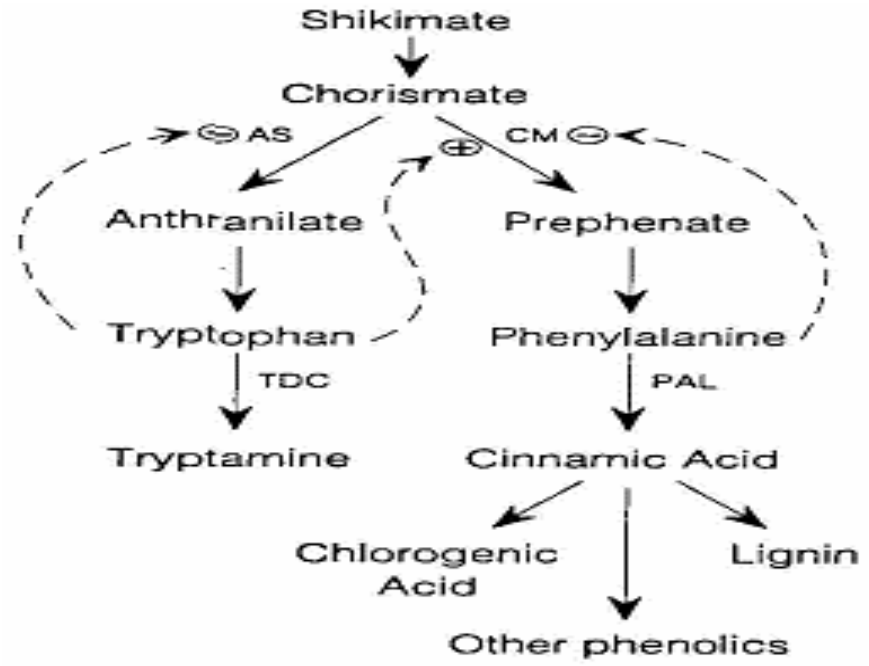


Figure 1.5. Schematic diagram of the Shikimate and Phenylpropanoid pathway. Dashed line indicate feedback inhibition (-) or activation (+). AS: Anthranilate synthase, CM : Chorismate mutase, TDC:Tryptophan decarboxylase and PAL: Phenylalanin- ammonia lyase.

1.6.1.1 Phenylalanine Ammonia-Lyase (PAL; EC 4.3.1.5)

Phenylalanine is the precursor molecule for synthesis of lignin along with several other secondary metabolites. Deamination of phenylalanine to cinnamate is catalyzed by the enzyme phenylalanine ammonia-lyase. The amino group cleaved from phenylalanine by

PAL is released as ammonia. No data are available regarding the fate of this released ammonia, but the potential magnitude of nitrogen loss through this reaction indicates that the ammonia is probably recaptured within the plant. One possible route for recapture of the ammonia is through the action of glutamine synthetase (Lam *et al*., 1995). This enzyme is one of the most intensively studied in plant secondary metabolism because of the key role it plays in phenylpropanoid biosynthesis.

PAL is found as a tetramer in vascular plants (Hanson and Havir, 1981; Jones, 1984). Functional PAL enzyme has been expressed from a parsley cDNA in *E. col.* (Schulz *et al*., 1989). PAL subunits are typically encoded by multigene families in angiosperms; with 2 to 40 different members depending on the species (Wanner *et al*., 1995). Genes encoding different PAL subunits show tissue-specific patterns of expression in severa1 angiosperms (Bevan *et al*., 1989; Liang *et al*., 1989; Lois and Hahlbrock, 1992). PAL has been suppressed by 85% and >98% in the stems of transgenic plants with resultant 52% (Sewalt *et al*., 1997) and 70% (Korth *et al*., 2001) reduction in Klason lignin content, respectively Lignin monomeric composition, determined by pyrolysis GC-MS, was characterized by a lower proportion of G units and a 1.7-fold increase in S/G ratio (Sewalt *et al*., 1997). Similarly, using an independent method, (Korth *et al*. 2001) determined a 4-fold increase in the S/G ratio caused by a more pronounced reduction in the level of G units than in S units. Because PAL catalyzes the first step of the phenylpropanoid pathway, reduction of its activity results in a wide range of abnormal phenotypes. The transgenic plants were stunted, had curled leaves, and had thinner cell walls in the secondary xylem with less lignin than those of the control (Elkind *et al*., 1990; Bate *et al*., 1994). These plants were also more susceptible to the fungal pathogen *Cercospora nicotianae* (Maher *et al*., 1994).

1.6.1.2 Cinnamate- 4-Hydroxylase (C4H; EC 1.14.13.11)

Cinnamate 4-hydroxylase belongs to a cytochrome P-450-linked monoxygenases family and catalyses the hydroxylation of *p*-cinnamic acid to *p*-coumarate. Molecular oxygen is

cleaved during this reaction, with one oxygen atom added to the aromatic ring and the other reduced to water. C4H has been purified and characterized to different degrees from severa1 plant species. Analysis of cDNA clones shows that C4H represents a class of cytochrome P450s distinct from previously characterized P450 enzymes (Fahrendorf and Dixon, 1993; Mizutani *et al*., 1993; Teutsch *et al*., 1993). C4H cDNA have been expressed in yeast, and active enzyme has been recovered (Fahrendorf and Dixon, 1993; Pierrel *et al*., 1994).

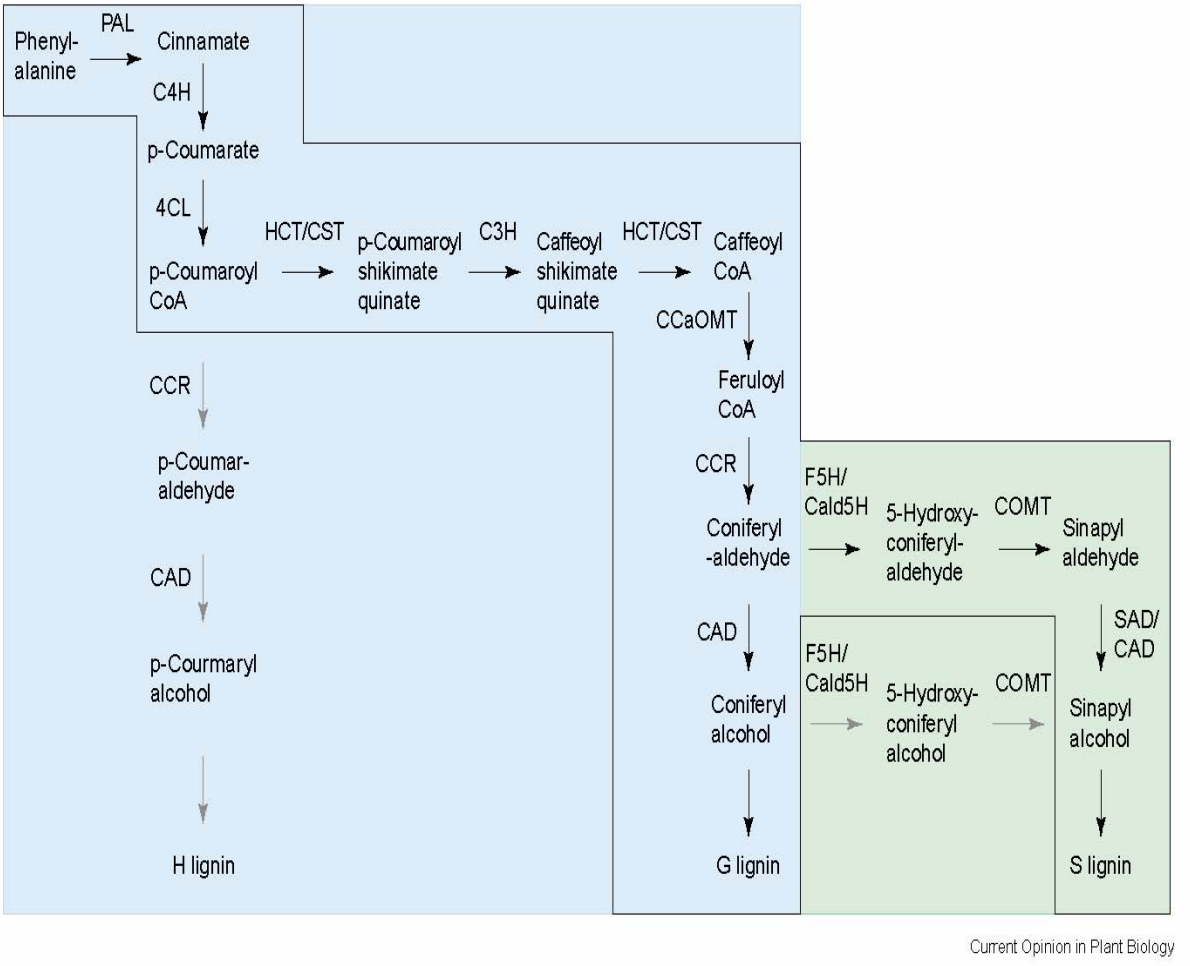


Figure 1.6. The predominant pathway for monolignol biosynthesis in xylem cells is outlined in black, with the dark arrows showing the primary substrates and products and the gray arrows showing the minor substrates and products. The blue shading indicates the pathway that is conserved between angiosperms and gymnosperms, whereas the green shading indicates the angiosperm-specific pathway. The enzymes and their abbreviations

are as follows: CAD, (hydroxy)cinnamyl alcohol dehydrogenase; CCoAOMT, caffeoyl CoA O-methyltransferase; CCR, (hydroxy)cinnamoyl CoA reductase; C3H, p-coumaroyl shikimate/quinate 3-hydroxylase; C4H, cinnamate 4-hydroxylase; 4CL, 4-coumarate CoA ligase; COMT, caffeic acid/5-hydryferulic acid O-methyltransferase; F5H/CAld5H, ferulate 5-hydroxylase/coniferylaldehyde 5-hydroxylase; PAL, phenylalanine ammonia-lyase; SAD, sinapyl alcohol dehydrogenase.

The plant enzyme is able to couple effectively with yeast NADPH- cytochrome P450 reductase and catalyze the hydroxylation of cinnamate in microsomes from transformed yeast with high efficiency and specificity. Feeding experiments using phenylalanine and cinnamic acid with different radiolabels have shown that C4H utilizes cinnamic acid produced *in situ* by PAL more readily than exogenously supplied cinnamic acid (Hrazdina and Wagner, 1985). These findings suggest that cinnamate is preferentially transferred from PAL to C4H, rather than coming to equilibrium with the cytosol and diffusing to C4H. This preferentia1 transfer of intermediates between enzymes of a pathway is an example of metabolic channeling (Hrazdina and Jensen, 1992). Channeling offers a mechanism by which the cell can maintain very low concentrations of toxic compounds while still allowing high levels of metabolic flux through the pathway. The degree to which channeling occurs in the remainder of the monolignol biosynthetic pathway is as yet unknown, but there have long been hypotheses that channeling might be important in regulation of monolignol biosynthesis (Hrazdina and Jensen, 1992). Formation of multienzyme complexes is one mechanism by which channeling may be achieved, and co-localization of all the enzymes of the pathway to the same subcellular compartment is another.

1.6.1.3 *p*-Coumarate 3-Hydroxylase (C3H; EC 1.10.3.1),

Little is known about *p*-coumarate 3-hydroxylase, the enzyme that catalyzes the hydroxylation of *p*-coumarate to form caffeate. The gene encoding *p*-coumarate 3-

hydroxylase (C3H) has only recently been cloned by two independent research groups. Using a functional genomics approach, Schoch *et al*., identified a CYP98A3 as possible candidate for C3H. CYP98A3 was highly expressed in developing xylem as determined by immunolocalization and had high activity when using 5-O-(4-coumaroyl) D-quinate and 5-O-(4-coumaroyl) shikimate as substrates and low activity when using *p*-coumaroyl-CoA. In parallel, by screening *Arabidopsis* mutants under UV light, (Franke *et al*. 2002a) isolated the reduced epidermal fluorescence 8 (*ref8*) mutant. By positional cloning, the *REF8* gene was identified as the cytochrome P450 dependent monooxygenase *CYP98A3*. The *ref8* mutant had collapsed xylem vessels, a higher cell wall degradability, and a higher susceptibility to fungal colonization (Franke *et al*., 2002b), associated with the accumulation of *p*-coumarate esters instead of sinapoylmalate and with a reduction in lignin content of 60-80%.

1.6.1.4 4-Coumarate: Coenzyme A Ligase (4CL; EC 6.2.1.12)

4-Coumarate: CoA ligase catalyzes the formation of CoA thioester of cinnamic acids in the biosynthesis of a wide variety of phenolic derivatives, including benzoic acid, condensed tannins, flavonoids, and the cinnamyl alcohols (Gross, 1985). 4CL depends strictly on ATP, and the reaction resembles the activation of fatty acids, proceeding through an intermediate acyl adenylate, which reacts with coenzyme A to form the thioester. Early studies on 4CL suggested that multiple isoforms might control formation of different phenylpropanoid products. Grand *et al*. (1983) identified three isoforms of 4CL with different substrate specificities in stems of poplar and hypothesized that expression of different 4CL enzymes could regulate the relative abundance of the different precursors for lignin. Transgenic plants with reduced 4CL activity have been produced in tobacco (Kajita *et al*., 1996, 1997), *Arabidopsis* (Lee *et al*., 1997), and aspen (Hu *et al*., 1999; Li *et al*., 2003). In tobacco, reduction of 4CL by over 90% resulted in 25% less lignin. In poplar and *Arabidopsis* with a >90% reduced 4CL activity, lignin content was reduced by 45–50%. A discrepancy between the results published by Kajita *et al*. (1997) and (Hu *et al*. 1999) is that the transgenic tobacco lines with the most severe

reduction in lignin content (25%) were characterized by a collapse of vessel cell walls and reduced growth (Kajita *et al*., 1997), whereas the transgenic poplars with a 45% reduction in lignin content had a normal cell morphology and a higher growth rate than the control (Hu *et al*., 1999). However, the increased growth was probably due to pleiotropic effects caused by the constitutive down-regulation of 4CL governed by the CaMV35S promoter. A combinatorial down-regulation of 4CL along with an over expression of F5H in xylem has been achieved by co-transformation of two *Agrobacterium* strains in aspen (Li *et al*., 2003). Additive effects of independent transformation were observed, in particular a 52% reduction in lignin content associated with a proportional increase in cellulose and a higher S/G ratio. The results show that stacking transgenes allows several beneficial traits to be improved in a single transformation step (Halpin & Boerjan, 2003).

1.6.1.5 Caffeate O-Methyltransferase (COMT; EC 2.1.1.68)

Caffeic acid is methylated to form ferulic acid by caffeic acid -O-methyltransferase, using S-adenosyl methionine (SAM) as the methyl group donor. This methylation reaction limits the reactivity of the 3-hydroxy group, thus reducing the number of sites on the aromatic ring that can form bonds to other monolignol molecules during polymerization. The same enzyme is also believed to catalyze the methylation of 5-hydroxyferulate to sinapate. COMT is clearly implicated in synthesis of monolignols based on genetic evidence from both monocots and dicots. The brown-midrib phenotypes in maize, sorghum, and pearl millet are known to involve changes in lignin content and composition (Kuc *et al*., 1968; Porter *et al*., 1978; Cherney *et al*., 1988). Specific brown-midrib genes have been associated with changes in OMT activity in sorghum (Cherney *et al*., 1988) and maize (Grand *et al*., 1985). The maize brown-midrib3 (bm3) mutation has recently been shown to correspond to the gene encoding COMT (Vignols et ai., 1995). Transgenic dicots with reduced OMT activity show changes in lignin content and monomer composition (Dwivedi *et al*., 1994; Ni *et al*., 1994). High-leve1 antisense expression of aspen OMT in tobacco is associated with depressed COMT activity and a

reduced ratio of syringyl to guaiacyl lignin subunits (Dwivedi *et al*., 1994). Despite these small discrepancies, the data from all of the COMT-suppressed tobacco and poplar plants indicate that COMT plays a predominant role in determining the incorporation of S units into the lignin polymer in these species. These results are consistent with those of Li *et al*. (2000) who showed, for several angiosperm species, that 5-hydroxy-coniferaldehyde is the preferred substrate for COMT in vitro. Thus this enzyme is also termed as 5-hydroxyconiferaldehyde O-methyltransferase (AldOMT).

1.6.1.6 Caffeoyl-Coenzyme A O-Methyltransferase (CCoA-OMT; EC 2.1.1.104)

Until recently, the methylation reactions at the C3 and C5 hydroxyl functions of the lignin precursors were thought to occur mainly at the cinnamic acid level by bi-functional COMT. However, the association of CCoAOMT expression with lignification (Ye *et al*., 1994; Ye & Varner, 1995; Ye, 1997; Martz *et al*., 1998; Chen *et al*., 2000) and the observation that down-regulation of COMT preferentially affected the amount of S units suggested the existence of an alternative pathway for the methylation of the lignin precursors at the hydroxycinnamoyl-CoA level . Down-regulation of CCoAOMT affected the Klason lignin content by 12–50% in transgenic tobacco. Other consequences of CCoAOMT down-regulation in transgenic poplar were an enhanced fluorescence of the vessel cell walls and the accumulation of *p*-hydroxybenzoic acid esterified to lignin. Simultaneous down-regulation of both COMT and CCoAOMT in tobacco (Zhong *et al*., 1998; Pincon *et al*., 2001a) and alfalfa (Guo *et al*., 2001) resulted in combinatorial and/or additive effects on lignin content and composition.

1.6.1.7 Cinnamoyl-Coenzyme A Reductase (CCR; EC 1.2.1.44)

Reduction of hydroxycinnamoyl-CoA thioester to the corresponding aldehydes is catalyzed by cinnamoyl-CoA reductase. CCR does not in general exhibit much specificity for one hydroxycinnamoyl-CoA substrate over another, although feruloyl-CoA is reported to be the best substrate for CCR from *Forsythia* (Gross and Kreiten, 1975), soybean (Wengenmayer *et al*., 1976), hybrid poplar (Sarni *et al*., 1984), and *Eucalyptus* (Goffner *et al*., 1994). Goffner *et al*. (1994) hypothesized that CCR plays a key regulatory role in lignin biosynthesis as the first committed step in the production of monolignols from phenylpropanoid metabolites and also regulates overall carbon flux towards lignin (Lacombe *et al*., 1997). Transgenic tobacco (Piquemal *et al*., 1998; Ralph *et al*., 1998; O'Connell *et al*., 2002) and *Arabidopsis* (Goujon *et al*., 2003b) down-regulated for CCR are characterized by an approximate 50% decrease in Klason lignin. The lignin S/G ratio was increased (mainly because of a decrease in the G unit amount) in transgenic tobacco, and variable depending on the growth conditions, in transgenic *Arabidopsis*. In tobacco, an orange-brown color was observed in the xylem. The presence of unusual phenolics (such as ferulic acid and sinapic acid) in the cell wall may account for this color, because semi-*in vivo* incorporation of these two hydroxycinnamic acids into stem sections resulted in a comparable phenotype (Piquemal *et al*., 1998). The transgenic plants with the lowest CCR activity and 50% reduced lignin had abnormal phenotypes, such as collapsed vessels, stunted growth, and abnormal leaf development. Important alterations in the fiber cell walls were observed, such as a loosening in the arrangement of the cellulose microfibrils that resulted in reduced cell wall cohesion (Pincon *et al*., 2001b; Goujon *et al*., 2003b).

1.6.1.8 Coniferaldehyde 5-Hydroxylase (CAld5H)

For many years, it was generally accepted that the 5-hydroxylation of the mono-methoxylated lignin precursor, catalyzed by the enzyme ferulic acid 5-hydroxylase (F5H), would take place at the hydroxycinnamic acid level. However, Osakabe *et al*., (1999) and Humphreys *et al*., (1999) demonstrated that the 5-hydroxylation of the mono-methoxylated precursor occurs preferentially at the cinnamaldehyde level, leading to the

renaming of the enzyme to coniferaldehyde 5-hydroxylase or CAld5H. It catalyzes hydroxylation of coniferaldehyde to 5-hydroxyconiferaldehyde (CAld5H; no EC number assigned). A number of reports have indicated that the hydroxylation step at C5 may also occur at the cinnamyl alcohol level (Matsui *et al*., 1994, 2000; Daubresse *et al*., 1995; Chen *et al*., 1999; Humphreys *et al*., 1999; Parvathi *et al*., 2001). An *Arabidopsis* mutant deficient in F5H (*fah1*) was described more than 15 years ago. The mutant produced a lignin deficient in S units (Chapple *et al*., 1992) with a consequently increased frequency of phenylcoumaran (β-5) and biphenyl (5-5) linkages (Marita *et al*., 1999). Li *et al*., (2003) have overexpressed a sweetgum F5H (CAld5H) under the control of a xylem-specific promoter (*Pt4CL1P*) in transgenic aspen and reported a 2.5-fold increase in the S/G ratio and no changes in lignin content.

1.6.1.9 Cinnamyl Alcohol Dehydrogenase (CAD; EC 1.1.1.195)

The reduction of hydroxycinnamaldehydes to hydroxycinnamyl alcohols is catalyzed by cinnamyl alcohol dehydrogenase. CAD is a Zn containing dimeric enzyme which requires $NADP^+$ as a coenzyme. CAD has been considered to be an indicator of lignin biosynthesis because of its specific role at the end of the monolignol biosynthetic pathway (Walter *et al*., 1988). However, CAD is expressed in cells that do not make lignin (O'Malley*et al*., 1992; Grima-Pettenati *et al*., 1994). CAD is also expressed in response to stress (Galliano *et al*., 1993), pathogen elicitors (Campbell and Ellis, 1992a), and wounding. CAD is therefore regulated by both developmental and environmental pathways, much like other well-studied enzymes of phenylpropanoid metabolism. Differences in substrate affinities of CAD enzymes from angiosperms and gymnosperms may play a role in controlling the formation of different types of lignin (Kutsuki *et al*., 1982a). *cad* mutants exist in pine (MacKay *et al*., 1997), maize (Halpin *et al*., 1998), and *Arabidopsis* (Sibout *et al*., 2003). CAD suppression has been associated with a red or red-brown color of the stem xylem. An unusual monomer, dihydroconiferyl alcohol, was shown to be incorporated into the lignin of the pine *cad* mutant and accounted for 30% of the lignin compared to only 3% in wild-type lignin (Ralph *et al*., 1997). Isoforms of CAD

with markedly different substrate affinities are detected in such species as soybean (Wyrambik and Grisebach, 1975), *Eucalyptus* (Goffner *et al*., 1992), and *Salix* (Mansell *et al*., 1976). Many species, however, are believed to contain a single form of the enzyme. For example, there is a single CAD enzyme encoded by a single gene in loblolly pine (OMalley *et al*., 1992; MacKay *et al*., 1995). CAD preparations from gymnosperms are generally much more active on coniferaldehyde, whereas angiosperm CAD preparations show more equal activities with coniferaldehyde and sinapaldehyde.

1.6.2 Transport and Storage of monolignols

Lignin monomers are biosynthesized and then translocated to the cell wall, where they are oxidized for polymerization. The phenolic glucosides coniferin and syringin have been considered as the transport forms of coniferyl and sinapyl alcohols before their aglycones are polymerized into lignin, but this has not been demonstrated yet. Genes have been cloned that encode coniferyl and sinapyl alcohol 4-O-glucosyltransferases and cell-wall-localized b-glucosidases for coniferin and syringin. Transgenic *Arabidopsis* plants in which the expression of these glucosyltransferases is downregulated have severely reduced levels of the glucosides. However, no significant lignin phenotype was observed (A Lanot, R Dixon, and D. Bowles, personal communication). Because coniferin and syringin do not accumulate to high levels in angiosperm xylem, and coniferyl and sinapyl alcohols might have the capacity to freely diffuse through the plasma membrane (Boija *et al*., 2006), these glucosides might play no role in monolignol export for developmental lignin. Accommodation of the multitude of alternative lignin monomers might support a non-specific rather than a glucosyl transferase/b-glucosidase-mediated transport route.

1.6.3 Polymerization

Lignin polymerization is thought to result from the oxidative coupling of a monolignol to the growing polymer. Polymerization continues if a phenolic group on the lignin polymer

is oxidized to a radical, either by a peroxidase (EC 1.11.1.7) or by a peroxidase-generated monolignol radical, and the phenolic radical on the polymer is coupled to a second monolignol radical. A radical mediation model was proposed to produce a lignin molecule by Takahama (1995). The radical mediation model postulates that the formation of monolignol radicals occur through interaction with cell-wall-bound peroxidases and the monolignol, and newly formed radicals diffuse to the growing lignin polymer and follow one of two possible fates: (a) if the lignin polymer is at a base oxidative state, the higher oxidation state on the monolignol can be transferred to the lignin molecule, thus returning the monolignol to the ground state; (b) if the lignin polymer is at a higher oxidation state, the monolignol radical can undergo an oxidative coupling reaction to form a covalent bond. The radical mediation mechanism was suggested based on the findings that coniferyl alcohol radicals oxidize sinapyl alcohol. Whether the monolignol radical can be transferred to the lignin macromolecule as well as to sinapyl alcohol, however, has not yet been shown. In contrast, another model proposes a direct oxidation of the growing polymer by peroxidase. The mechanism requires that peroxidases directly oxidize both lignin macromolecules and monolignols, and that a phenolic radical on the lignin macromolecule be coupled to a monolignol radical. It is unclear, however, whether peroxidase can oxidize directly three-dimensional lignin macromolecules as substrates. Several reports on peroxidase activity or gene expression in lignin-forming tissues have appeared, but only a few isoenzymes or genes have been specifically associated with lignification (Marjamaa *et al.,* 2006). Peroxidases from *Populus alba* (L.) and *Z. elegans* were shown to oxidize Sinapyl alcohol efficiently; (Abaldon *et al.,* 2005). Antisense expression of the TP60 peroxidase gene in tobacco resulted in an equal reduction of both G and S units, suggesting the existence of a feedback regulation to decrease the monolignol synthesis and transportation under reduced oxidative capacity in the apoplast.

1.7 Why *Leucaena*?

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 comes from 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.

1.8 *Leucaena*

Leucaena is a native of Central America and has been naturalized pan-tropically. Members of the genera are vigorous, drought tolerant, highly palatable, high yielding and rich in protein and grow in a wide range of soils (Jones, 1979; Hughes, 1998). However, these attributes are limited by the occurrence of anti-nutritive factors in the fodder, such as tannins and mimosine (Jones, 1979; Hegarty *et al*., 1964b; 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. Out of them, 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: *Leucaena leucocephala*

Family: Leguminosae

Subfamily: Mimosideae

Sub tribe: Mimoseae

Genus: *Leucaena*

Species: *leucocephala*

Sub species: *Glabrata* (Rose S. Zarate); *Ixtahuacana* (huguese) and *Leucocephala* (Benth) Var. Peru and Cunningham

Related genera: *Desmanthus*; *Schleinitzia*; *Calliandropsis*; *Neptunia; Alantsilodendron; Gagnebena; Dichrostachy and Canaloa.*

1.8.1 *Leucaena leucocephala* species growth form

L. leucocephala has three growth forms that are commonly found in tropics (Hughes, 1998). The first and common type is a shrubby free shading form known as the Hawaiian *Leucaena*, which is weedy and low yielding. The second growth form is giant type, which is a tall tree of about 20 meters, with large leaves, pod and seeds, and almost without branches. It is native of Central America and Mexico, and produces twice the biomass of the common type. It is suitable for timber, wood product and industrial fuel. The last is Peru type, which is a multi-branched, semi-erect, medium height (about 10 meters) and the most productive form.

1.9 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. Bole wood has a specific gravity of 0.54 at the age of 6 to 8 years. 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 Hughes, 1995). *Leucaena* wood is among the best hardwoods for the paper and rayon making. It produces pulp that is high in hemi-cellulose, low in silica, ash, lignin, alcohol-benzene soluble and hot water soluble. Pulp yield is 50 to 52%. Its short fiber is suitable for rayon production (Pottinger and Hughes, 1995).

1.10 Why CAld5H from *Leucaena*?

This study is the first instance towards isolation and characterization of lignin biosynthetic pathway gene(s) for the development of transgenic *Leucaena* plants. This study is intended to lay down criteria for the development of the transgenic plants which would ultimately strive for: (a) altered ratio of S/G lignin (b) altered lignin content, (c) increased cellulose content and (d) enhanced growth. These criteria would be met by isolation, cloning and characterization of the selected and key lignin biosynthesis pathway genes *viz.* cinnamyl alcohol dehydrogenase (CAD), 4-coumarate-CoA-ligase (4CL), Caffeate Omethyltransferase (COMT or AldOMT), caffeoyl-CoA-3-O-methyltransferase (CCoAOMT), cinnamoyl-CoA-reductase (CCR), and Coniferaldehyde 5-hydroxylase (CAld5H), UDP glucose-glucosyl transferase (UDPG-GT) and coniferin-β-glucosidase (CBG). The above genes of the lignin biosynthesis pathway in the target plant species could then be down/up regulated by transformations with the gene constructs in sense or antisense orientations under the control of suitable tissue specific promoters. It will, however, be beyond the scope of this thesis to cover all the genes listed above. Pulp and paper industry will benefit by either reduced lignin synthesis or by alteration of the lignin components in a manner where in the sinapyl alcohol to guaiacyl alcohol ratios are in favour of sinapyl alcohol. The later of the two may be achieved by regulating the Coniferaldehyde 5-hydroxylase (CAld5H), which will divert the flux towards the formation of sinapyl alcohol. Hence, the present study was aimed at the isolation of CAld5H cDNA clone from *Leucaena leucocephala,* heterologous protein expression, studying the pattern of CAld5H gene expression in *Leucaena* and finally genetic transformation analysis in tobacco as well as *Leucaena leucocephala.*

1.11 CAld5H is a cytochrome P450 protein (CYP84A1)

Coniferaldehyde 5-hydroxylase is the enzyme responsible for the last hydroxylation of the syringyl-type lignin precursors. This enzyme belongs to Cytochrome P450 monoxygenase family. Cytochrome P450 proteins, named for the absorption band at 450 nm of their carbon-monoxide bound form, are one of the largest super families of enzyme proteins. The P450 genes (also called CYP) are found in the genomes of virtually all organisms, but their number has exploded in plants. Their amino-acid sequences are extremely diverse, with levels of identity as low as 16% in some cases, but their structural fold has remained the same throughout evolution. P450s are heme-thiolate proteins; their most conserved structural features are related to heme binding and common catalytic properties, the major feature being a completely conserved cysteine serving as fifth (axial) ligand to the heme iron. Canonical P450s use electrons from NAD(P)H to catalyze activation of molecular oxygen, leading to regiospecific and stereospecific oxidative attack of a plethora of substrates. The reactions carried out by P450s, though often hydroxylation can be extremely diverse and sometimes surprising. They contribute to vital processes such as carbon source assimilation, biosynthesis of hormones and of structural components of living organisms, and also carcinogenesis and degradation of xenobiotics. In plants, chemical defense seems to be a major reason for P450 diversification. In prokaryotes, P450s are soluble proteins. In eukaryotes, they are usually bound to the endoplasmic reticulum or inner mitochondrial membranes. The electron carrier proteins used for conveying reducing equivalents from NAD(P)H differ with subcellular localization. P450 enzymes catalyze many reactions that are important in drug metabolism or that have practical applications in industry; their economic impact is therefore considerable.

1.11.1 Characteristic Structural Features

P450s can be divided into four classes depending on how electrons from NAD(P)H are delivered to the catalytic site. Eukaryotic class I enzymes are found associated with the inner membrane of mitochondria and catalyze several steps in the biosynthesis of steroid hormones and vitamin D and require both an FAD-containing reductase and an iron sulfur redoxin. Class II enzymes are the most common in eukaryotes. P450s and NADPH-P450 reductases are dissociated and independently anchored on the outer face of the endoplasmic reticulum (ER) by amino-terminal hydrophobic anchors and require only an FAD/FMN-containing P450 reductase for transfer of electrons. Class III enzymes are self-sufficient and require no electron donor and involve in the synthesis of signaling molecules such as prostaglandins in mammals and jasmonate in plants, whereas P450s from class IV receive electrons directly from NAD(P)H.

Sequence identity among P450 proteins is often extremely low and may be less than 20%, and there are only three absolutely conserved amino acids. The determination of an increasing number of P450 crystal structures, however, shows that this unusual variability does not preclude a high conservation of their general topography and structural fold (Graham SE *et al.,* 1999). Highest structural conservation is found in the core of the protein around the heme and reflects a common mechanism of electron and proton transfer and oxygen activation. The conserved core is formed of a four-helix (D, E, I and L) bundle, helices J and K, two sets of 1 sheet, and a coil called the 'meander'. These regions comprise (Fig.1. 7a) first, the heme-binding loop, containing the most characteristic P450 consensus sequence (Phe-X-X-Gly-X-Arg-X-Cys-X-Gly), located on the proximal face of the heme, just before the L helix, with the absolutely conserved cysteine that serves as fifth ligand to the heme iron; second, the absolutely conserved Glu-X-X-Arg motif in helix K, also on the proximal side of heme and probably needed to stabilize the core structure; and finally, the central part of the I helix, containing another consensus sequence considered as P450 signature (Ala/Gly-Gly-X-Asp/Glu-Thr-Thr/Ser), which corresponds to the proton transfer groove on the distal side of the heme. The most variable regions are associated with either amino-terminal anchoring or targeting of membrane-bound proteins, or substrate binding and recognition; the latter regions are located near the substrate-access channel and catalytic site and are often

referred to as substrate-recognition sites or SRSs (Gotoh O, 1992). They are described as flexible, moving upon binding of substrate so as to favor the catalytic reaction. Other variations reflect differences in electron donors, reaction catalyzed or membrane localization (Fig.1.7b). Most eukaryotic P450s are associated with microsomal membranes, and very frequently have a cluster of prolines (Pro-Pro-X-Pro) that form a hinge, preceded by a cluster of basic residues (the halt-transfer signal) between the hydrophobic amino-terminal membrane anchoring segment and the globular part of the protein (Fig.1.7a). Additional membrane interaction seems to be mediated essentially by a region, located between the F and G helices that show increased hydrophobicity (Williams PA *et al.,* 2000). A signature for mitochondrial enzymes is, in plants Class II P450s are involved in biosynthesis or catabolism of all types of hormones, in the oxygenation of fatty acids for the synthesis of cutins, and in all of the pathways of secondary metabolism in lignification and the synthesis of flower pigments and defense chemicals (which are also aromas, flavors, antioxidants, phyto-estrogens, anti-cancer drugs and other drugs).

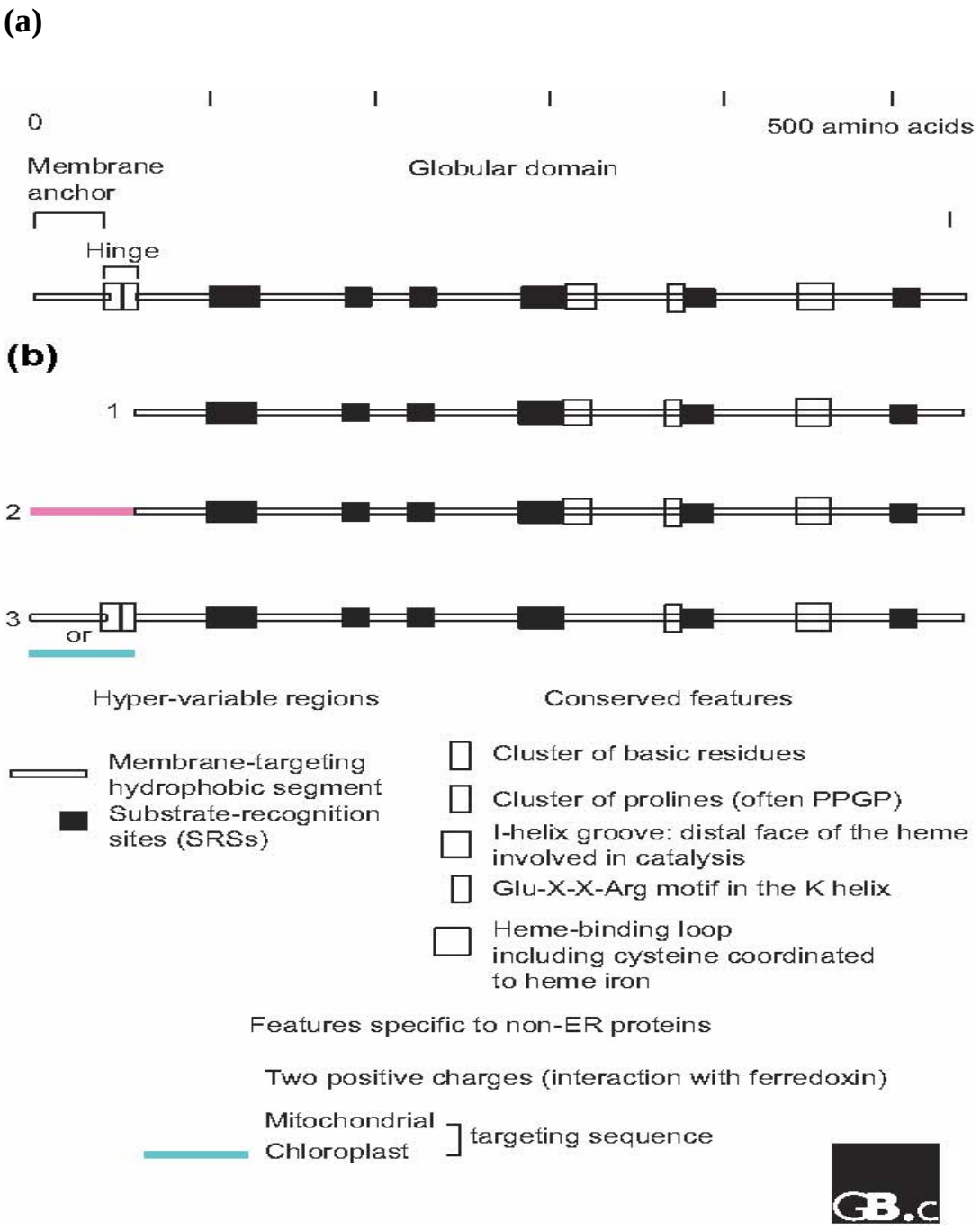


Figure 1.7. Primary structure of P450 proteins. **(a)** Typical features of an ER-bound P450 protein (class II enzyme). The function of the different domains and regions indicated by colors are described in the text. **(b)** Variants of this canonical structure most commonly found: 1, soluble class I; 2, mitochondrial class I; 3, membrane-bound or plastidial class III.

Chapter 2

Materials and Methods

This chapter includes general laboratory techniques routinely followed in the studies. 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. leucocephala* cultivar K-636. Seeds of *L. leucocephala* (K-636) were treated with conc. H_2SO_4 for 2 – 3 min. and 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. Seeds were soaked in sterile water for two days at 30-37 ºC 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 autoclaving. The culture bottles were incubated at 25± 2 ºC and 60% relative humidity under 16 h photoperiod of light intensity 24.4 µmol /m^2/s. One month old axenic cultured plants were the raw material for further experiments.

2.1.2 *Nicotiana tabacum*

Tobacco seeds *Nicotiana tabacum* var. Anand 119 was germinated on wet sterile paper towel. Germinated seeds were transferred to ½x 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 autoclaving. The culture bottles were incubated at 25± 2 ºC and 60% relative humidity under 16 h photoperiod. One month old axenic cultured pant leaves were the raw material for further experiments.

2.2 Glassware

Glassware used in all the experiments were procured 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 1 h 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). Autoclaving of the glassware and above items was done at 121 ºC 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 procured from "Laxbro", India. Microfuge 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), Affigel 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. coniferaldehyde, coniferyl alcohol, ferulic acid the cofactor such as NADPH and the δ-aminolevulenic acid and the $FeCl_3$ which has been used in protein expression as heme domain precursors are also obtained from Sigma-Aldrich (USA).

2.5 Equipment

Table 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 Documentation system	Bio-Rad
7	Thermo Cycler PCR machine/Real Time PCR	MJResearch/Stratagene
8	Spectrophotometer	Applied Biosystem
9	Power pack	Bio-Rad
10	Agarose Gel	Bangalore Genei/ Bio-Rad

	Electrophoresis Units	
11	protein Gel Electrophoresis Units	Hoeffer Scientific/ BioRad
12	Speed Vac concentrater	Savant/Eppendorf
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
19	Bioanalyser agilent 2100	Agilent Technology
20	Typhoon Trio + Scanner	GE Healthcare (USA)
21	FPLC AKTA Explorer	GE Healthcare (USA)
22	Particle bombardment system	Bio Rad
23	ELISA Plate Reader	Amersham (USA)

2.6 Buffers and Solutions

2.6.1 Buffers and Solutions for DNA Electrophoresis

Table 2.2 Buffers and Solutions for DNA Electrophoresis.

Name	Ingredients	Preparation and Storage
50x TAE	2 M Tris Acetic acid 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 100 ml with H_2O	The solutions were filter sterilized using 0.22 micron filter and stored at RT

2.6.2 Buffers and Solutions for g DNA Isolation, Southern and Slot Blot

Table 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 0.5 M NaOH	RT
Neutralization buffer	1.5 M NaCl 1.0 M Tris HCl (pH 7.4)	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.*

Table 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 ºC
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 ºC
Ampicillin	100 mg/mL in SMQ	Sterile filtration and storage at -20 ºC
Tetracycline	12.5 mg/mL in 70% ethanol	Sterile filtration and storage at -20 ºC
Kanamycin	50 mg/mL in SMQ	Sterile filtration and storage at -20 ºC
Rifampcin	50 mg/mL in DMSO	Sterile filtration and storage at -20 ºC
Hygromycin	25 mg/mL in SMQ	Sterile filtration and storage at -20 ºC

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

Table 2.5 Buffers and Solutions for plasmid isolation (Alkaline lysis method).

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

2.6.5 Buffers and Solutions for Gel Electrophoresis (PAGE)

Table 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 ºC 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 4 µL	
Separating gel (10%)	Distilled Water 3.3 mL 1.5 M Tris-HCl (pH 8.8), 2.0 mL Acrylamide/Bis 30%, 4.0 mL 10% (w/v) SDS, 0.1 mL (SDS-PAGE) 10% (w/v) APS, 0.1 mL TEMED, 7 µL	
2x Protein loading buffer	Distilled Water, 2.7 mL 0.5 M Tris-HCl (pH 6.8), 1.0 mL Glycerol, 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 ºC
10x SDS-electrode buffer	Tris base 15.1 g Glycine 94 g SDS 0.5 g Adjust pH-8.3 And make up the volume up to 500 ml.	Store at 4 ºC dilute 1:10 before use
Staining solution	Coomassie-blue R 250, 0.25 g Methanol, 40 mL Acetic acid, 10 mL Water up to 100 mL	
Destaining solution	Methanol, 40 mL Acetic acid, 10 mL Distilled water up to 100 mL	

Silver staining Fixing solution	40% Methanol, (150 mL) 10% acetic acid, (50 mL) Distilled water up to 100 mL	Store at room temperature
Sensitizing solution	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 (225 µL) Distilled water up to (300 mL)	Prepare fresh and in dark
Developing solution	6% Na_2CO_3 (18 g) 0.02% formaldehyde (150 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

Table 2.7 Buffers and solutions for protein purification (based on imidazole conc.).

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 denaturing conditions

Table 2.8 Buffers and solutions for protein extraction under denaturing conditions

Name	Ingredient	Preparation and Storage
Lysis buffer	50 mM Tris-HCl (pH 8.0) 5 mM EDTA 100 mM NaCl 0.5% TritonX-100 0.7 mM DTT 0.1 mM PMSF (Freshly added) 10 mM $MgSO_4$ Lysozyme 100 µg/mL (Added freshly)	Stored at 4 ºC
Sonication buffer	100 mM Tris HCl (pH8.0) 50 mM Glycine	Stored at 4 ºC
Dispersion buffer:	100 mM Tris-HCl (pH8.0) 50 mM Glycine 8 M Urea /6 M GuHCl	Stored at 4 ºC

2.6.8 Buffers and solution for the CAld5H enzyme assay

Table 2.9 Buffers and solutions for the CAld5H enzyme Assay

Name	Ingredients	Preparation and Storage
Phosphate buffer	100 mM Na_2HPO_4	Adjust pH by adding either or NaH_2PO_4 and Store at 4 ºC
Sodium Chloride	150 mM NaCl	RT

2.6.9 Different media used for Studies

Table 2.10 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 ºC
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 +ºC
TB buffer	10 mM PIPES 15 mM $CaCl_2$ 250 mM KCl	pH was adjusted 6.8 with KOH. $MnCl_2$ was added to 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 ºC
Terrific Broth	12 g Bactotryptone 24 g Yeast extract 9.4 g K_2HPO_4 2.2 g KH_2PO_4 4 mL Glycerol Make up the vol. 1L.	pH was adjusted to 7.5 and autoclaved.

2.6.10 Buffers and Solutions for ELISA/Immunocytolocalization/ GUS Assay/particle coating.

Table 2.11. Buffers and Solutions for ELISA/Immunocytolocalization/GUS Assay/ particle coating.

Name	Ingredients	Preparation and Storage
Crude protein extraction buffer	100mM Tris (pH 7.4) 2% PVPP 2% PEG 6000 10 mM PMSF 5 mM DTT	Stored at RT
PBS T	1.44 g Na_2HPO_4 0.24 g KH_2PO_4 0.2 g KCl 8 g NaCl 0.05% v/v Tween-20	Stored at RT
Antibody dilution Buffer	PBS with 0.25% BSA	store at 4 ºC
Substrate buffer	200 mM TrisCl pH 9.5 0.5 mM $MgCl_2$	Store at 4 ºC

1X PBS	10 mM NaH_2PO_4-Na_2HPO_4 Buffer (pH 7.2) 130 mM NaCl	Adjust pH to pH 7.2 and Stored at 4 ºC
0.5 X SSC	75 mM NaCl 7.5 mM Na Citrate	Adjust pH to pH 7.0 and Stored at 4 ºC
Color development buffer	100 mM Tris (pH 9.5) 150 mM NaCl, 50 mM $MgCl_2$	Stored at 4 ºC
BCIP/NBT mix	0.577 mM BCIP 0.122 mM NBT	Stored at 4 ºC In Dark
Stop Solution	10 mM EDTA	Stored at RT
Other reagents used	Polyvinyl alcohol, Ethanol, Tertiary butanol, Paraffin, Xylene, Glycerol	

X-GluC	1 mM X-Gluc (5-bromo, 4-chloro, 3-indolyl-β-D-glucuronide: Cyclohexylammonium salt (X-GlcA) in dimethylformamide 100 mM sodium dihydrogen phosphate dihydrate 0.5% Tween-20	Adjust pH to 7.0 and Stored at 4 ºC
Phloroglucinol	2% Phloroglucinol in 95% Ethanol	Stored at 4 ºC
Xho buffer	150 mM Nacl 10 mM Tris Cl pH 8.0	Stored at -20 ºC
Spermidine	0.1 M Spermidine in H_2O	Stored at -20 ºC
$CaCl_2$	2.5 M $CaCl_2$	Stored at -20 ºC

2.6.11 Component of murasighe and Skoog Media/ Different inducing media and hormones

Table 2.12 Component of Murasighe and Skoog media / Different inducing media and hormones

Name	Ingredients	Preparation and Storage
Major component	20.61 mM NH_4NO_3 18.75 mM KNO_3 2.99 mM $CaCl_2.2H_2O$ 1.5 mM $MgSO_4.7H_2O$ 1.24 mM KH_2PO_4	Stored At 4 ºC
Minor components	0.147 mM $MnSO_4$ 5.3 x 10^{-2} mM $ZnSO_4$ 1.56 x 10^{-4} mM $CuSO_4$ 1.50 x 10^{-4} mM $COCl_2.6H_2O$ 4.99 x 10^{-3} mM KI 0.1mM H_3BO_4 1.03 x 10^{-3} mM $Na_2Mo_4.2H_2O$	Stored At 4 ºC
Vitamins	5.55 x 10^{-2} mM Myoionsitol 4.06 x 10^{-3} mM Nicotinic	Stored At 4 ºC

	acid 2.43 x 10^{-3} mM Pyridoxine HCl 2.96 x 10^{-4} mM Thymine HCl 2.66 x 10^{-2} mM Glycine	
Iron	0.1mM $FeSO_4.7H_20$ 0.1mM Na_2EDTA	Stored At 4 ºC
BAP	1.776 mM BAP (dissolve in NaOH and make up the volume by adding ethanol)	Stored At 4 ºC
NAA	2.148 mM NAA (Dissolve in H_2O)	Stored At 4 ºC
Shoot Induction Medium	Major (40x)-25 mL Minor (100x)-10 mL Iron (100x)-10 mL Vitamins (200x)-5 mL BAP (4.4 µM))-2.5 mL NAA (5.37 µM)-250 µL Glucose-1.5% Sucrose-2.0% pH-5.6 to 5.8	Stored At 4 ºC

	Agar-0.8% (make up to volume 1L)	
Resuspension Medium	Major (40x) 25mL Minor (100x) 10 mL Iron (100x) 10 mL Vitamins (200x) 5 mL BAP (4.4µM) 2.5mL NAA (5.37µM) 250µL Glucose 1.5% Sucrose 2.0% pH-5.6 to 5.8 Acetosyringone-200 µM $MgSO_4$ 40 mM (make up to volume 1L)	Stored At 4 ºC
Root Induction Medium	Major(40x) 25 mL Minor(100x) 10 mL Iron (100x) 10 mL Vitamins (200x) 5 mL NAA(5.37 µM) 250 µL Glucose 1.5% Sucrose 2.0%	Stored At 4 ºC

	pH-5.6 to 5.8 Agar 0.8% (make up to volume 1L)	
Selection medium	7.09 μM hygromycin 200/250 μM cefotaxime	

2.7 Host Cells

Table 2.13 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 ºC 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 medium 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 inoculated in 5 ml of LB medium containing tetracycline (12.5 mg / L). LB medium (50 mL) was inoculated 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 ºC, suspended in ice-cold 100 mM $CaCl_2$ and kept on ice for 30 min. Cells were centrifuged, the pellet suspended in 1 mL of ice-cold 100 mM $CaCl_2$ and stored as aliquots of 200 µL at 4 ºC or in -70 ºC after adding 15% glycerol or 7% DMSO.

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 ºC for 1.5 min (heat shock) and immediately kept back on ice. To each tube 800 µL of LB broth was added and further incubated at 37 ºC for 1 h. Cells were pelleted by centrifugation at 5000 rpm for 5 min. The supernatant media was discarded and pellet was dissolved in 100 µL volume of LB broth. The resuspended cells were plated on LB agar medium with appropriate antibiotics, IPTG and X-gal as per need (Sambrook *et al*., 1989).

Solutions	Stock	Final conc.
1) IPTG stock solution	200 mg mL^{-1} in sterile distilled water	40 µg mL^{-1}
2) X-gal stock solution	20 mg mL^{-1} dimethylformamide	40 µg mL^{-1}

2.8.2.3 *Agrobacterium tumefaciens* transformation and selection

The method adopted for the preparation of *A. tumefaciens* competent cells was essentially as reported by An (1987). A single colony of *A. tumefaciens* (GV2260) was inoculated in 50 mL LB broth containing Rifampicin 150 µg mL^{-1} and incubated at 28 ºC with shaking at 200 rpm till O.D. 0.5 at 600 nm. Cells were centrifuged at 5,000 g for 10 min at 4 ºC and washed twice with ice cold 150 mM $CaCl_2$. The cells were pelleted and resuspended in 1mL of ice cold 20 mM $CaCl_2$. Aliquots of 200 µL were made. For transformation 1 µg DNA of the desired binary plasmid vector was added to an aliquot of the competent *A. tumefaciens* cells and incubated on ice for 30 min. The cells were then snap frozen in liquid nitrogen and allowed to thaw at 37 ºC. After thawing 1mL LB medium was added and the tubes incubated at 28 ºC for 2 h with gentle shaking. The cells were centrifuged at 4,000 g for 5 min and 100 µL supernatant was retained. The cells were resuspended in the 100 µL supernatant and plated on LB agar medium with appropriate antibiotic(s). The plates were incubated at 28 ºC for two days to allow the transformed colonies to grow for further processing.

2.8.3 Colony Screening by PCR

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 picked up from the agar plate containing transformants with the help of microtip and added to 1.5 mL eppendorf PCR tube containing 25 µL of sterile miliQ water. The microtip was agitated in the water to suspend the bacterial cells. Subsequently, 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 ºC before the amplification cycles will aid the bacterial lysis to enhance PCR product amplification success. The resulting PCR products were checked on an agarose gel for the presence of the amplicon of expected size.

2.8.4 Isolation of Nucleic Acids and Polymerase 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 processed 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 protocol 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 with shaking (200 rpm) at 37 ºC in LB broth, with appropriate antibiotic(s). About 1.5 to 3 mL culture was centrifuged for 1 min at 7000 g to pellet the bacterial cells. The pellet was resuspended in 100 µL of GTE 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 normally for 2-3 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 centrifugation for 10 min at 12000 g at 4 ºC. The supernatant was transferred to another microfuge tube, RNase A was added to a final concentration of 20 µg/mL (Sambrook *et al*., 1989) and incubated at 37 ºC for 20 min. To the above solution 400 µL of chloroform was added, vortex for 2 min and centrifuged for 5 min at 12, 000 g at room temperature. The upper aqueous layer was transferred to a clean tube, 1/10th volume sodium acetate and one volume absolute ethanol was added

with gentle mixing and kept at -20 ºC for 1-2 h. The sample was centrifuged at 12,000 g for 10 min at 4 ºC. 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 12,000 g for 15 min at 4 ºC. 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 ºC.

2.8.4.2 Isolation of plant Genomic DNA

Extraction buffer: 100 mM Tris-HCl and 20 mM sodium EDTA adjust pH to 8.0 with HCl; add NaCl to 1.4 M and 2.0% (w/v) CTAB (cetyltrimethylammonium bromide). Dissolve CTAB by heating to 60°C. Store at 37 ºC. Add β-mercaptoethanol to 0.2 % before use. Chloroform: isoamyl alcohol 24:1(v/v), 5 M NaCl, RNAase A (10 mg mL^{-1}), 95% ethanol, 75% ethanol, TE buffer: 10 mM Tris-HCl and 1 mM EDTA, adjust pH to 8.0 and autoclave.

Genomic DNA was isolated by using the protocol 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, autoclaved 50 mL centrifuge tube and 100 mg insoluble polyvinylpoly-pyrrolidone (PVPP) as well as 20-40 µL of β-mercaptoethanol was also added. The tube was inverted several times but very gently to thoroughly mix the slurry, incubated at 65 ºC for 30 min and then allowed to cool down to room temperature. 12 mL of chloroform: isoamylalcohol mixture was added and the contents mixed by inverting the tube gently till an emulsion formed. The mixture 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 and two volumes of cold (-20 ºC) 95% ethanol was added and the sample kept at 4 ºC 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, and DNA pellet washed with cold (4 ºC) 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^{-1}) per 100 µL DNA and incubated at 37 ºC for 30 min. The sample was extracted with chloroform: isoamyl alcohol to remove RNAase A. DNA was re-precipitated and dissolved in 40-100 µL TE buffer. Purity of DNA was checked spectrophotometrically by measuring the absorbance ratio (A_{260}/A_{280}) and also by visualization on 0.8 % agarose gel with 1X TAE. Genomic DNA was stored at 4 ºC.

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^{-1}) 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™ GEL elution kit, Biosciences, USA) added. The tube was incubated at 70 ºC 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 collection 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 30 s, 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 as well as ethanol. 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 ºC. This eluted PCR product or any DNA fragments are of good quality and can be visualized on 1% Agarose gel by taking an aliquots of 3-4 µL. The Eluted DNA /PCR product can be used in subsequent studies.

2.8.4.5 RNA Extraction

RNase free environment was created and maintained as described by Blumberg (1987). All glass and plasticware were DEPC (0.1% in water) treated overnight and autoclaved. 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 13,000 g for 15 min at 4 ºC. The supernatant was transferred to 1.5 mL tubes and the chloroform: isoamyl alcohol step was repeated. The aqueous phase was transferred to 1.5 mL tubes and 0.6 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 ºC. 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 ºC 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 as well as with the Bioanalyser system. The amount of RNA was in the range of 1- 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 ºC in a heating block. Sample was removed from the heating

block, and placed at 20 – 30 ºC 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 microcentrifuge tube and centrifuged for 1 min at maximum speed. Spin column was transferred to a new RNase-free 1.5 mL microcentrifuge tube, appropriate amount of OW2 buffer added to the column and again centrifuged for 1 min at 12,000 g, flow-through was discarded. Spin column was transferred to a new RNase free 1.5 mL microcentrifuge tube. Appropriate amount of hot (70 ºC) OEB buffer was pipetted onto the column and resuspended by pipetting up and down; spin it down for 1 min at 12,000 g. The Eluted mRNA sample was quantified spectrophotometrically and used for further downstream processes.

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 1 cm light path quartz cuvette. Absorbance value (A_{260}) should fall between 0.1-1.0 to be accurate. Sample dilution was adjusted accordingly. An absorbance of 1.0 at 260 nm corresponds to 50 µg DNA/ mL.
The Bioanalyser system (Agilent 2100) was also used to check the quantity of the nucleic acids (according to manufacturer's instructions). Some other useful parameters like gel electrophoresis pattern of the samples (Nucleic acids and protein), the integrity of the RNA and DNA, concentration of the individual bands can also be analyzed simultaneously with this system. The basis of the system to perform multiple analysis simultaneously is that, one has to load the samples, standards, suitable ladder and the gel matrix in the different wells of a single microchip provided with the system and then the system primarily performs electrophoretic separation of all the samples loaded through the inbuilt micro-capillaries of the chip and then after analyses the various parameters in one go.

2.8.4.8 First strand cDNA synthesis by Reverse Transcription

Complementary DNA (cDNA) was 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 process 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 mixture was isothermally denatured at 70 ºC 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 for 5 min, the reaction was incubated at 42 ºC for up to 1-1.5 h. The cDNA synthesized was directly added to amplification reactions. The first strand reaction was set up as follows:

Reagent	**Volume**
Experimental RNA (1 µg)	1.0 µL
Primer [Oligo$(dT)_{15}$ or Random (10 pmol)	1.0 µL
DEPC treated Water	3.0 µL
Final volume	**5.0 µL**

The tubes were incubated at 70 ºC for 5 min and then chilled in ice-water for 5 min. Tubes were briefly spun in a microcentrifuge to collect the condensate and maintain the

original volume. The tubes were kept closed on ice until addition of the reverse transcription reaction mix. The reverse transcription reaction mix was prepared by adding the following components of the ImProm-II Reverse Transcription System in a sterile 1.5 mL microcentrifuge tube on ice.

Reagent	**Volume**
ImProm-II. 5X Reaction Buffer	4.0 µL
$MgCl_2$ (15 mM)	2.0 µL
dNTP Mix (7.5 mM)	1.0 µL
RNasin® Ribonuclease Inhibitor (40 U/ 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 reaction mixure was incubated at 25 ºC for 5 min for primer annealing and then at 42 ºC for 1 h for cDNA first strand synthesis. Reverse transcriptase was thermally inactivated by incubation at 70 ºC for 15 min prior to proceeding 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 (1 U/µL)	0.3 µL
Total volume	**15.0** µL

PCR cycle conditions

1 cycle 95 ºC 5 min

35 cycles 95 ºC –1 min/ 45-65 ºC -30-90 s (Ta = Tm-5 ºC)/72 ºC - 1.5 min.

1 cycle 72 ºC 10 min

1 cycle 4 ºC 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 to perform RACE. The reactions were set up as per the manufacturer's guidelines with

minor modification. 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	7.0 µL
10X CIAP Buffer	1.0 µL
RNaseOut™ (40 U/µL)	1.0 µL
CIP (10 U/µL)	1.0 µL
DEPC water	0 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 ºC at 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 of 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	7.0 µL
10X TAP Buffer	1.0 µL
RNaseOut™ (40 U/µL)	1.0 µL
TAP (0.5 U/µL)	1.0 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 ºC 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 ºC for 5 minutes and kept on ice

Reagent	Volume
10 X Ligase Buffer	1.0 µL
10 mM ATP	1.0 µL
RNaseOut™ (40 U/ µL)	1.0 µL
T4 RNA ligase (5 U/µL)	1.0 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 ºC 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. The decapped full length mRNA ligated with GeneRacer™ RNA - Oligo was used to prepare cDNA by reverse transcription. 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 ºC 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 ºC for 1 h. The reaction was inactivated by incubating the reaction mixture at 70 ºC 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 ºC for 20 min. Centrifuged the sample and kept at -20 ºC for further use in PCR.

PCR reaction 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 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, 5 U/ 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 conditions

Temperature	Time	Cycles
94 ºC	2 min	1
94 ºC	30 S	5
72 ºC 1 min/ 1 kb	2 min	
94 ºC	30 S	5
70 ºC 1 min/ 1 kb	2 min	
94 ºC	30 S	25
66 ºC	30 S	
72 ºC	2 min	
72 ºC	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 on the next page.

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, 5 U/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 condition

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

10-20 μL of nested PCR product was analyzed 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.

2.8.5 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 preferred 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.

2.8.5.1 Pre-protocol 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 ºC 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 real-time 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. Reading 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 50 to 300 nM in increments of 50 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 100 nM to 900 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. In the present studies, 30 nM concentration of reference dye has been used.

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 protocol) or the annealing/extension step (2-step cycling protocol).

2.8.5.2 Preparing the RT-qPCR 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 3005 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	12.5 µL
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**

2.8.5.3 PCR Cycling Programs

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

Temperature	Time	No. of Cycles
95 ºC	10 min	1
95 ºC	30 S	40
55-57 ºC	30 S	
72 ºC	1 min	

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

2.8.6 Nucleic Acids Hybridization

For Southern hybridization (Southern, 1975) the DNA samples were electrophoresed on an agarose gel in 1X TAE buffer containing 0.5 µg mL^{-1} ethidium bromide. The gel was rinsed with deionized water (D/W) and placed in depurination solution for 15 min. The gel was rinsed with deionized water and immersed in denaturation solution for 30 min with gentle shaking. The gel was again rinsed with deionized water and transferred to neutralization solution for 45 min. The gel was next set up for capillary transfer of DNA

to solid membrane support. A tray was filled with the transfer buffer (20X SSC). A platform was made and covered with a wick, made from 2 sheets of Whatman 3 MM filter paper saturated with transfer buffer and the gel was placed on it. It was surrounded with Saran Wrap to prevent the transfer buffer from being absorbed directly by the paper towels stacked above the membrane. A sheet of Hybond-N^+ membrane (Amersham, UK) of the exact gel size was wetted with deionized water followed by transfer buffer (20X SSC) and then placed on top of the gel. A glass rod was rolled over the membrane to remove any trapped air bubbles. One piece of Whatman 3MM paper wetted with 20X SSC was placed on the membrane followed by Whatman 3MM paper pre-wetted in 2X SSC. On this paper another dry Whatman 3MM paper was placed followed by a stack of absorbent paper towels. A glass plate and a ~0.5 kg weight were placed on the top of the paper towels. Transfer of DNA was allowed to proceed for 18 h. The membrane was marked for orientation, removed carefully and washed with 6X SSC. The membrane was air dried and baked for 2 h at 80 ºC to immobilize DNA onto the nylon membrane. Hybridization and autoradiography were carried out as is described in the following section.

2.8.6.1 Southern Hybridization/ Slot blot

Genomic DNA extracted from *L. leucocephala* was quantified and checked for its digestibility and quality as well. 20 µg of gDNA was digested up to completion with individual restriction enzymes in separate eppendorf tube. The restriction enzyme was selected on the basis of its presence within the gene of interest. The digested DNA was electrophoretically separated on 0.8% agarose gel.

For slot blot hybridization DNA or RNA samples were diluted according to experimental requirements. The DNA samples were denatured by adding 1/10th volume of 3 M NaOH and incubated at 65 ºC for 10 min. To the denatured sample an equal volume of 6X SSC was added. Two layers of Whatman 3 MM filter paper wetted with sterile deionized water and 6X 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 6X 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 baked for 2 h at 80 ºC to immobilize DNA. Radiolabelling of probe, hybridization and autoradiography were carried out as described in the following sections.

2.8.6.2 Random Primer Labeling

Random primer labeling of the DNA probes (Feinbeng and Vogelstein, 1983) 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 mmol)	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 ºC for 1 to 1.5 hr and stopped by incubation at 80 ºC 10 min and snaps chilled on ice.

2.8.7 Hybridization

The blots made as in sections 2.8.5.0 and 2.8.5.1 above were prehybridized at 62 ºC 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 62 ºC for 14-18 h. The solution was discarded and the membrane washed with low stringency buffer at 55-65 ºC for 5 min followed by a high stringency wash at 55-65 ºC for 5 min. The moist blot was wrapped in saran wrap and placed in the developing / intensifying screen provided with the Typhoon trio+ scanner and closed it. After 6-8 hrs of exposure, the remove the membrane and scan the screen with Typhoon trio scanner system. One can increase the exposure time if the band intensity is weak. The membrane scanning is done usually in the resolution mode instead of sensitivity mode of the scanner. Good quality of band pictures can be obtained by scanning at 500 micron resolution with the system. This instrument will provide many more options and functions to get exquisite band pictures of the exposed membrane. After scanning the intensifying screen, the signals can easily be erased from it, upon exposure to a white light (white light lamp is provided with the system) for 10 min and the same screen can be used to develop the other membrane.

2.8.8 Expression and purification of recombinant protein

Heterologous expression and activity of LCald5H was done in *E.coli* (BL21) host cell. The transformants were grown at 37 ºC in Luria –Bertani medium containing kanamycin (50 µg/ml). One hundred mililitres of terrific broth containing kanamycin was inoculated

with 1ml of the overnight grown culture and incubated at 37 ºC and 200 rpm shaking. When the optical density (A_{600}) reached 0.6-0.8, 1 mM δ-aminolevulinic acid (ALA) and 0.5 mM $FeCl_3$ were added to the culture. After further incubation for 20 min at 25 ºC, the expression was induced by the addition of 1 mM IPTG, and cells were further incubated at 25 ºC. For protein visualization cells were harvested up to 5 h at each hour, after IPTG induction and isolated by centrifugation. The pelleted cells were resuspended in sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) sample buffer and heated at 90 ºC. Protein were separated by SDS-PAGE on 10% gels and stained with Coomassie blue. Protein over-expression was evident from SDS- PAGE gel visualization and also the expression level was increased gradually as time passes. The Cytosolic as well as insoluble fractions was prepared and analyzed for the LlCAld5H protein expression which shows the presence of LlCAld5H protein in inclusion body.

2.8.8.1 Protein isolation from inclusion body

Cells were isolated by centrifugation and resuspended in 100 mM potassium phosphate buffer, pH 7.5, containing 2 mM EDTA, 20% glycerol, 1.5 mM DTT, 1 mM PMSF and 0.4% Triton. Lysozyme (0.5 mg/ml) was added, and the suspension was incubated at 37 ºC. Bacterial cells were further lysed by sonication on ice with an MSE soniprep cell disrupter by using 20 pulses for 10 s with 20 s-cooling intervals. Cell debris was removed by centrifugation and the pellet was resuspended in sonication buffer which contain 100 mM Tris (pH 8.0) and 50 mM glycine. Cells were sonicated once again for 5-6 pulses 10 s on and 20 second cooling.

2.8.8.2 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 or by lowering the pH of the washing

and elution buffer. The column(s) and buffers were equilibrated to room temperature. The Ni^{+} beads column was equilibrated with binding buffer for 30 min 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 ºC for 1 h for binding of recombinant protein to Ni^{+} beads. Flow through was collected in different tubes after 1 h and column was washed with two bed volume of washing buffer (chapter 2: section 2.6.6). 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). 6x His-tagged 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 separated and analyzed by SDS-PAGE.

2.8.8.3 Enzyme Activity

Overexpressed LlCAld5H is forming inclusion body, which limits the unfolding of the protein. The enzyme activity was detected from crude lysate by varying certain parameters for protein expression such as temperature and IPTG concentration. To get the LlCAld5H protein in active form various concentration of IPTG ranging from 0.01 to 1 mM and different temperature conditions were tried. The protein expression in the cytosolic fraction was detected with ELISA done with Anti-LlCAld5H. In the preliminary studies the good expression level co-coordinated with activity was detected at 20-25 ºC and 0.03 mM IPTG after 24-30 hr of incubation. The absorbance of NADPH at 340 nm was measured spectrophotometrically which gradually declines due to the reduction of NADPH. Prior to initialization of enzyme assay the auto-oxidation of NADPH was checked. The enzyme assay was set up as given in the table below.

Table 2. 14. Enzyme assay

Reagents	Blank	Test
100 mM Phoshphate buffer pH 7.5	800 µL	700 µL
Substrate 34 µg	0 µL	100 µL
NADPH 100 µg	100 µL	100 µL

Crude Enzyme	100 µL	100 µL
	Total= 1.0 mL	Total= 1.0 mL

2.8.8.4 Polyacrylamide gel electrophoresis (PAGE)

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

2.8.8.5 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.

2.8.8.6 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.

2.8.8.7 Preparation of the sample

Equal parts of the protein sample and the loading buffer were mixed in a microcentrifuge tube and kept in a boiling water bath for 3-5 min for SDS-PAGE. Then the samples were centrifuged at 10,000 g for 10 min, supernatant collected and used for electrophoresis. Gel was run at room temperature at constant ampere 20 mA or 80 V. The protein samples for native PAGE were mixed with native dye and gel was run at 80 V at low temperature ranging from 4-10 ºC.

2.8.8.8 Loading and Running the polyacryl-amide Gel

Once well has been formed in the gel, flushed it 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 70-80 units. The run was stopped when the dye reached the bottom of the polyacryl-amide gel.

2.8.8.9 Silver staining of the gel Solutions

Protein separated on SDS or native PAGE, was transferred to the fixer solution for 1 h. This was followed by 2 x 20 min washes in 50% ethanol. The gel was transferred to thiosulfate solution for 1 min and rinsed thrice with deionized H_2O (20 s each). The gel was silver stained (HCHO was added to the solution just before use) for 20 min with intermittent shaking in dark. The gel was then rinsed thrice with deionized H_2O (20 s each) and transferred to the developer till the bands developed. The gel was washed with deionized H_2O and stored in fixer.

2.8.9 Histology and Immunocytolocalization

Free hand transverse sections were fixed overnight under vacuum in freshly prepared cold 4% buffered formaldehyde (4% paraformaldehyde in 1X PBS). The sections were

dehydrated by passages through increasing ethanol: water series (30%, 50%, 70%, 85%, 95% and 100% ethanol) for 30 min each. This was followed by passages through tertiary butanol: ethanol series (25:75, 50:50, 75:25, and 100:0). The sections were rehydrated by treating with 70% and 50% ethanol and 0.5X SSC for 2 min. The rehydrated sections were soaked in two changes of 1X PBS for 10 min each. Next, the sections were washed in 1X PBS containing 0.1% BSA for 5 min and subjected to 30 min of blocking with 10% BSA at room temperature in a humidified chamber. Post blocking washes included three washes of 15 min each with 1X PBS containing 0.1% BSA. Primary antibody incubation was carried out overnight in a humidified chamber at 4 ºC using 75 µL of diluted antibody (1 µg mL^{-1}) in 1 X PBS containing 0.1% BSA. Negative controls included either the use of pre-immune serum or the omission of both antibody and pre-immune serum. Following the primary antibody incubation, the sections were washed thrice for 15 min each in 1X PBS containing 0.1% BSA. A secondary antibody, 0.2% Anti-rabbit-IgG-goat alkaline phosphate conjugate antibody (diluted in 1X PBS with 0.1% BSA), was added to the tissue sections at this stage and incubated at 37 ºC in a humidified chamber for 2 h in dark. Post secondary antibody washes were carried out at room temperature using 1X PBS with 0.1% BSA. Color was developed in dark by addition of 60 µL of color development reagent (20 µL BCIP/NBT mix in color development buffer containing 10% polyvinyl alcohol) to the color development buffer (100 mM Tris, pH 9.5, 150 mM NaCl, 50 mM $MgCl_2$) treated slides. In present study we have got the ready to use BCIP/NBT solution therefore the volume from this solution was added sufficiently to submerge the sections in the watch glass. The slides were placed in humidified (color development buffer) chamber at RT in dark for 45 min. Upon color development, 10 mM EDTA was used to stop the reaction, rinsed with water, air dried and cover slip-mounted using glycerol and then it was observed under microscope, and microscopic image captured.

2.8.9.1 Histochemical Staining

Free hand transverse sections were prepared for histochemistry as described above. Phloroglucinol staining was done by covering pretreated slide with phloroglucinol for

half an hour and then concentrated HCl was added (Harris *et al.,* 1980; Speer *et al.,* 1987). Lignified tissue takes red color. A cover slip mounted on the specimen using glycerol was sealed to the slide using nail enamel. The slides were viewed under a light microscope and pictures captured.

2.8.9.2 GUS histochemical assay

The GUS staining solution was prepared by taking 1 mM X-Gluc (5-bromo, 4chloro, 3-indolyl-β-D-glucuronide: Cyclohexylammonium salt (X-GlcA) from a 20 mM stock made in dimethylformamide, 100 mM sodium dihydrogen phosphate dihydrate and 0.5% Tween-20. The pH of the solution was adjusted to 7.0 with 1N NaOH. The histochemical GUS assay was performed as described by Jefferson (1987) to monitor *GUS* gene expression in putative transgenic calli. The GUS assay was carried out on 10% of the randomly selected calli immediately after co-cultivation as well as 21 days after co-cultivation. Embryo axes were immersed in GUS assay solution and incubated at 28 ºC O/N. *GUS* gene expression was observed and photographed by using an Axioplan 4 microscope from Carl Zeiss.

2.8.10 *Agrobacterium* mediated tobacco transformation

The tobacco regeneration and transformation protocol is modified from Horch *et al*. (1985). *Agrobacterium tumefaciens* strain GV2260 harboring the binary plasmid vector to be transformed in tobacco was inoculated in 5mL YEP media containing rifampicin (250 mg L^{-1}) and kanamycin (50 mg L^{-1}). The culture was allowed to grow overnight at 28 ºC with shaking at 200 rpm. Next day 1 mL inoculum from this tube was taken and added to 50 mL YEP flask containing appropriate antibiotics. The culture was incubated at 28 ºC for 24 hours with shaking at 200 rpm. Fresh leaves were taken from axenic cultures of *Nicotiana tabacum* var. Anand 119 and ~ 6 mm diameter disc were punched out. The leaf discs were inoculated on MS medium agar plates containing 4.4 µM BAP and 5.37 µM NAA (MSBN) and kept for 48 h in dark. *A. tumefaciens* GV2260 cells grown for 24 h were centrifuged for 10 min at 5,000 g. Cells were suspended in 20 mL of 50 mM $MgSO_4$ and 200 µM acetosyrigone and allowed to stand for 10 min with

intermittent shaking. The tobacco leaf discs from axenic tobacco cultures were co-cultivated in this suspension for 10 min with intermittent shaking. The leaf discs were then transferred to MSBN plates without any antibiotics. The adaxial side of the leaf disc was in contact with the medium. The plates were incubated in dark at 28 ºC. After two days the leaf discs were harvested, washed with liquid MSBN and dried on sterile filter paper. Up to 10 leaf discs per plate were transferred to MSBN agar plates containing hygromycin (7.09 µM) and cefotaxime (200 µM). The cultures were incubated at 28 ºC with 18/6 hours photoperiod in diffused light (24.4 µmol/m^2 s^{-1}). Leaf discs were subcultured every week till elongated shoots were excisable. Elongated shoots were excised and transferred to root induction medium. The rooted plants were hardened and further analyzed.

2.8.11 ELISA (Enzyme-Linked Immunosorbent Assay)

Fresh tissues were collected, frozen in liquid nitrogen and crushed to a fine powder. Crude protein was extracted with 2 mL of protein extraction buffer. Total Protein was quantified using Bradford reagent. Equal amount of protein was coated on 96 well micro titer plates. Antigen (recombinant protein) was diluted in PBS to an optimal concentration and coated on 96 well microtiter plate (100 µL/ well). Plate was incubated for 2 h at room temperature or over night at 4 ºC and washed twice with PBS after incubation. Non-specific sites were blocked with blocking buffer 300 µL/ well and incubated for 1h at room temperature. After washing out the unbound blocking agent with PBS, primary antibody (Anti rabbit IgG) was added and incubated for 1 h at room temperature. The unbound primary antibody was washed twice with PBS and secondary antibody conjugated with alkaline phosphatase was added and incubated for 2 h at room temperature. Enzyme specific substrate BCIP/ NBT or PNPP was added and incubated for 45 min for the color development and the absorbance was measured at 405 nm.

Chapter 3


Isolation, Cloning and
Characterization of CAld5H gene
from *Leucaena leucocephala*

3.1 Introduction

CAld5H enzyme belongs to cytochrome P450 monoxygenase family. Cytochrome P450 proteins, named for the absorption band at 450 nm of their carbon-monoxide bound form, are one of the largest super families of enzyme proteins. Carbon monoxide can bind ferrous P450 instead of di-oxygen, inducing a shift of the maximum of absorbance of the heme (called the Soret peak) to 450 nm (Omura, T. and Sato, R., 1964). CO is bound with high affinity and prevents binding and activation of O_2. The result is an inhibition of P450 activity. P450s are the large group of heme-containing enzymes, most of which catalyze NADPH- and O_2 –dependent hydroxylation reactions. Class II P450s from plants are involved in biosynthesis or catabolism of all types of hormones, in the oxygenation of fatty acids for the synthesis of cutins, and in all of the pathways of secondary metabolism - in lignification and the synthesis of flower pigments and defense chemicals (which are also aromas, flavors, antioxidants, phyto-estrogens, anti-cancer drugs and other drugs) [Kahn R and Durst F., 2000]. The well characterized portion of the catalytic sequence involves four steps. The first step is substrate binding, with displacement of the sixth ligand solvent inducing a shift in the maximum of absorbance, spin state and redox potential of the heme protein system; the second is one-electron reduction of the complex to a ferrous state, driven by the increase in redox potential that results from the previous step; the third is binding of molecular oxygen to give a super-oxide complex; and the fourth is a second reduction step leading to an 'activated oxygen species' (Schlichting I. *et al*., 2000). In plants, two P450s catalyze reactions of the general phenylpropanoid pathway: cinnamate- 4-hydroxylase (C4H) and Ferulate 5-hydroxylase (F5H). The hydroxylation reaction in the biosynthesis of S-unit (syringyl) was first considered to occur at the ferulate level (C. Grand, 1984), and hence, the enzyme was called Ferulate 5-hydroxylase. However, more recent studies have revealed that F5H can also function at later steps in the pathway, mainly at the coniferyl aldehyde or coniferyl alcohol level (J.M. Humphreys *et al*., 1999, L. Li *et al*., 2000). This enzyme was therefore alternatively renamed coniferaldehyde- 5-hydroxylase (CAld5H) (K. Osakabe *et al*., 1999).

Although most of the animal P450s that have been studied have broader range of substrate specificity and function in the catabolism of xenobiotics, the plant P450s that have been

studied to date have anabolic roles and are highly substrate-specific (Chapple, C., 1998). F5H/CAld5H is unusual in that it is a multifunctional plant P450 with three physiologically relevant substrates. The *K*m for the substrates such as coniferaldehyde, coniferyl alcohol and for the ferulic acid are 1 μM, 3 μM and 1000 μM respectively. This study demonstrates that the coniferaldehyde is the most preferred substrate for the enzymes (J. M. Humphrey *et al.,* 1999). The present study focuses on the down-regulation of lignin in pulp producing plant such as *Leucaena*. Pulping efficiency depends on lignin reactivity (which is associated with syringyl/guaiacyl S/G monolignol ratio). Wood- pulping kinetics revealed that every unit increase in the S/G ratio would roughly doubles the rate of lignin removal (Chapple *et al*., 1992). Considerable evidence is now available that shows that in angiosperm trees, the syringyl monolignol pathway branches out from guaiacyl pathway through coniferaldehyde and is regulated in sequence by three genes encoding coniferaldehyde 5-hydroxylase (CAld5H), 5-hydroxyconiferaldehyde *O*-methyltransferase (COMT) and sinapyl alcohol dehydrogenase (SAD). It has also been revealed by enzyme kinetics studies that CAld5H has a 6-to 50- times slower turn over rate than the other two syringyl pathway enzymes (Li *et al.,* 2000), pointing to a key role for CAld5H in limiting syringyl monolignol biosynthesis and , therefore, lignin S/G ratio. Though there are many P450s protein members including CAld5H/F5H have been characterized from different plants, till date there are no studies has been done on CAld5H from *Leucaena leucocephala*. In the present study, polymerase chain reaction (PCR) based approach has been used to fish out a partial cDNA clone. Subsequently, a full-length cDNA clone of CAld5H has been isolated and characterized 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 Fishing out the coniferaldehyde 5-hydroxylase gene from *Leucaena leucocephala*

PCR based approach was followed to fish out the coniferaldehyde- 5-hydroxylase gene.

Though cDNA/gDNA library screening approach was also tried but it didn't successfully translate into isolation of gene in question. Usually, Rapid amplification of cDNA ends

(RACE) is an important technique to get the full length gene. The prerequisite for RACE reaction is to have a partial known sequence of the gene to be isolated. Therefore, it is mandatory to have a partial cDNA clone of the gene.

3.2.2 Primers Design

The CAld5H sequences available in the NCBI gene data bank were aligned and multiple sets of primers were designed from the conserved regions. Several sets of degenerate primers were also designed based on the conserved domains of the CAld5H peptide sequences.

3.2.2.1 Multiple sequence alignment of amino acid sequences of CAld5H gene

All the reported Plant CYP450 sequences were aligned using multiple sequence alignment program Clustal W 1.8. The amino acid sequences of CAld5H gene from various plants such as, *Liquidambar styraciflua* (AF139532), *Trifolium pratense* (AB236753), *Broussonetia papyrifera* (AY850934), *Arabidopsis thaliana* (NM_119790), *Oryza sativa Japonica* Group (AB207253), *Solanum lycopersicum x Solanum peruvianum* (AF150881), *Brassica napus* (DQ679758), *Camptotheca acuminata* (AY621153), *Glycine max* (DQ340235), *Medicago sativa* (DQ222912), *Populus x Canadensis* (AM921699), *Triticum monococcum* (EU099349) and *Arabidopsis thaliana* (NM_100010) were aligned. (GenBank accession numbers are given in bracket). Multiple sets of degenerate primers were designed from the most conserved domains (heme binding domain, NADPH binding domain, stop transfer signal etc.). The codon usage table (Appendix 1) for *Leucaena leucocephala* was used to select the best suited nucleotide at the wobble position of the codon, while designing degenerate primers.

3.2.2.2 Multiple sequence alignment of nucleotide sequence of selected CAld5H

The corresponding nucleotide sequences from the consensus regions of the CAld5H gene were aligned using Clustal W (1.8) multiple sequence alignment tool. The homologous

regions were manually edited and primers were designed from them. Few primer sets were also designed directly from the nucleotide sequences of some of the plant of *Fabaceae* family i.e. *Glycine max* and *Medicago sativa*, using primer 3 software (http://frodo.wi.mit.edu/primer3).

3.2.3 Nucleic acid 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 in 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

Total RNA was extracted either from *in vitro* grown seedlings or from the xylem tissue of mature plant 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 in chapter 2; section 2.8.4.8.

3.2.4.1 Amplification of partial cDNA fragments

PCR reactions were setup with cDNA as template. Primers designed from conserved regions of CAld5H gene were used.

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

RACE ready cDNA was synthesized using Smart RACE cDNA Kit (Invitrogen, USA) as described in chapter 2; section 2.8.4.10.

3.2.5.1 3’ RACE

Primary PCR was done using RACE Ready cDNA as template. Forward gene specific primer (3’RACE1) and Generacer 3’ as reverse primer were used. The resultant PCR product was diluted (1:50) and used as template for secondary PCR. To perform secondary PCR, 3’RACE2 primer was used as a forward primer and Generacer 3’nested primer as reverse primer.

3.2.5.2 5’ Rapid Amplification cDNA Ends

Primary PCR was done using RACE Ready cDNA as template. Gene specific primer 5’RACE1 was used as reverse primer and Generacer 5’ as forward primer. The resultant PCR product was diluted (1:50) and used as template for secondary PCR. To perform secondary PCR, 5’RACE2 primer was used as a nested reverse primer and Generacer 5’nested primer as forward primer. Details of PCR setup and cycling were described in chapter 2; section 2.8.4.10.

3.2.5.3 PCR amplification of full-length CAld5H cDNA clone

To get the full length CAld5H cDNA clone standard PCR reaction was done using startcal as a forward primer and stopcal as reverse reaction. As the name suggest, the primers were designed based on the start and stop region of CAld5H nucleotide sequences obtained from 5’ and 3’RACE.

3.2.6 Southern hybridization / Slot blot Analysis

Southern hybridization (Southern, E., 1975) and slot blot was done according to the protocol described in chapter 2; section 2.8.5.

3.2.7 Characterization of the *L. leucocephala* CAld5H gene using bioinformatics tools

Nucleotide and amino acid sequence analysis of the *L. leucocephala* coniferaldehyde 5-hydroxylase 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 , Signal P software (to predict the signal peptide), and to find the restriction enzyme map the webmapdna software http://pga.mgh.harvard.edu/web_apps/web_map/start was used. Phylogenetic trees are constructed after alignment of amino acid sequences and nucleotide sequences in Clustal X and Neighbor-Joining trees in Mega 4.0 software.

3.3 Results and Discussion

3.3.1 Multiple sequence alignment of amino acid sequences of reported CAld5H protein

Five amino acid sequences of CAld5H/F5H/CYP450 from *Glycine max, Medicago sativa, Liquidambar Styraciflua, Broussonetia papyrifera* and *Brassica napus* were selected from NCBI database for protein sequence alignment. These sequences were aligned using Clustal W (1.8) multiple sequence alignment program. They shares 70-85% similarity. The conserved regions selected as a target site for primer designing has been shown as red letter in figure 3.1.

```
Glycine           MD---WQSMMGNL--DPFQRTILILVPLTLLLLLLLSRTRPRPPYPPGPKGFPIIGNMFM
Medicago          MDSLLKYPIMENFKEEPFLMAIMFILPL-ILLLGLVSRILKRPRYPPGPKGLPVIGNMLM
Liquidambar       MDSS----LHEAL--QPLPMTLFFIIPL-LLLLGLVSRLRQRLPYPPGPKGLPVIGNMLM
Broussonetia      MDTKSITLLQEAL--KPLPMAIFLVIPL-LFLLGHILRSRGRLPYPPGPKGLPIIGSMSM
Brassica          MESSISQTLSQVL--DPTTGILIVVSPL--IFVGLITR-RRRPPYPPGPRGWPIIGNMSM
                  *:      :   :  .*    ::.: **  :::  : *   *  *****:* *:**.* *

Glycine           MDQLTHRGLANLAKHYGGIFHLRMGFLHMVAISDPDAARQVLQVQDNIFSNRPATIAISY
Medicago          MDQLTHRGLANLAKKYGGIFHLRMGFLHMVAISDADAARQVLQVQDNIFSNRPATVAIKY
Liquidambar       MDQLTHRGLAKLAKQYGGLFHLKMGFLHMVAVSTPDMARQVLQVQDNIFSNRPATIAISY
Broussonetia      MDQLTHRGLAALAKQYGGIFHLKMGYLHMVAISSPETARQVLQLQDNIFSNRPANIAIRY
Brassica          MDQLTHRGLANLAKKYGGLCHLRMGFLHMYAVSSPDVAKQVLQVQDSVFSNRPATIAISY
                  ********** ***:***: **:**:*** *:* .: *:****:**.:******.:** *
```

Figure 3.1. Multiple sequence alignment of CYP450s amino acid sequences from different plant species.

3.3.2 Multiple sequence alignment of nucleotide sequences of the coniferaldehyde 5-hydroxylase/ F5H/CYP450 from different plants.

The nucleotide sequences of the CAld5H gene from five different plants were aligned using Clustal W 1.8 program. The nucleotide sequences corresponding to conserved amino acid

region (YLTYDRADM and FGGTETV**)** were selected to design forward and reverse primers respectively (Red marked sequences in Figure 3.1). Though the amino acid is conserved in the selected regions (Figure 3.1), their nucleotide counterparts don't show much homology especially in case of reverse primer (Figure 3.2). In an attempt to design an unambiguous primer, the codon usage table for *Leucaena leucocephala* was used and the nucleotide bases were modified at the wobble positions (Appendix-1).

```
Glycine          --------------------GGAATTCGGCACGAGGG--CGAAATCATCCAAACACATTC
Medicago         GGCACGAGGCATTCAAACCCAAAAAACTATAGGAATATTCACAATCATATAGCTGCATAA
Liquidambar      ---TGCAAACCTGCACAAACAAAGAGAGAGAAGAAGAAAAAGGAAGAGAGGAGAGAGAGA
Brassica         ------------------------------------------------------------
Broussonetia     ------------------------------------------------------------

Glycine          CATCTATTAAGATTTTAGAGAGTGAAAA-TGGATTGGCAAAGCATGATGGGC---AACCT
Medicago         ATTAACACAAAACCATATGGATTCTCTTCTAAAATATCCAATCATGGAGAACTTTAAGGA
Liquidambar      GAGAGAGAGAGAAGCCATGGATTCTTCTCT---TCATGAAGCCTTG--CAAC---CACTA
Brassica         ------------CGCATGGAGTCTTCTATA--TCACAAACACTAAGCCAAG---TATTA
Broussonetia     ----------GCAAATATGGATACCAAA-AGTATCACCCTCCTACAAGAGGC-----CCT
                               *  **

Glycine          GGACCCATTCCAAAGAACAATCTTAATCCTCGTCCCACTAACACTACTCCTGCTGCTGTT
Medicago         AGAACCATTTCTTATGGCAATCATGTTCATCTTACCACTAA---TACTCTTGTTGGGTTT
Liquidambar      ---CCCATGACGC-----TGTTCT--TCATTATACCTTTGC---TACTCTTATTGGGCCT
Brassica         GATCCCACCACGGG----TATTCTCATCGTTGTCTCACCTC---TCATCTTCGTCGGCCT
Broussonetia     CAAGCCCTTGCCCATGGCCATCTTCTTAGTCATCCCCCTCC---TCTTCCTCCTCGGCCA
                     **    *         *  *  *  *  *  *        *  ** *  *

Glycine          ACTATCTCGTACCCGTCCAAGACCCCCCTATCCACCAGGCCCTAAGGGCTTCCCAATCAT
Medicago         AGTGTCACGAATCCTCAAAAGGCCAAGATATCCACCAGGACCAAAAGGGTTACCTGTTAT
Liquidambar      AGTATCTCGGCTTCGCCAGAGACTACCATACCCACCAGGCCCAAAAGGCTTACCGGTGAT
Brassica         CATCACACGGC---GACGAAGGCCTCCGTACCCACCCGGTCCACGTGGTTGGCCCATCAT
Broussonetia     TATCCTCCGGTCCCGGGGGAGGCTGCCGTATCCGCCGGGCCCGAAGGGCCTTCCGATCAT
                   *    **          ** *     ** ** ** ** **    **    **  * **

Glycine          AGGAAACATGTTTATGATGGACCAGCTAACCCACCGCGGCCTCGCCAACCTGGCCAAACA
Medicago         AGGTAACATGCTAATGATGGACCAATTAACCCACCGTGGTCTAGCCAACTTAGCCAAAAA
Liquidambar      CGGAAACATGCTCATGATGGATCAACTCACTCACCGAGGACTCGCCAAACTCGCCAAACA
Brassica         AGGCAATATGTCAATGATGGACCAACTCACTCACCGTGGTTTAGCCAACTTAGCTAAAAA
Broussonetia     CGGGAGCATGTCGATGATGGACCAGCTGACCCACCGCGGCCTCGCCGCCCTCGCCAAGCA
                  ** *  ***   ******** **  * ** ***** **  * ***    * ** **  *

Glycine          CTACGGCGGAATCTTCCACCTCCGCATGGGCTTCCTCCACATGGTCGCCATCTCCGACCC
Medicago         ATATGGAGGCATCTTTCACCTACGCATGGGATTCCTCCACATGGTAGCTATTTCCGACGC
Liquidambar      ATACGGCGGTCTATTCCACCTCAAGATGGGATTCTTACACATGGTGGCCGTTTCCACACC
Brassica         GTACGGTGGCCTGTGCCATCTCCGCATGGGATTTCTCCACATGTACGCCGTTTCCTCGCC
Broussonetia     GTACGGCGGCATCTTCCACCTCAAGATGGGGTACCTCCACATGGTGGCGATCTCGTCGCC
                  ** ** **  * *  ** **    ***** *   * ******   **  * **     *

Glycine          CGACGCCGCGCGACAGGTTCTCCAAGTCCAAGACAACATCTTCTCCAACCGCCCAGCCAC
Medicago         AGACGCCGCACGACAAGTTCTCCAAGTTCAAGACAACATCTTTTCCAACCGGCCAGCAAC
Liquidambar      CGACATGGCTCGCCAAGTCCTTCAAGTCCAAGACAACATCTTCTCGAACCGGCCAGCCAC
Brassica         AGACGTGGCTAAGCAAGTCCTTCAAGTCCAGGACAGCGTCTTCTCAAACCGACCAGCAAC
Broussonetia     GGAAACGGCCCGGCAGGTCCTCCAACTCCAGGACAACATCTTCTCGAACCGGCCAGCGAA
                  **    **    ** ** ** *** * ** **** * **** ** ***** ***** *
```

```
Glycine          CATCGCCATCAGCTACCTTACCTACGACCGCGCCGACATGGCTTTTGCCCACTACGGCCC
 Medicago         TGTAGCCATTAAATACCTAACTTACGACCGTGCTGACATGGCGTTCGCTCACTACGGTCC
 Liquidambar      CATAGCCATCAGCTACCTCACCTATGACCGAGCCGACATGGCCTTCGCTCACTACGGCCC
 Brassica         TATAGCTATAAGCTATTTGACTTATGACCGAGCCGACATGGCGTTTGCTCACTACGGACC
 Broussonetia     TATCGCCATCAGGTACCTGACGTACGACCGGGCGGACATGGCGTTCGCCCACTACGGCCC
                    * ** ** *  **  * ** ** ***** ** ********* ** ** ******** **

 Glycine          CTTCTGGCGCCAGATGCGGAAACTCTGCGTCATGAAGCTCTTCAGCCGCAAGCGCGCCGA
 Medicago         CTTTTGGCGCCAGATGCGAAAACTTTGCGTGATGAAGCTTTTCAGCCGCAAGCACGCAGA
 Liquidambar      GTTTTGGCGTCAGATGCGTAAACTCTGCGTCATGAAATTATTTAGCCGGAAACGAGCCGA
 Brassica         GTTTTGGAGACAGATGAGGAAAGTTTGTGTCATGAAGGTGTTTAGCCGTAAACGAGCCGA
 Broussonetia     GTTCTGGCGGCAGATGCGGAAGCTGTGCGTGATGAAGCTCTTCAGCCGCAAGCGGGCGGC
                   ** *** * ****** * **  * ** ** *****  * ** ***** ** *  ** *

 Glycine          GTCCTGGCAGTCAGTCCGCGATGAGGTCGACGCCGCCGTTCGCGCCGTCGCTAGCAGCGT
 Medicago         GTCTTGGCAATCTGTTAGAGACGAGGTTGACCATGCTATCCGAACTGTTTCGGACAACAT
 Liquidambar      GTCGTGGGAGTCGGTCCGAGACGAGGTCGACTCGGCAGTACGAGTGGTCGCGTCCAATAT
 Brassica         GTCATGGGCTTCTGTTCGTGATGAAGTGGACAAAATGATCCGGTCGGTTTCTAGTAACGT
 Broussonetia     GTCGTGGGAGTCGGTGCGCGGGGAGGTGGAGAAGACGGTCCTCTCCGTGGCGGGCAGCAC
                  *** ***   ** **  * *  ** ** **        * *     **  *    *

 Glycine          CGGAAAGCCCGTCAACATTGGAGAATTAGTGTTTAACCTCACCAAGAACATCATCTACCG
 Medicago         AGGCAACCCTGTGAACATTGGAGAACTGGTGTTCAATTTAACTAAAAACATTATATATCG
 Liquidambar      TGGGTCGACGGTGAATATCGGCGAGCTGGTTTTTGCTCTGACGAAGAATATTACTTACAG
 Brassica         TGGTAAGTCTATCAACGTTGGTGAGCAAATTTTTGCACTGACCCGAAACATAACTTACCG
 Broussonetia     CGGGACGGCGGTGAACGTGGGAGAGCTGGTGTTCAAGACGACGATGGATATAATCTACAG
                   **     *  * **  * ** **     * **       **     * ** *  **  *

 Glycine          CGCCGCGTTCGGGTCGAGTTCCCAAGAAGGCCAGGACGAGTTCATTAAAATACTGCAGGA
 Medicago         AGCGGCTTTCGGGTCGAGCTCAAGAGAAGGACAAGATGAGTTTATTGGAATATTGCAAGA
 Liquidambar      GGCGGCTTTTGGGACGATCTCGCATGAGGACCAGGACGAGTTCGTGGCCATACTGCAAGA
 Brassica         GGCAGCGTTCGGGTCAGCATGTGAAAAGGGACAAGACGAGTTCATAAGAATTTTACAAGA
 Broussonetia     AGCGGCGTTCGGGGCGAGCAACAAGGAAGGGCAGGACGAGTTCATAAGCATATTGCAGGA
                   ** ** ** *** *           * *  ** ** *****  *    **  * ** **

 Glycine          GTTCTCCAAGCTCTTTGGCGCGTTTAATATTGCGGATTTTATACCCTATCTCGGGTGCGT
 Medicago         GTTTTCCAAATTGTTTGGAGCTTTTAATATTTCCGACTTTGTACCTTGTTTTGGAGCTAT
 Liquidambar      GTTTTCGCAGCTGTTTGGTGCTTTTAATATAGCTGATTTTATCCCTTGGCTCAAATGGGT
 Brassica         GTTCTCTAAGCTTTTTGGAGCCTTCAACGTAGCGGATTTCATACCGTATTTCGGGTGGGT
 Broussonetia     GTTTTCGAAGCTGTTCGGGGCCTTCAACATGGCGGACTTCATTCCGTGGCTGACTTGGGT
                  *** **  *  * ** ** ** ** **  *  * ** **  * ** *   *       *

 Glycine          GGATCCACAAGGTTTGAACTCGAGACTCGCTAGGGCACGTGGCGCGCTCGATAGCTTCAT
 Medicago         TGACCCTCAAGGGCTTAATGCTAGGCTTGTGAAGGCTCGTAAAGATTTGGATAGTTTCAT
 Liquidambar      ---TCCTCAGGGGATTAACGTCAGGCTCAACAAGGCACGAGGGGCGCTTGATGGGTTTAT
 Brassica         CGATCCGCAAGGAATAAACAAGCGGCTCGTGAAGGCCCGTAATGACCTAGACGGATTTAT
 Broussonetia     GGATCCGCAGGGGCTGAACGGCAGGCTGAAGAGGGCGCGGAAGGCGTTGGATGCGTTTAT
                      ** ** **  * **     * **    * *** **    *   * **    ** **

 Glycine          TGATAAGATCATCGACGAGCACGTGCATAAGATGAAGAATGATAAGAGCAGTGAAATTGT
 Medicago         AGACAAAATCATAGATGAACATGTGCAGAAGAAGAAAAGTG--------------TTGT
 Liquidambar      TGACAAGATCATCGACGATCATATACAGAAG---GGGAGTAAAAAC----------TCGG
 Brassica         TGACGATATCATCGATGAACACATAAAGAAGAAAGAGAATCAAAACAGTGTTGATGCTGG
 Broussonetia     AGACAAGATAATAGACGAGCACGTGGTGAAG---AAGAATGA---------------GG
                   **  * ** ** ** ** **  *    ***      * *               *

 Glycine          TGATGG---AGAAACGGACATGGTGGATGAGTTGCTGGCGTTCTACAGCGAGGAGGCGAA
 Medicago         TGATGA---AGAAACGGATATGGTGGATGAGTTGCTTGCTTTCTATAGTGAGGAGGCTAA
 Liquidambar      AGGAG--GTTGATACTGATATGGTAGATGATTTACTTGCTTTTTACGGTGAGGAAGCCAA
 Brassica         AGATGTTGTTGATACCGATATGGTTGATGATCTTCTTGCTTTTTACAGTGAAGAGGCGAA
 Broussonetia     CGGCC-----GAGACCGACATGGTCGATGATCTTCTGGCCTTTTACAGTAACGAGGAGGC
                   *        ** ** ** ***** *****  * ** ** ** **  *  * ** *

 Glycine          GTTGAACAATGAATCGGACGATTTGCAGAACTCTATCAGACTCACTAAGGATAACATCAA
 Medicago         ATTGAATAATGAATCAGATGATTTGCATAATTCCATCAAACTTACCAAGGATAACATCAA
 Liquidambar      A---GTAAGCGAATCTGACGATCTTCAAAATTCCATCAAACTCACCAAAGACAACATCAA
 Brassica         ATTAGTGAGCGAGACAGCGGATCTTCAGAACTCCATCAAACTTACCCGTGACAATATCAA
 Broussonetia     CAAGGTGAATGACGCGGAGGATCTGCACAACGCCATTAGGCTCACTAGAGATAATATCAA
                         *  **  * *  *** * ** **  * ** *  ** **    ** ** *****
```

```
Glycine          GGCTATCATTATGG-ACGTGATGTTCGGAGGCACGGAAACGGTAGCGTCAGCGATCGAGT
Medicago         AGCCATCATAATGG-ACGTGATGTTTGGAGGAACGGAAACGGTAGCATCAGCAATCGAAT
Liquidambar      AGC---TATCATGG-ACGTAATGTTTGGAGGGACCGAAACGGTGGCGTCCGCGATTGAAT
Brassica         AGCAATCATCATGG-ACGTTATGTTTGGAGGAACGGAAACGGTAGCGTCAGCGATAGAGT
Broussonetia     GGCCATCATCATGGTAAGTACTAAGTAATTAGCTTGTCCCAATCAT-TTTGTGTGTGTGT
                  **    ** **** * **  *             *   *  *    *  *     *  *

Glycine          GGGCCATGGCGGAGCTCATGAGAAGCCCAGAAGATCAAAAGCGGGTCCAACAAGAGCTGG
Medicago         GGGCTATGGCAGAGTTAATGAAAAGCCCAGAAGATCTTAAAAAAGTCCAACAAGAACTAG
Liquidambar      GGGCCATGACGGAGCTGATGAAAAGCCCAGAAGATCTAAAGAAGGTCCAACAAGAACTCG
Brassica         GGGCATTGACTGAGTTACTACGGAGCCCAGAGGATCTAAAACGAGTCCAACAAGAACTCG
Broussonetia     GTGTTTTTTTAAATCGAAGATAAAGAATTCAAACTCTTATTTATTTTATGGTAGAGTCGC
                 * *   *     *           **     *   **  *      *      ***

Glycine          CGGATGTAGTGG--GCCTGGACCGTCGGGCCGAAGAGTCCGATTTCGAGAAACTCACTTA
Medicago         CAGAAGTTGTCG--GCCTGAGCCGACAGGTCGAAGAGCCCGATTTCGAGAAACTAACTTA
Liquidambar      CCGTGGTGGTGG--GTCTTGACCGGCGAGTCGAAGAGAAAGACTTCGAGAAGCTCACCTA
Brassica         CTGAAGTTGTCG--GACTTGACCGACGTGTGGAAGAATCAGACATCGAGAAGTTGACTTT
Broussonetia     ATGTATCTGCAATCACATAAATAAATAATAATAATCAGCCGAATTGGTGGACGTTAC--A
                   *     *        *              **      **  * * * *  * **

Glycine          TCTCAAATGTGCCCTCAAAGAGACCCTCCGCCTCCACCCTCCGATACCGCTCCTCCTCCA
Medicago         TCTGAAATGCGCTCTTAAGGAAACCCTACGTCTTCACCCACCAATTCCTTTGCTTCTTCA
Liquidambar      CTTGAAATGCGTACTGAAGGAAGTCCTTCGCCTCCACCCACCCATCCCACTCCTCCTCCA
Brassica         TTTGAAATGCACACTCAATGAAACCCTAAGGCTACACCCACCGATCCCACTCCTCCTCCA
Broussonetia     CAAAAAGGGAAAAAGAAAGAAAAGAGTGAGCATTCAAATTTGGAAA--ATTAGTCGTGAA
                     **  *       **  *     *  *  * **       *      *  *  *  *

Glycine          CGAGACGGCGGAGGACGCCACCGTCGGCGGCTACTTCGTCCCCAGGAAGGCGCGTGTTAT
Medicago         CGAAACTGCGGAAGAAGCAACGGTTAATGGTTATTTCATTCCAAAGCAAGCGCGTGTGAT
Liquidambar      CGAGACTGCCGAGGACGCCGAGGTCGGCGGCTACTACATTCCGGCGAAATCGCGGGTGAT
Brassica         CGAAACCGCAGAGGACACTGAGATCGACGGTTACTTCGTTCCCAAGAAATCTCGCGTTAT
Broussonetia     ATGAATTTTC-AATATGTATCTTTCCACTATTGACCATTATGAAAGATATTTTTAAAAAT
                     *      *  *        *       *      *      *            **



Glycine          GATCAACGCG----TGGGCCATTGGGAGGGACAAGAATTCCTGGGAGGAACCCGA-AACT
Medicago         GATAAACGCA----TGGGCTATTGGAAGAGACGCAAATTGTTGGGAAGAACCCGA-GAGT
Liquidambar      GATCAACGCG----TGCGCCATCGGCCGGGACAAGAACTCGTGGGCCGACCCAGA-TACG
Brassica         GATCAACGCG----TTTGCGATTGGACGCGACAAGAACTCTTGGGTTGATCCCGA-AACG
Broussonetia     AACCAGTAAGAAAGTAGGAAATTAGTTGGGATATGAAAATTGAAGGAGAGTCTAATGAAA
                  *  *         *  *  **  *  * **    **       *  **  *  *  *

Glycine          TTTAAGCCCGCCCGGTTCCTTAA----ACCGGGCGTGCCCGATTTCAAGGGGAGCAACTT
Medicago         TTTAAGCCATCACGGTTTTTGAA----ACCAGGTGTGCCCGATTTTAAAGGGAGTAATTT
Liquidambar      TTTAGGCCCTCCAGGTTTCTCAA----AGACGGTGTGCCCGATTTCAAAGGGAACAACTT
Brassica         TTTAGACCGTCCAGGTTTTTGGA----ACCGGGCGTACCAGATTTCAAAGGGAGTAACTT
Broussonetia     AATGATTTTTTATAGTTGAAAAATATTAAAAAGTTGACTAAGTGTTAAGGAAAACTAAAT
                   *           ***     *    *    *    *    * * ** *  *   *  *

Glycine          CGAGTTCATTCCATTCGGGTCGGGTCGAAGATCCTGCCCCGGAATGGTGTTGGGGCTCTA
Medicago         TGAGTTTATTCCGTTTGGGTCAGGACGTAGATCCTGTCCAGGTATGCAGTTGGGTTTGTA
Liquidambar      CGAGTTCATCCCATTCGGGTCAGGTCGTCGGTCTTGCCCCGGTATGCAACTCGGACTCTA
Brassica         CGAGTTTATACCATTCGGGTCGGGTCGTCGGTCGTGCCCGGGTATGCAGCTCGGGTTATA
Broussonetia     AAATAATATGT-ATATAAACACGCCATTTAACGTTATTCGCAACCGAAAAATAAAATGTA
                   *    **    *        *           *   *      *          * **

Glycine          CGCGCTTGAATTG-GCGGTGGCTCATCTTCTTCACTGCTTCACGTGGGAATTGCCAGATG
Medicago         CGCGCTTGATTTG-GCGGTAGCTCATTTACTTCATTGCTTTACTTGGGAGTTGCCGGATG
Liquidambar      CGCGCTAGAGACG-ACTGTGGCTCACCTCCTTCACTGTTTCACGTGGGAGTTGCCGGACG
Brassica         CGCGCTTGAACTA-GCCGTGGCCCATATATTACATTGCTTCACGTGGAAATTACCTGATG
Broussonetia     -ACGTTACATATGCAAAGCATGATATTTGTTTAAACAATTTTAATTCGTGATCGTATACG
                   ** *  *        *      *  *  *  *    **    *      *     * *

Glycine          GAATGAAACCAAGTGAGATGGACATGGGTGACGTGTTCGGACTCACCGCTCCAAGGTCCA
Medicago         GAATGAAACCAAGTGAGATGGATATGAGTGATGTATTTGGACTCACTGCTCCGAGAGCAA
Liquidambar      GGATGAAACCGAGTGAACTCGAGATGAATGATGTGTTTGGACTCACCGCGCCAAGAGCGA
Brassica         GCATGAAACCAAGCGAGCTTGATATGAGCGACGTGTTTGGTCTGACGGCTCCTAAAGCCA
Broussonetia     TATTAGAA---AGTTAATAGAATTTCAGGAAAATCTTAGCAGATTTTGATAGGATATTCA
                    *  **   **  *     *  *     *  * ** *        *     *     *
```

```
Glycine          CGCGACTCATTGCTGTACCAACCAAGCGCGTGGTGTGCCCTCTCTTTTAAAAATAAAAAA
Medicago         GTCGACTCATTGCTATTCCTACTAAGCGTGTCTTGTGTCCTCTGGATTAGAAAGAAAAAG
Liquidambar      TTCGACTCACCGCCGTGCCGAGTCCACGCCTTCTCTGTCCTCTCTATTGATCGAATGATT
Brassica         CGCGTCTCTACGCTGTCCCGTGCACGCGCCTGATTTGTGCTCTCTAATTC----------
Broussonetia     ATTATGCTTTTTTTTTTATCACTCAAAAAAAAAAAAGGGTCAGGGCATTAACGATATGGCG
```

Figure 3.2. Multiple sequence alignment of the coniferaldehyde 5-hydroxylase nucleotide sequences (Clustal W 1.8) from *Glycine max* (DQ340235), *Medicago sativa* (DQ222912), *Liquidambar Styraciflua* (AF139532), *Broussonetia papyrifera* (AY850934) and *Brassica napus* (DQ679758). GenBank accession numbers are given in bracket.

3.3.2.1 Forward and Reverse Primer

A set of forward and reverse primer was designed based on the above mentioned strategy.

The primer sequences are given below.

Forward primer (AgF2): **5' -TAC CTT ACC TAC GAC CGC GCC GAC ATG- 3'**

Reverse Primer (AmR5): 5' -TTT GGA GGA ACG GAA ACG GTA G- 3'

3.3.3 Genomic DNA extraction and PCR

Genomic DNA was isolated from the tender leaves of *L. leucocephala* according to the protocol described by Lodhi *et al*. (1994). A_{260}/A_{280} ratio of the isolated gDNA was 1.8 and the proper digestibility by restriction enzymes has indicated the good quality of the gDNA. The concentration of the gDNA was measured by spectrophotometer and also by comparing the band intensity with known concentration of λ-bacteriophage DNA (Figure 3.3).

50-100 ng of genomic DNA template was used in PCR reaction. Various primer combinations and cycling parameters were checked in the PCR. This resulted into no amplification in either of the conditions.

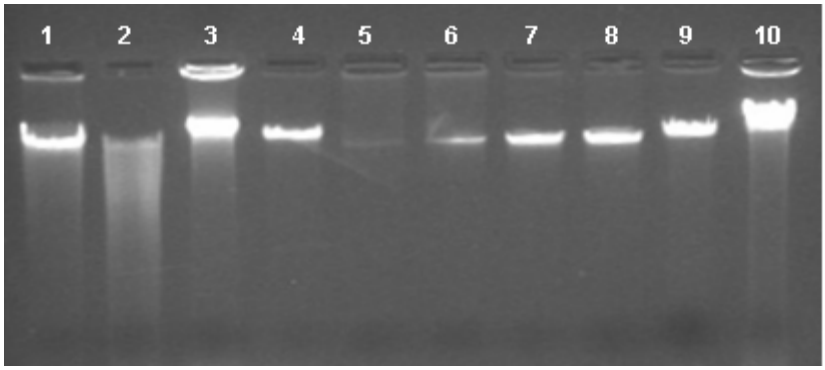


Figure 3.3. Genomic DNA of *Leucaena leucocephala* resolved on 0.7% agarose gel at 70 V. Lane 1-4 and 9-10 is gDNA. Lanes 5-8 contain different concentration of λ bacteriophage DNA with increment of 50 ng/well (from 50-200 ng).

3.3.3.1 RNA isolation, cDNA synthesis and PCR

Total RNA was isolated either from the xylem tissue of the mature plants or from the *in vitro* grown seedlings according to the protocol described in the chapter 2; section 2.8.4.5**.** The quantity and integrity of the isolates was checked by Agilent 2100 Bioanalyser system and also by agarose gel electrophoresis (Figure 3.4). First strand cDNA was synthesized using 1 µg of total RNA by AMV reverse transcriptase based ImPromII™ Reverse Transcription System (Promega, USA) as described in chapter 2; section 2.8.4.8.

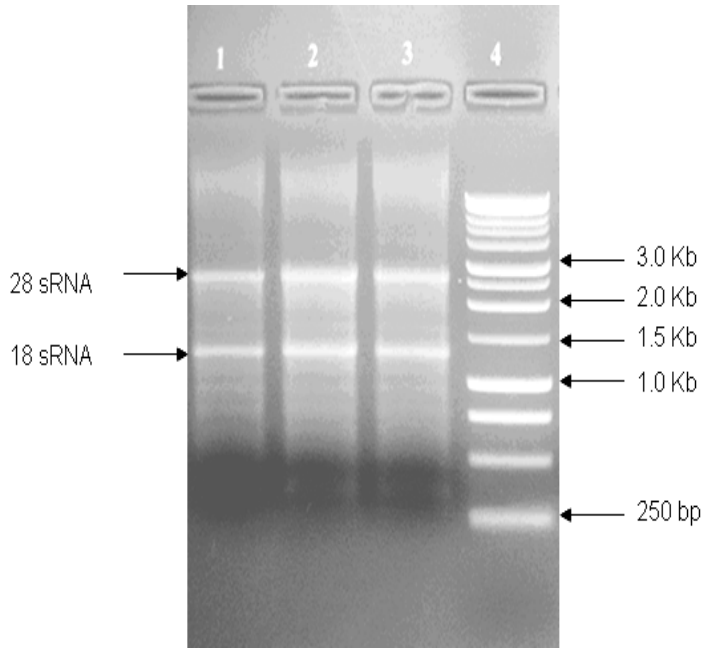


Figure 3.4. 1% denaturing agarose gel showing total RNA extracted from *Leucaena* xylem (lanes 1-3). Both 28s RNA and 18s RNA band is visible on the gel. Lane 4 is 1 Kb DNA ladder (Promega, USA). RNA was resolved at 100 V for 40 min.

3.3.3.2 PCR amplification of partial cDNA fragments of *Leucaena* CAld5H

PCR amplification was performed using *Leucaena* cDNA as template and forward primer AgF2 (5'-TACCTTACCTACGACCGCGCCGACATG-3') and reverse primer AmR5 (5'-TTTGGAGGAACGGAAACGGTAG- 3'). An amplification product of expected size i.e. approximately 600 bp was obtained (Figure 3.5). The amplified fragment was eluted from the agarose gel using gel extraction kit (Axyprep, USA) and cloned into pGEMT-Easy vector (Promega, USA). Recombinant clones were confirmed by *EcoRI* digestions which have released a~ 600 bp insert and sequenced. The sequence was subjected to NCBI nucleotide blast which shows 60- 70% sequence similarity with reported CAld5H/F5H sequences among various plants. The deduced amino acid sequence shows maximum sequence identity of 83% with CYP 84A16 of *Glycine max*. It also shows considerably very

high sequence identity (80%) with Ferulate 5-hydroxylase of *Medicago sativa* and CYP87A17 monoxygenase of *Medicago trunculata*.

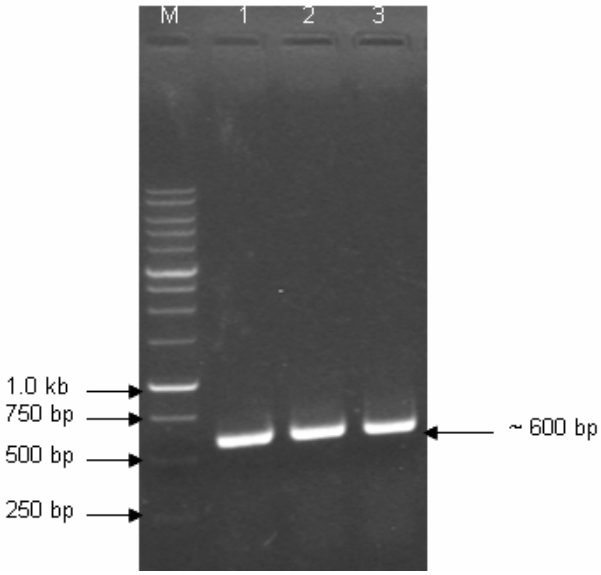


Figure 3.5. PCR amplification product of ~600 bp (lane 1-3) separated on 1% agarose gel.

Lane M is DNA size marker (1.0 Kb marker from Biozymes).

This partial clone (as it misses 5' and 3' regions from the complete gene) will be referred as LlCAld5H620 in the subsequent discussions. The sequence was submitted to NCBI GenBank database and is available with accession number DQ986906.

3.3.3.3 Nucleotide sequence of LlCAld5H620.

Forward primer (AgF2)

```
  1 TACCTTACCT ACGACCGCGC CGACATGGCT TTCGCTCACT ACGGTCCCTT
 51 CTGGCGGCAG ATGCGCAAGA TCTGCGTCAT GAAGCTCTTC AGCCGGAAGC
101 GCGCAGAGTC GTGGCAGTCG GTGCAAGATG AGGTGGAGAC GGTGGTCCGT
151 ACGGTGGGGG CGAATACCGG GAAGGAAGTG AATATCGGGG AGCTGGTGTT
201 TTCGCTGACA AAAAACATAA CGTATAGGGC GGCTTTCGGG TCGAGCTCGA
```

```
251 GGGAGGGACA GGACGAGTTC ATTGGAATAC TTCAGGAGTT CTCCAAGTTG
301 TCTGGAGCTT TCAATATAGC GGACTTTATA CCGAGTCTCA GCTGGGTTGA
351 CCCACAAGGG CTAAACGCCA GACTGGCCAA GGCTCGTGGC TCACTGGACA
401 GTTTCATAGA CAAGATCATC GACGAACACA TGCAAAACAA GAAGAAGGAA
451 GATGAAGAAG GTGACATGGT GGATGAGCTG CTGGCTTTTT ACAGCGAGGA
501 CCACTCTAAA GTCAGCGAAT CGGACGATCT GCACAACTCC ATCAAACTCA
551 CCAAGGATAA CATCAAAGCC ATCATTATGG ACGTGATGTT TGGAGGAACG
601 GAAACGGTAG
```

Reverse primer (AmR5)

Figure 3.6. Nucleotide sequence of partial length clone LlCAld5H620.

3.3.4 Slot Blot Analysis.

The *L. leucocephala* gDNA was used to perform slot blot and LlCAld5H620 clone was used as a probe as described in chapter 2: section 2.8.5. Based on signal intensity analysis it was inferred that a small LlCAld5H gene family of 2-3 members may be present in the *L. leucocephala* genome. Similar finding was reported in sweetgum by Osakabe *et al*, 1999.

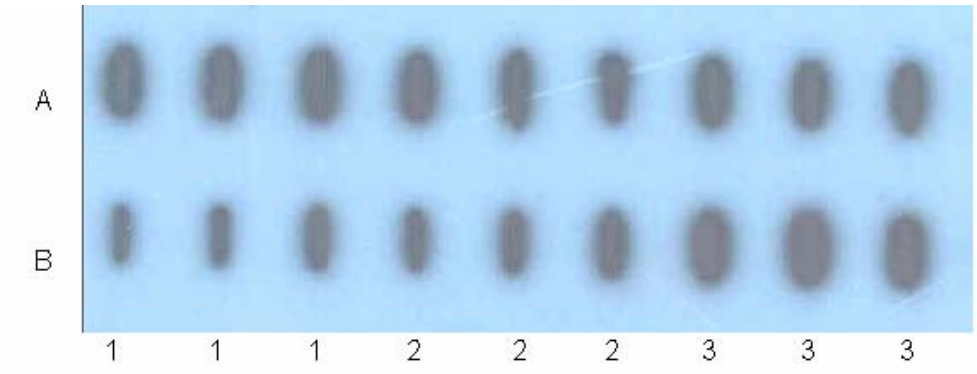


Figure 3.7. Slot blots of *L. leucocephala* genomic DNA hybridized with radiolabelled LlCAld5H620 clone. (a) Signals from 10^6 gDNA representations of *L. leucocephala*. (b) Signals from 10^6 x 1, 2 and 3 copies of the LlCAld5H clone (each blotted in triplicate).

3.3.4.1 Southern hybridization

To further validate the results from slot blot experiment and to unveil the extent of representation of CAld5H gene in the *L leucocephala* genome, Southern hybridization was performed. Good quality of genomic DNA was extracted from tender leaves of *Leucaena plant* according to the protocol described by Lodhi et al (1994). According to the standard protocols, 15- 20 µg of genomic DNA were digested with *Bst*X1, *Eco*R1, *Eco*RV, *Nco*I and, *Pvu*II restriction endonucleases. The completely digested DNA was electrophoretically separated on 0.7% agarose gel at 60 V for 8 h (Figure 3.8 a), and transferred to positively charged Hybond N^+ nylon membrane, Amersham). Hybridization was performed with a radiolabelled probe corresponding to a ~600 bp internal fragment (LlCAld5H620). Post hybridization membrane was exposed on intensifying screen. The screen was scanned for signal development using a very robust and much sensitive instrument named Typhoon Trio + scanning system.

Southern blot performed under stringent conditions revealed that from the restriction digested DNA samples multiple bands hybridized to the homologous probe used (Figure 3.8 b). The hybridization signals from very low to very high molecular weight DNA fragments indicated possible multiple gene copies in the *L. leucocephala* genome. Low molecular weight signals from *Bst*XI and *Nco*I and *Pvu*II digests suggested the presence of these restriction sites within the gene(s). As with the CAld5H gene family in *L. leucocephala* multiple *CAld5H* gene copies have been reported in sweetgum (Osakabe *et al*, 1999) and in *Populus trichocarpa* (R. sibout *et al*., 2002).

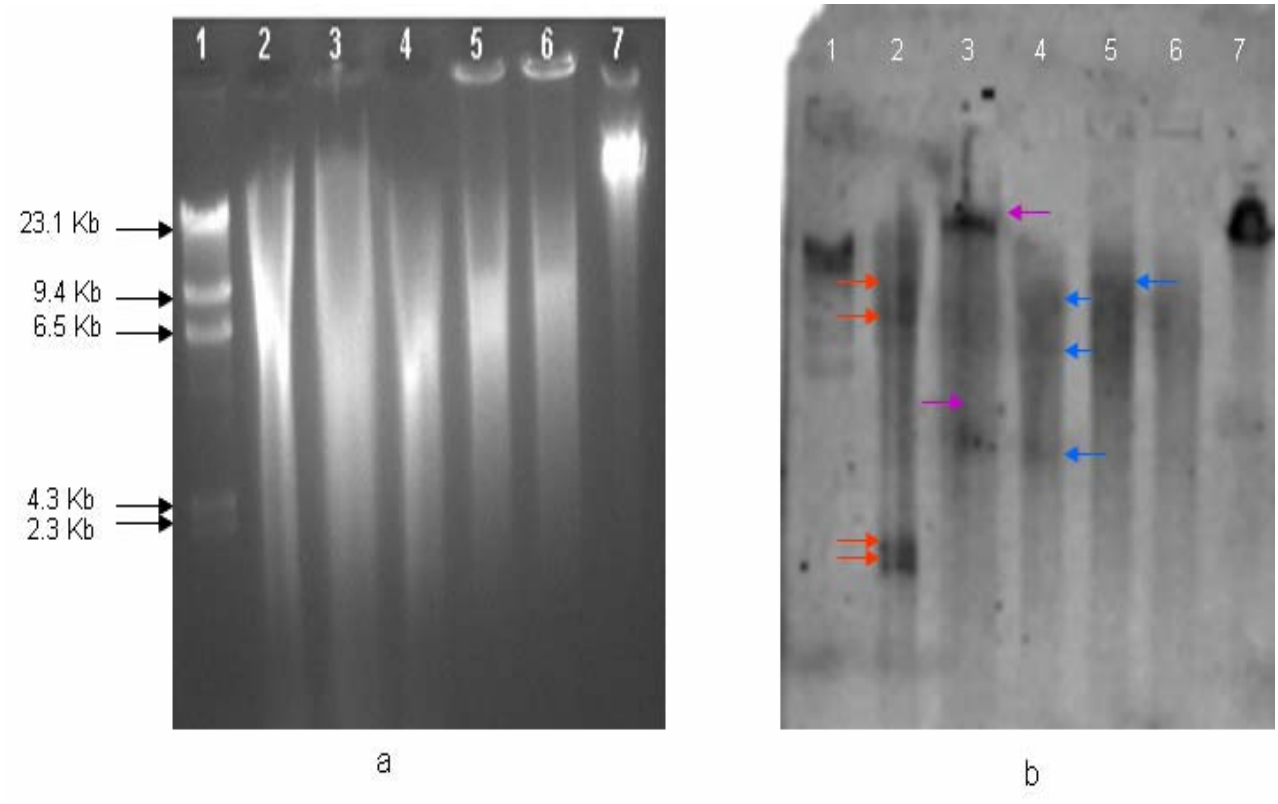


Figure 3.8. (a) 0.7% agarose gel showing *L. leucocephala* gDNA digested with *Bst*XI (lane 2), *Eco*R1 (lane 3), *Eco*RV (lane 4), *Nco*I (lane 5) and *Pvu*II (lane 6) and undigested gDNA similarly analyzed (lane 7). The λ- phage DNA digested with *Hin*dIII is used as DNA size marker (lane 1). (b) Southern hybridization of (a) using LlCAld5H620 cDNA as a probe.

3.3.5 Rapid Amplification of cDNA Ends (RACE)

3.3.5.1 Primer designing for RACE

Based on LCald5H620 sequence one 5' RACE and a 3' RACE primer as well as one set of 5' and 3' nested primers were designed as shown below underlined.

TACCTTACCTACGACCGCGCCGACATGGCTTTCGCTCACTACGGTCCCTTCTGGCGGCAGATGCGCAAGAT

CTGCGTCATGAAGCTC**TTCAGCCGGAAGCGCGCAGAGTCGT**GGCAGTCGGTGCAAGATGAGGTGGAGACGG

3'Race primer

TGGTCCGTACGGTGGGGGCGAATACCGGGAAGGAAGTGAATATCGGGGAGCTGGTGTTTTCGCTGACAAAA

AACATAACGTATAGGGCGGCTTTCGGGTCGAGCTCGAGGGAGGGACAGGACGAGTTCATTGGAATACTTCA

GGAGTTCTCCAAGTTGTCTGGAGCTTTCAATATAGCGGACTTTATAC**CGAGTCTCAGCTGGGTTGACCCAC**

5’nested primer

AAGGGCTAAACGCCAGACTGGCCAAG**GCTCGTGGCTCACTGGACAGTTTCATA**GACAAGATCATCGACGAA

3’Nesed primer

CACATGCAAAACAAGAAGAAGGAAGATGAAGAAGGTGACATGGTGGATGAGCTGCTGGCTTTTTACAGCGA

GGACCACTCTAAAGTCAGCGAATCGGACGATCT**GCACAACTCCATCAAACTCACCAAGGA**TAACATCAAAG

5’Race primer

CCATCATTATGGACGTGATGTTTGGAGGAACGGAAACGGTAG

3.3.5.1 5’ RACE

5’ Rapid amplification of cDNA ends was performed using gene specific reverse primer (5’-TCCTTGGTGAGTTTGATGGAGTTGTGC - 3’) and 5’Generacer forward primer provided with the kit as described in chapter 2; section 2.8.4.10. 5’ RACE ready cDNA was used as a template. The reaction mixture was diluted (1:50) and secondary PCR was done using nested gene specific primer (5’ Nested GeneRacer-primer) and a nested 5’ GeneRacer primer provided with the kit. Agarose gel electrophoresis has revealed an amplification product of approximately 750 bp as depicted in figure 3.9 a. The amplification product was coned into pGEMT-Easy vector (Promega, USA). The resulting construct was confirmed by *Eco*RI restriction digestion (Figure 3.9 b) and by sequencing. The nucleotide sequence was subjected to NCBI blast which shows 77% similarity with *Liquidambar styraciflua* CAld5H and 75% identities with *Campotheca acuminata* and *Populus trichocarpa* F5H nucleotide sequences. The sequence analysis shows that the RACE amplification has well extended upstream to start codon, and unveiled a 40 bp 5’UTR along with 700 bp of CAld5H gene fragment as shown in figure 3.10.

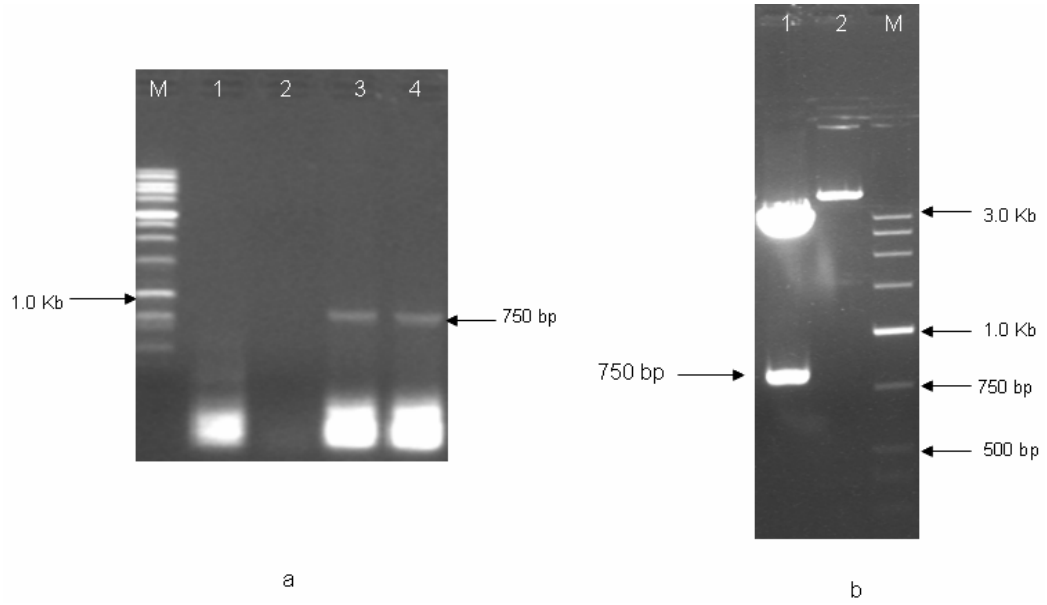


Figure 3.9. (a) 1% agarose gel showing primary RACE product (lane 1), 1:50 diluted primary race product (lane 2), Secondary RACE product of ~750 bp (lane 3 and 4). (b) *EcoRI* digested recombinant clone revealing ~750 bp fragment (lane 1) on 1% agarose gel. Lane 2 contains recombinant but undigested clone of ~ 3.75 Kb. Lane M in both the gel shows DNA size marker.

GGACACTGACATGGACTGAAGGAGTA

5'NUP

GAAAACTCCGAAGTTGCAATCTCATAACCAATACACTTAGCC

```
  1 ATGGATTCCC TTCTTCTTCA TCTCCAACCT CTTGCCATGG CCTTGTTGTT
 51 CGCAATCCCA CTCTTACTCC TGCTCGGACT CCTCTCTCGA TTCCGTAAAA
101 GGCCACCCTA CCCTCCGGGG CCCAAAGGCC TCCCCATCAT CGGCAACATG
151 GCCATGATGG ACCAGCTCAC GCACAGAGGA CTAGCCAAGC TGGCCAACCT
201 CTACGGCGGC ATACTCCACC TCCGGATGGG CTTCCTCCAC ATGGTAGCCA
251 TCTCCGACGC CGACTCCGCT CGCCAGGTGC TTCAGGTGCA CGACAACATC
301 TTCTCCAACC GCCCCGCCAC CATAGCCATC AGCTACCTCA CCTACGTCCG
351 GGCCGACATG GCTTTCGCTC ACTACGGTCC CTTCTGGCGG CAGATGCGCA
401 AGATCTGCGT CATGAAGCTC TTCAGCCGGA AGCGCGCAGA GTCGTGGCAG
451 TCGGTGCAAG ATGAGGTGGA GACGGTGGTC CGTACGGTGG GGGCGAATAC
```

```
501 CGGGAAGGAA GTGAATATCG GGGAGCTGGT GTTTTCGCTG ACAAAAAACA

551 TAACGTATAG GGCGGCTTTC GGGTCGAGCT CGAGGGAGGG ACAGGACGAG

601 TTCATTGGAA TACTTCAGGA GTTCTCCAAG TTGTTTGGAG CTTTCAATAT

651 AGCGGACTTT ATACCGAGTC TCAGCTGGGT TGACCCACAA GGGC
```

```
DEDUCED 317 AMINO ACIDS
M D S L L L H L Q P L A M A L L F A I P L L L L L G L L S R F R K R
P P Y P P G P K G L P I I G N M A M M D Q L T H R G L A K L A N L Y
G G I L H L R M G F L H M V A I S D A D S A R Q V L Q V H D N I F S
N R P A T I A I S Y L T Y V R A D M A F A H Y G P F W R Q M R K I C
V M K L F S R K R A E S W Q S V Q D E V E T V V R T V G A N T G K E
V N I G E L V F S L T K N I T Y R A A F G S S S R E G Q D E F I G I
L Q E F S K L F G A F N I A D F I P S L S W V D P Q G L N A R L A K
A R G S L D S F I D K I I D E H M Q N K K K E D E E G D M V D E L L
A F Y S E D H S K V S E S D D L H N S I K L T K D N I K A I I M D V
M F G G T E T V A I T
```

Figure 3.10. Partial cDNA sequence of CAld5H resulted from 5'RACE. The red colored region depicts 5'UTR of the gene. The reverse gene specific primer has shown by shaded region where as the 5' nested universal primer sequences has shown as blue underlined. The lower panel shows amino acid sequence deduced from corresponding nucleotides.

This partial cDNA sequence from the 5'region of the gene was aligned to the nucleotide sequences of LCAld5H620. The both sequences are very much identical in their overlapping regions (334 bp) as shown in pink letters of figure 3.12.

```
5'Race          ATGGATTCCCTTCTTCTTCATCTCCAACCTCTTGCCATGGCCTTGTTGTTCGCAATCCCA
seqence620      ------------------------------------------------------------

5'Race          CTCTTACTCCTGCTCGGACTCCTCTCTCGATTCCGTAAAAGGCCACCCTACCCTCCGGGG
seqence620      ------------------------------------------------------------

5'Race          CCCAAAGGCCTCCCCATCATCGGCAACATGGCCATGATGGACCAGCTCACGCACAGAGGA
seqence620      ------------------------------------------------------------
5'Race          CTAGCCAAGCTGGCCAACCTCTACGGCGGCATACTCCACCTCCGGATGGGCTTCCTCCAC
seqence620      ------------------------------------------------------------

5'Race          ATGGTAGCCATCTCCGACGCCGACTCCGCTCGCCAGGTGCTTCAGGTGCACGACAACATC
seqence620      ------------------------------------------------------------
```

```
5'Race          TTCTCCAACCGCCCCGCCACCATAGCCATCAGCTACCTCACCTACGTCCGGGCCGACATG
seqence620      --------------------------------TACCTTACCTACGACCGCGCCGACATG
                                                ***** ******* *** *********


5'Race          GCTTTCGCTCACTACGGTCCCTTCTGGCGGCAGATGCGCAAGATCTGCGTCATGAAGCTC
seqence620       GCTTTCGCTCACTACGGTCCCTTCTGGCGGCAGATGCGCAAGATCTGCGTCATGAAGCTC
                 ************************************************************

5'Race          TTCAGCCGGAAGCGCGCAGAGTCGTGGCAGTCGGTGCAAGATGAGGTGGAGACGGTGGTC
seqence620      TTCAGCCGGAAGCGCGCAGAGTCGTGGCAGTCGGTGCAAGATGAGGTGGAGACGGTGGTC
                 ************************************************************

5'Race          CGTACGGTGGGGGCGAATACCGGGAAGGAAGTGAATATCGGGGAGCTGGTGTTTTCGCTG
seqence620      CGTACGGTGGGGGCGAATACCGGGAAGGAAGTGAATATCGGGGAGCTGGTGTTTTCGCTG
                 ************************************************************

5'Race          ACAAAAAACATAACGTATAGGGCGGCTTTCGGGTCGAGCTCGAGGGAGGGACAGGACGAG
seqence620      ACAAAAAACATAACGTATAGGGCGGCTTTCGGGTCGAGCTCGAGGGAGGGACAGGACGAG
                 ************************************************************

5'Race          TTCATTGGAATACTTCAGGAGTTCTCCAAAGTTGTTTGGAGCTTTCAATATAGCGGACTT
seqence620      TTCATTGGAATACTTCAGGAGTTCTCCAA-GTTGTCTGGAGCTTTCAATATAGCGGACTT
                 ***************************** ***** ************************

5'Race          TATACCGAGTCTCAGCTGGGTTGACCCACAAGGC
seqence620      TATACCGAGTCTCAGCTGGGTTGACCCACAAGGGC
                 *********************************
```

Figure 3.11. Sequence alignment of by 5'RACE product with LCald5H620 clone.

These two cDNA fragments were added up to 921 bp of the CAld5H sequences after deducting the overlapping nucleotides (Figure 3.12). But it lacked the 3'region of the gene.

```
   1 ATGGATTCCC TTCTTCTTCA TCTCCAACCT CTTGCCATGG CCTTGTTGTT

  51 CGCAATCCCA CTCTTACTCC TGCTCGGACT CCTCTCTCGA TTCCGTAAAA

 101 GGCCACCCTA CCCTCCGGGG CCCAAAGGCC TCCCCATCAT CGGCAACATG

 151 GCCATGATGG ACCAGCTCAC GCACAGAGGA CTAGCCAAGC TGGCCAACCT

 201 CTACGGCGGC ATACTCCACC TCCGGATGGG CTTCCTCCAC ATGGTAGCCA

 251 TCTCCGACGC CGACTCCGCT CGCCAGGTGC TTCAGGTGCA CGACAACATC

 301 TTCTCCAACC GCCCCGCCAC CATAGCCATC AGCTACCTCA CCTACGTCCG

 351 GGCCGACATG GCTTTCGCTC ACTACGGTCC CTTCTGGCGG CAGATGCGCA

 401 AGATCTGCGT CATGAAGCTC TTCAGCCGGA AGCGCGCAGA GTCGTGGCAG

 451 TCGGTGCAAG ATGAGGTGGA GACGGTGGTC CGTACGGTGG GGGCGAATAC

 501 CGGGAAGGAA GTGAATATCG GGGAGCTGGT GTTTTCGCTG ACAAAAAACA

 551 TAACGTATAG GGCGGCTTTC GGGTCGAGCT CGAGGGAGGG ACAGGACGAG
```

601 TTCATTGGAA TACTTCAGGA GTTCTCCAAG TTGTTTGGAG CTTTCAATAT

651 AGCGGACTTT ATACCGAGTC TCAGCTGGGT TGACCCACAA GGGCTAAACG

701 CCAGACTGGC CAAGGCTCGT GGCTCACTGG ACAGTTTCAT AGACAAGATC

751 ATCGACGAAC ACATGCAAAA CAAGAAGAAG GAAGATGAAG AAGGTGACAT

801 GGTGGATGAG CTGCTGGCTT TTTACAGCGA GGACCACTCT AAAGTCAGCG

851 AATCGGACGA TCTGCACAAC TCCATCAAAC TCACCAAGGA TAACATCAAA

901 GCCATCATTA TGGACGTGAT G

Figure 3.12. A (561 bp) + B (694 bp)-(OVERLAPPING: 334bp) =921 bp: (A corresponds to LCAld5H620 clone and B is 5'RACE clone). The overlapping region between two sequences has shown as pink colored alphabets.

3.3.5.2 3' Rapid Amplification of cDNA Ends

To isolate the 3' sequence of the CAld5H gene, 3' RACE was done using a forward primer (5'-TTCAGCCGGAAGCGCGCAGAGTCGT -3') designed from LlCAld5H620 partial sequence and a reverse primer (3'GeneRacer) provided with the kit. 3' RACE ready cDNA was used as a template. An amplicon of approximately 1.4 Kb was obtained that was further verified by performing secondary PCR with 3'nested gene specific primer (5'-GCTCGTGGCTCACTGGACAGTTTCATA -3') and 3' nested universal primer (provided with kit). Both the amplification products of sizes ~ 1.4 and 1.1 Kb were individually cloned into pGEMT vector (Promega, USA). The restriction digestion of the recombinant clones have yielded insert of expected size (Figure 3.13 b). Sequences obtained from 5' RACE and 3'RACE product were independently aligned with LlCAld5H620 sequence.

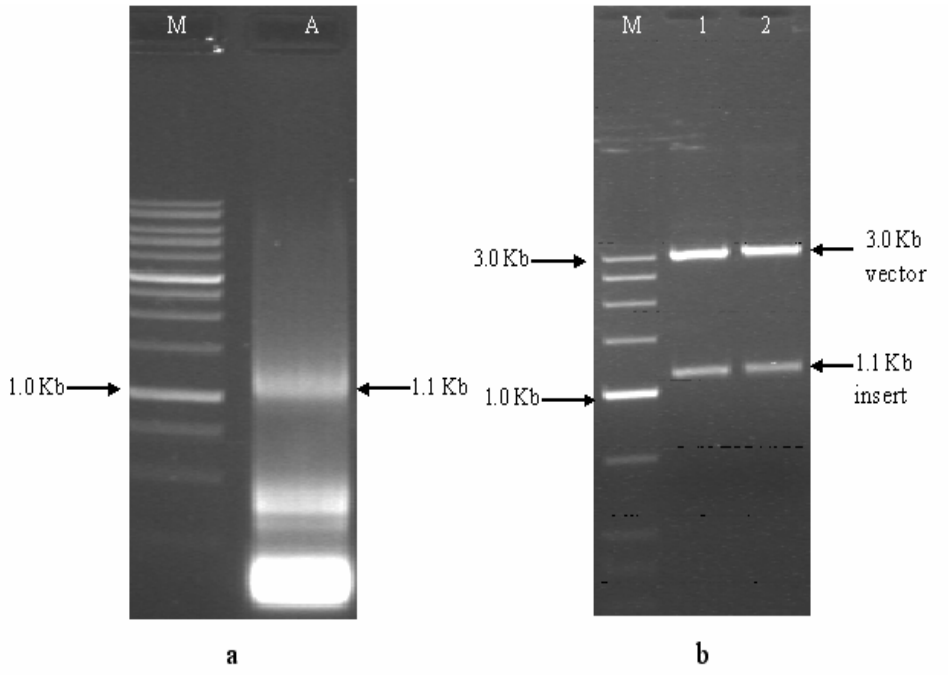


Figure 3.13. (a) 1.1 Kb amplicon shown in lane A, obtained as secondary 3' RACE products with a nested gene specific primer and nested 3' Generacer primer. (b) Restriction digestion showing 1.1 Kb insert and linearized 3.0 Kb vector backbones. Lane M in both the gels (1% agarose) is DNA size marker.

Alignment studies show that all three gene fragments eg.LlCAld5H620, 5' race product and 3' race products are absolutely identical in their overlapping regions and hence must have been amplified from same gene but not from the gene variants. These three clones were aligned and after trimming out the overlapping region the complete nucleotide sequence of CAld5H gene along with the 5' and 3' UTR was deduced (Figure 3.14A).

```
  1 TTCAGCCGGA AGCGCGCAGA GTCGTGGCAG TCGGTGCAAG ATGAGGTGGA
 51 GACGGTGGTC CGTACGGTGG GGGCGAATAC CGGGAAGGAA GTGAATATCG
101 GGGAGCTGGT GTTTTCGCTG ACAAAAAACA TAACGTATAG GGCGGCTTTC
151 GGGTCGAGCT CGAGGGAGGG ACAGGACGAG TTCATTGGAA TACTTCAGGA
201 GTTCTCCAAG TTGTTTGGAG CTTTCAATAT AGCGGACTTT ATACCGAGTC
251 TCAGCTGGGT TGACCCACAA GGGCTAAACG CCAGACTGGC CAAGGCTCGT
301 GGCTCACTGG ACAGTTTCAT AGACAAGATC ATCGACGAAC ACATGCAAAA
```

```
 351 CAAGAAGAAG GAAGATGAAG AAGGTGACAT GGTGGATGAG CTGCTGGCTT
 401 TTTACAGCGA GGACCACTCT AAAGTCAGCG AATCGGACGA TCTGCACAAC
 451 TCCATCAAAC TCACCAAGGA TAACATCAAA GCCATCATTA TGGACGTGAT
 501 GTTCGGAGGG ACGGAGACGG TGGCGTCGGC AATAGAGTGG GCCATGGCAG
 551 AACTGATGAG AAGCCCAGAA GATCTGAAGA GGGTCCAACG GGAACTGGCC
 601 GAGGTGGTGG GTCTTGATCG GAGAGTGCAA GAGTCCGACA TGGAGAAGCT
 651 GACTTACCTG AAATGCGCTC TGAAGGAGAC TCTCCGCCTC CACCCACCGA
 701 TCCCCTTGCT ACTCCACGAG ACCGCGGAAG ACGCCGTGGT TTCGGGCTAT
 751 TTCGTTCCGA AGAAGTCACG CGTGATGATC AACGCCTGGG CCATCGGGAG
 801 GGACCGCAAC GCCTGGGAGG ACCCCGACAG CTTCAAGCCG GCCCGCTTCC
 851 TCGGGGAGGG CGTCCCAGAT TTCAAAGGCA GCAACTTCGA GTTCATCCCG
 901 TTCGGGTCGG GTCGAAGGTC CTGCCCGGGG ATGCAACTGG GCCTCTACGC
 951 CTTGGAAATG GCCGTCGCTC ACCTTCTCCA CTGTTTCTCG TGGGAGTTGC
1001 CGGACGGCAT GAAGCCGAGT GAGATGGACA TGAGCGACGT CTTCGGACTC
1051 ACCGCTCCCA GGGCGAGTAG ACTTGTCGCG GTCTCCAGGA AGCGCTTGGT
1101 CTGCCCTCTC TTTTAA
```

3′ UTR of CAld5H gene

```
   1 GGGCAGGGAG CATGGCACGT AACGTGAGCG AATGAGACAG GTCTGGTTTC
  51 TGTGTTTTAT TATTTTTTGT AGAAATAAAT TATATTTCAA TTATTCATAA
 101 ATATGAAGGG AAACGTTCGA AGAAAGAAGA AAGAAAGAAT GGAAACGGCA
 151 ATAGCTGGAT AGGTGGCAAA ATAAGGAGAG TCGTGATGTG GTGTGGATAA
 201 ACTTTGTATG AATTCCTACT GTAATAATCT CCCCTCCTTT TGGGTTCAAA
 251 AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA CACT
```

Figure 3.14 A. Nucleotide sequence obtained as 3’ RACE product. The lower panel represents the sequences from 3’ untranslated region of the gene.

```
-42 gaaaactccg aagttgcaat ctcataacca atacacttag cc
  1 ATGGATTCCC TTCTTCTTCA TCTCCAACCT CTTGCCATGG CCTTGTTGTT
 51 CGCAATCCCA CTCTTACTCC TGCTCGGACT CCTCTCTCGA TTCCGTAAAA
101 GGCCACCCTA CCCTCCGGGG CCCAAAGGCC TCCCCATCAT CGGCAACATG
151 GCCATGATGG ACCAGCTCAC GCACAGAGGA CTAGCCAAGC TGGCCAACCT
201 CTACGGCGGC ATACTCCACC TCCGGATGGG CTTCCTCCAC ATGGTAGCCA
251 TCTCCGACGC CGACTCCGCT CGCCAGGTGC TTCAGGTGCA CGACAACATC
301 TTCTCCAACC GCCCCGCCAC CATAGCCATC AGCTACCTCA CCTACGTCCG
351 GGCCGACATG GCTTTCGCTC ACTACGGTCC CTTCTGGCGG CAGATGCGCA
401 AGATCTGCGT CATGAAGCTC TTCAGCCGGA AGCGCGCAGA GTCGTGGCAG
451 TCGGTGCAAG ATGAGGTGGA GACGGTGGTC CGTACGGTGG GGGCGAATAC
501 CGGGAAGGAA GTGAATATCG GGGAGCTGGT GTTTTCGCTG ACAAAAAACA
551 TAACGTATAG GGCGGCTTTC GGGTCGAGCT CGAGGGAGGG ACAGGACGAG
601 TTCATTGGAA TACTTCAGGA GTTCTCCAAG TTGTTTGGAG CTTTCAATAT
651 AGCGGACTTT ATACCGAGTC TCAGCTGGGT TGACCCACAA GGGCTAAACG
701 CCAGACTGGC CAAGGCTCGT GGCTCACTGG ACAGTTTCAT AGACAAGATC
751 ATCGACGAAC ACATGCAAAA CAAGAAGAAG GAAGATGAAG AAGGTGACAT
801 GGTGGATGAG CTGCTGGCTT TTTACAGCGA GGACCACTCT AAAGTCAGCG
851 AATCGGACGA TCTGCACAAC TCCATCAAAC TCACCAAGGA TAACATCAAA
901 GCCATCATTA TGGACGTGAT GTTCGGAGGG ACGGAGACGG TGGCGTCGGC
951 AATAGAGTGG GCCATGGCAG AACTGATGAG AAGCCCAGAA GATCTGAAGA
1001 GGGTCCAACG GGAACTGGCC GAGGTGGTGG GTCTTGATCG GAGAGTGCAA
1051 GAGTCCGACA TGGAGAAGCT GACTTACCTG AAATGCGCTC TGAAGGAGAC
1101 TCTCCGCCTC CACCCACCGA TCCCCTTGCT ACTCCACGAG ACCGCGGAAG
1151 ACGCCGTGGT TTCGGGCTAT TTCGTTCCGA AGAAGTCACG CGTGATGATC
1201 AACGCCTGGG CCATCGGGAG GGACCGCAAC GCCTGGGAGG ACCCCGACAG
1251 CTTCAAGCCG GCCCGCTTCC TCGGGGAGGG CGTCCCAGAT TTCAAAGGCA
1301 GCAACTTCGA GTTCATCCCG TTCGGGTCGG GTCGAAGGTC CTGCCCGGGG
```

```
1351 ATGCAACTGG GCCTCTACGC CTTGGAAATG GCCGTCGCTC ACCTTCTCCA
1401 CTGTTTCTCG TGGGAGTTGC CGGACGGCAT GAAGCCGAGT GAGATGGACA
1451 TGAGCGACGT CTTCGGACTC ACCGCTCCCA GGGCGAGTAG ACTTGTCGCG
1501 GTCTCCAGGA AGCGCTTGGT CTGCCCTCTC TTTTAAgggc agggagcatg
1551 gcacgtaacg tgagcgaatg agacaggtct ggtttctgtg ttttattatt
1601 ttttgtagaa ataaattata tttcaattat tcataaatat gaagggaaac
1651 gttcgaagaa agaagaaaga aagaatggaa acggcaatag ctggataggt
1701 ggcaaaataa ggagagtcgt gatgtggtgt ggataaactt tgtatgaatt
1751 cctactgtaa taatctcccc tccttttggg ttcaaaaaaa aaaaaaaaaa
1801 aaaaaaaaaa aaaaaaaaaa aaaaaa
```

Figure 3.14 B. Deduced nucleotide sequence of complete LlCAld5H gene. 5' and 3' UTRs sequences have shown in lower case. The overlapping sequences of 5'and 3'RACE products have shown in pink letters. Stop codon has represented with red letters. Underlined letters in 3' UTR represents primary cleavage site selection in pre- mRNA (CP Joshi, 1987) and the shaded nucleotides are the poly-A tail.

3.3.6 Isolation of Full-length cDNA of the CAld5H from *Leucaena leucocephala.*

To isolate complete open reading frame of coniferaldehyde 5-hydroxylase gene in one clone from the two independently existing clones (5' and 3' race clones), a set of forward primer starting from ATG and reverse primer from the end of the coding sequence was designed. The forward and reverse primers were named as startcal (5'-ATGGATTCCCTTCTTCTTCATCTC- 3') and stopcal (5' –AAAGAGAGGGCAGAC CAAGCGCTTCC- 3') respectively. *Leucaena* xylem cDNA was used as template and PCR was done. An amplicon of ~1.5 Kb was obtained as expected full length CAld5H gene (Figure 3.15 a). The amplicon was excised from the agarose gel and purified by Axygen purification kit (according to manufacturer's instruction).

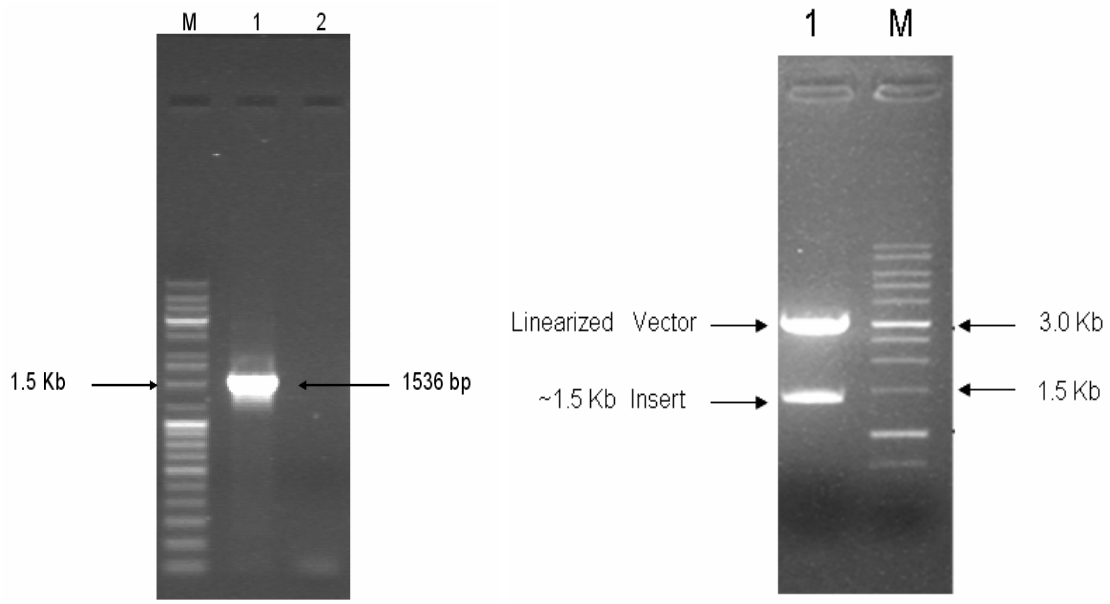


Figure 3.15. (a) Full length LlCAld5H amplicon migrated at ~1.5 Kb on 1% agarose gel (lane 1). Lane 2 is no template control. (b) 1% agarose gel showing restriction digest of positive clones released a ~1.5 Kb insert as a plausible full length candidate gene. Lane M in both the figures are medium range DNA ruler from Bangalore Genei.

The purified product was cloned into pGEMT-easy vector (Promega, USA). The recombinant clone was confirmed by restriction digestion which released a ~1.5 Kb fragment (Figure 3.15 b). The clone referred as LlCAld5H and was sequenced. The sequence shows maximum homology (83%) with CAld5H gene of *Glycine max* and significance identity with various plants P450s gene as well.

3.3.7 Characterization of LlCAld5H cDNA sequence

The full length cDNA sequence of LlCAld5H gene was submitted to NCBI GenBank database under accession no. EU041752. The *Leucaena* CAld5H cDNA sequence contains a 1536 bp open-reading frame that encodes a protein of 511 amino acids with a molecular weight 57.3 kDa. The putative initiation codon ATG is flanked by a G at + 4, in keeping

with the nucleotide commonly found flanking the initiator methionine in plants (Lutcke *et al.,* 1987). The sequence analysis shows that race amplification has been well extended downstream to the stop codon (TAA) and has yielded a 290 bp 3'UTR. Upstream to poly-A tail an ubiquitous polyadenylation signal sequence AATAA was present in the 3'UTR (Figure 3.14) which is present in most eukaryotes and forms complex with U4 SnRNP for primary cleavage site selection in pre- mRNA (CP Joshi, 1987).

```
   1 ATGGATTCCC TTCTTCTTCA TCTCCAACCT CTTGCCATGG CCTTGTTGTT
  51 CGCAATCCCA CTCTTACTCC TGCTCGGACT CCTCTCTCGA TTCCGTAAAA
 101 GGCCACCCTA CCCTCCGGGG CCCAAAGGCC TCCCCATCAT CGGCAACATG
 151 GCCATGATGG ACCAGCTCAC GCACAGAGGA CTAGCCAAGC TGGCCAACCT
 201 CTACGGCGGC ATACTCCACC TCCGGATGGG CTTCCTCCAC ATGGTAGCCA
 251 TCTCCGACGC CGACTCCGCT CGCCAGGTGC TTCAGGTGCA CGACAACATC
 301 TTCTCCAACC GCCCCGCCAC CATAGCCATC AGCTACCTCA CCTACGTCCG
 351 GGCCGACATG GCTTTCGCTC ACTACGGTCC CTTCTGGCGG CAGATGCGCA
 401 AGATCTGCGT CATGAAGCTC TTCAGCCGGA AGCGCGCAGA GTCGTGGCAG
 451 TCGGTGCAAG ATGAGGTGGA GACGGTGGTC CGTACGGTGG GGGCGAATAC
 501 CGGGAAGGAA GTGAATATCG GGGAGCTGGT GTTTTCGCTG ACAAAAAACA
 551 TAACGTATAG GGCGGCTTTC GGGTCGAGCT CGAGGGAGGG ACAGGACGAG
 601 TTCATTGGAA TACTTCAGGA GTTCTCCAAG TTGTTTGGAG CTTTCAATAT
 651 AGCGGACTTT ATACCGAGTC TCAGCTGGGT TGACCCACAA GGGCTAAACG
 701 CCAGACTGGC CAAGGCTCGT GGCTCACTGG ACAGTTTCAT AGACAAGATC
 751 ATCGACGAAC ACATGCAAAA CAAGAAGAAG GAAGATGAAG AAGGTGACAT
 801 GGTGGATGAG CTGCTGGCTT TTTACAGCGA GGACCACTCT AAAGTCAGCG
 851 AATCGGACGA TCTGCACAAC TCCATCAAAC TCACCAAGGA TAACATCAAA
 901 GCCATCATTA TGGACGTGAT GTTCGGAGGG ACGGAGACGG TGGCGTCGGC
 951 AATAGAGTGG GCCATGGCAG AACTGATGAG AAGCCCAGAA GATCTGAAGA
1001 GGGTCCAACG GGAACTGGCC GAGGTGGTGG GTCTTGATCG GAGAGTGCAA
1051 GAGTCCGACA TGGAGAAGCT GACTTACCTG AAATGCGCTC TGAAGGAGAC
1101 TCTCCGCCTC CACCCACCGA TCCCCTTGCT ACTCCACGAG ACCGCGGAAG
1151 ACGCCGTGGT TTCGGGCTAT TTCGTTCCGA AGAAGTCACG CGTGATGATC
```

```
1201 AACGCCTGGG CCATCGGGAG GGACCGCAAC GCCTGGGAGG ACCCCGACAG
1251 CTTCAAGCCG GCCCGCTTCC TCGGGGAGGG CGTCCCAGAT TTCAAAGGCA
1301 GCAACTTCGA GTTCATCCCG TTCGGGTCGG GTCGAAGGTC CTGCCCGGGG
1351 ATGCAACTGG GCCTCTACGC CTTGGAAATG GCCGTCGCTC ACCTTCTCCA
1401 CTGTTTCTCG TGGGAGTTGC CGGACGGCAT GAAGCCGAGT GAGATGGACA
1451 TGAGCGACGT CTTCGGACTC ACCGCTCCCA GGGCGAGTAG ACTTGTCGCG
1501 GTCTCCAGGA AGCGCTTGGT CTGCCCTCTC TTTTAA
```

Figure 3.16. Nucleotide sequence of LlCAld5H, showing complete open reading frame.

3.3.7.1 Restriction map of LlCAld5H

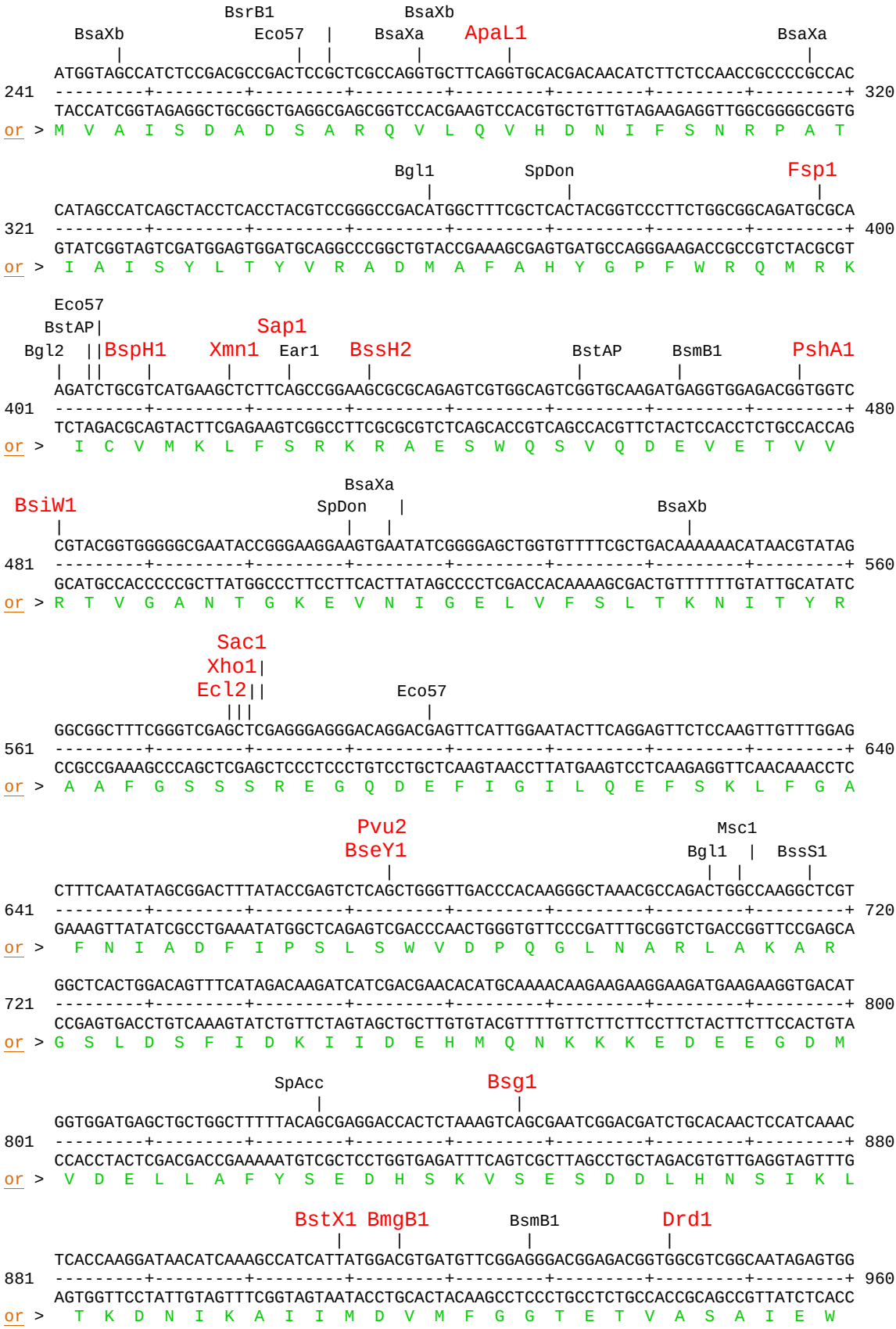
Restriction map of LCAld5H gene was created by using online bioinformatics software named Webmap DNA, http://pga.mgh.harvard.edu/web_apps/web_map/start and edited at http://pga.mgh.harvard.edu/web_apps/dna_utilities.html. The restriction map created through this tool is very robust and comparatively versatile in the sense that one can spot the restriction site on the double stranded DNA. Moreover, it also provides the translated amino acid sequence. At the end of the pages the list of unique restriction enzymes as well as the enzymes found in the sequence will be displayed as the output data file.

```
*** Sites protected by E. coli methylation are not shown in this map ***

                                    Nco1          BsaXb                BseR1
                                      |             |                    |
     ATGGATTCCCTTCTTCTTCATCTCCAACCTCTTGCCATGGCCTTGTTGTTCGCAATCCCACTCTTACTCCTGCTCGGACT
  1  ---------+---------+---------+---------+---------+---------+---------+---------+ 80
     TACCTAAGGGAAGAAGAAGTAGAGGTTGGAGAACGGTACCGGAACAACAAGCGTTAGGGTGAGAATGAGGACGAGCCTGA
or > M  D  S  L  L  L  H  L  Q  P  L  A  M  A  L  L  F  A  I  P  L  L  L  L  L  G  L

                                        Apa1
BsaXa                                 PspOM   |  Stu1                  Msc1
    |                                    |    |    |                    |
     CCTCTCTCGATTCCGTAAAAGGCCACCCTACCCTCCGGGGCCCAAAGGCCTCCCCATCATCGGCAACATGGCCATGATGG
 81  ---------+---------+---------+---------+---------+---------+---------+---------+ 160
     GGAGAGAGCTAAGGCATTTTCCGGTGGGATGGGAGGCCCCGGGTTTCCGGAGGGGTAGTAGCCGTTGTACCGGTACTACC
or >  L  S  R  F  R  K  R  P  P  Y  P  P  G  P  K  G  L  P  I  I  G  N  M  A  M  M  D

                                  Msc1
                                Bgl1  |     BsaXb            BspE1       BsaXa
                                  |   |       |                 |          |
     ACCAGCTCACGCACAGAGGACTAGCCAAGCTGGCCAACCTCTACGGCGGCATACTCCACCTCCGGATGGGCTTCCTCCAC
161  ---------+---------+---------+---------+---------+---------+---------+---------+ 240
     TGGTCGAGTGCGTGTCTCCTGATCGGTTCGACCGGTTGGAGATGCCGCCGTATGAGGTGGAGGCCTACCCGAAGGAGGTG
or >   Q  L  T  H  R  G  L  A  K  L  A  N  L  Y  G  G  I  L  H  L  R  M  G  F  L  H
```


BsrB1
BsaXb
BsaXb
Eco57
BsaXa
ApaL1
BsaXa
ATGGTAGCCATCTCCGACGCCGACTCCGCTCGCCAGGTGCTTCAGGTGCACGACAACATCTTCTCCAACCGCCCCGCCAC
241
320
TACCATCGGTAGAGGCTGCGGCTGAGGCGAGCGGTCCACGAAGTCCACGTGCTGTTGTAGAAGAGGTTGGCGGGGCGGTG
or > M V A I S D A D S A R Q V L Q V H D N I F S N R P A T
Bgl1
SpDon
Fsp1
CATAGCCATCAGCTACCTCACCTACGTCCGGGCCGACATGGCTTTCGCTCACTACGGTCCCTTCTGGCGGCAGATGCGCA
321
400
GTATCGGTAGTCGATGGAGTGGATGCAGGCCCGGCTGTACCGAAAGCGAGTGATGCCAGGGAAGACCGCCGTCTACGCGT
or > I A I S Y L T Y V R A D M A F A H Y G P F W R Q M R K
Eco57
BstAP
Sap1
Bgl2
BspH1
Xmn1
Ear1
BssH2
BstAP
BsmB1
PshA1
AGATCTGCGTCATGAAGCTCTTCAGCCGGAAGCGCGCAGAGTCGTGGCAGTCGGTGCAAGATGAGGTGGAGACGGTGGTC
401
480
TCTAGACGCAGTACTTCGAGAAGTCGGCCTTCGCGCGTCTCAGCACCGTCAGCCACGTTCTACTCCACCTCTGCCACCAG
or > I C V M K L F S R K R A E S W Q S V Q D E V E T V V
BsaXa
BsiW1
SpDon
BsaXb
CGTACGGTGGGGGCGAATACCGGGAAGGAAGTGAATATCGGGGAGCTGGTGTTTTCGCTGACAAAAAACATAACGTATAG
481
560
GCATGCCACCCCCGCTTATGGCCCTTCCTTCACTTATAGCCCCTCGACCACAAAAGCGACTGTTTTTTGTATTGCATATC
or > R T V G A N T G K E V N I G E L V F S L T K N I T Y R
Sac1
Xho1
Ecl2
Eco57
GGCGGCTTTCGGGTCGAGCTCGAGGGAGGGACAGGACGAGTTCATTGGAATACTTCAGGAGTTCTCCAAGTTGTTTGGAG
561
640
CCGCCGAAAGCCCAGCTCGAGCTCCCTCCCTGTCCTGCTCAAGTAACCTTATGAAGTCCTCAAGAGGTTCAACAAACCTC
or > A A F G S S S R E G Q D E F I G I L Q E F S K L F G A
Pvu2
BseY1
Msc1
Bgl1
BssS1
CTTTCAATATAGCGGACTTTATACCGAGTCTCAGCTGGGTTGACCCACAAGGGCTAAACGCCAGACTGGCCAAGGCTCGT
641
720
GAAAGTTATATCGCCTGAAATATGGCTCAGAGTCGACCCAACTGGGTGTTCCCGATTTGCGGTCTGACCGGTTCCGAGCA
or > F N I A D F I P S L S W V D P Q G L N A R L A K A R
GGCTCACTGGACAGTTTCATAGACAAGATCATCGACGAACACATGCAAAACAAGAAGAAGGAAGATGAAGAAGGTGACAT
721
800
CCGAGTGACCTGTCAAAGTATCTGTTCTAGTAGCTGCTTGTGTACGTTTTGTTCTTCTTCCTTCTACTTCTTCCACTGTA
or > G S L D S F I D K I I D E H M Q N K K K E D E E G D M
SpAcc
Bsg1
GGTGGATGAGCTGCTGGCTTTTTACAGCGAGGACCACTCTAAAGTCAGCGAATCGGACGATCTGCACAACTCCATCAAAC
801
880
CCACCTACTCGACGACCGAAAAATGTCGCTCCTGGTGAGATTTCAGTCGCTTAGCCTGCTAGACGTGTTGAGGTAGTTTG
or > V D E L L A F Y S E D H S K V S E S D D L H N S I K L
BstX1
BmgB1
BsmB1
Drd1
TCACCAAGGATAACATCAAAGCCATCATTATGGACGTGATGTTCGGAGGGACGGAGACGGTGGCGTCGGCAATAGAGTGG
881
960
AGTGGTTCCTATTGTAGTTTCGGTAGTAATACCTGCACTACAAGCCTCCCTGCCTCTGCCACCGCAGCCGTTATCTCACC
or > T K D N I K A I I M D V M F G G T E T V A S A I E W

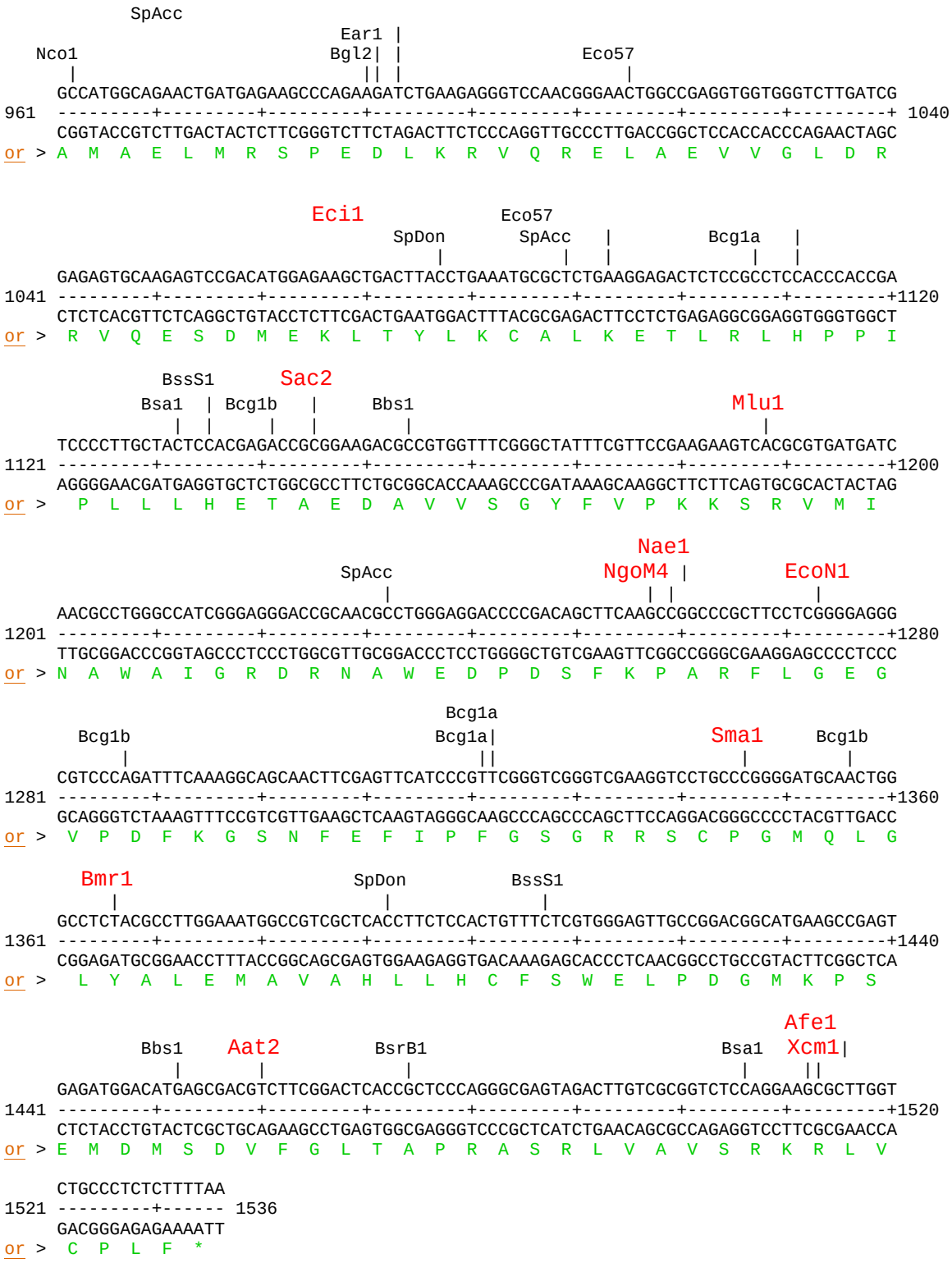


Figure 3.17. Restriction map analysis of LlCAld5H coding sequence. The unique enzyme is shown in red color. The translated amino acids are shown in green.

3.3.7.2 Multiple sequence alignment of nucleotide sequences of selected plant CPP450s

The *Leucaena* coniferaldehyde-5 –hydroxylase (EU041752) nucleotide sequence was used for homology search using BLASTN program (available at www.ncbi.nlm.nih) against somewhat similar sequences available in database. Eleven nucleotide sequences of coniferaldehyde 5 –hydroxylase / Ferulate 5-hydroxylase /Cytochrome P450 monoxygenase namely, *Glycine max* (DQ340235), Medicago sativa (DQ222912), *Liquidambar styraciflua* (AF139532), *Brassica napus* (DQ679758), *Trifolium pratense* (AB236753), *Arabidopsis thaliana* (NM_119790), *Broussonetia papyrifera* (AY850934), *Solanum lycopersicum x Solanum peruvianum* (AF150881), *Oryza sativa Japonica* Group (AB207253), Arabidopsis thaliana (NM_100010), were selected for multiple sequence alignment. These sequences have 70-75 % identity and approximately 90% query coverage with the *Leucaena* CAld5H (EU041752) nucleotide sequence. They are used for multiple sequence alignment using program Clustal X. The conserved nucleotide sequences are shown in figure 3.18.

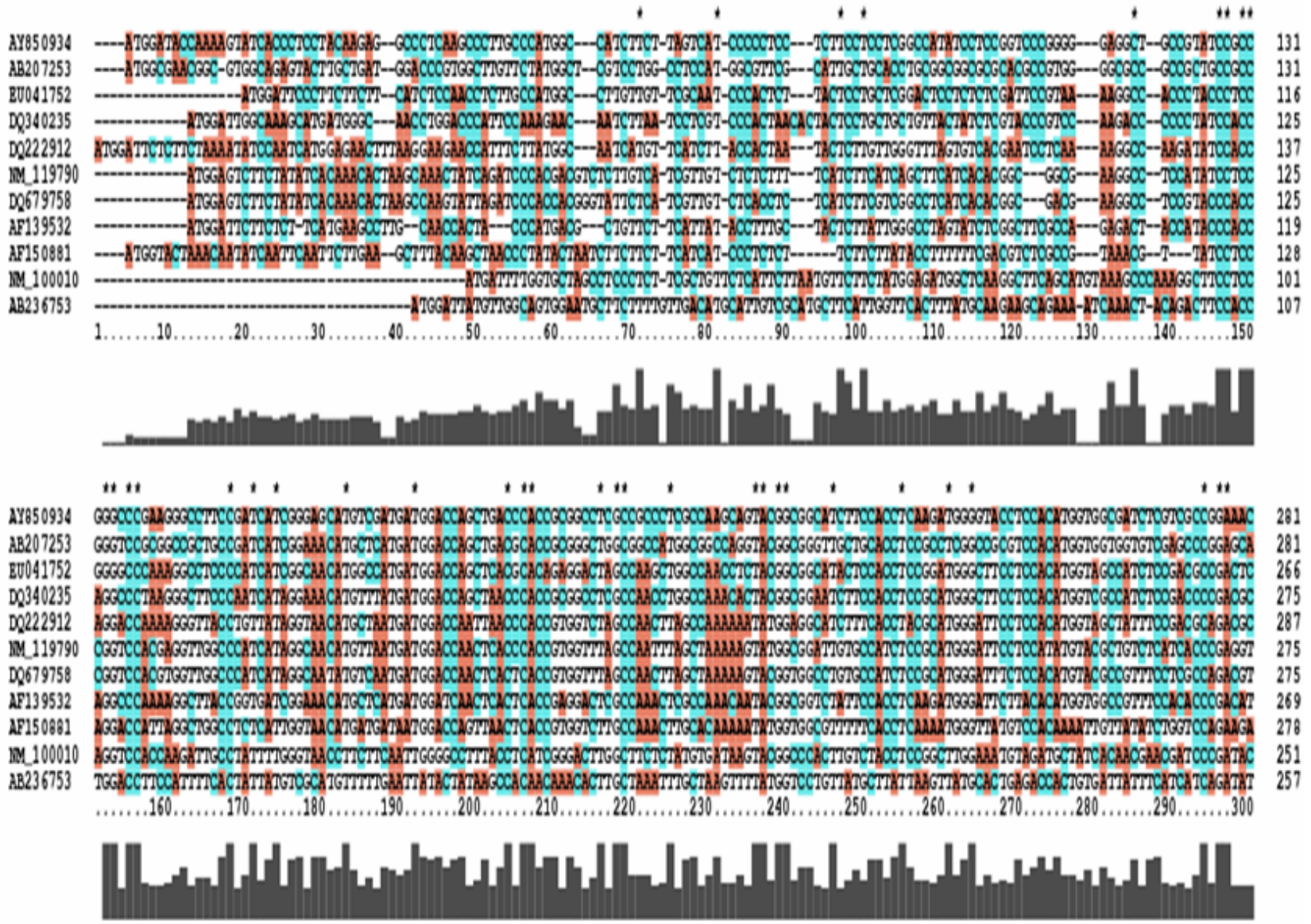


Fig 3.18: continued

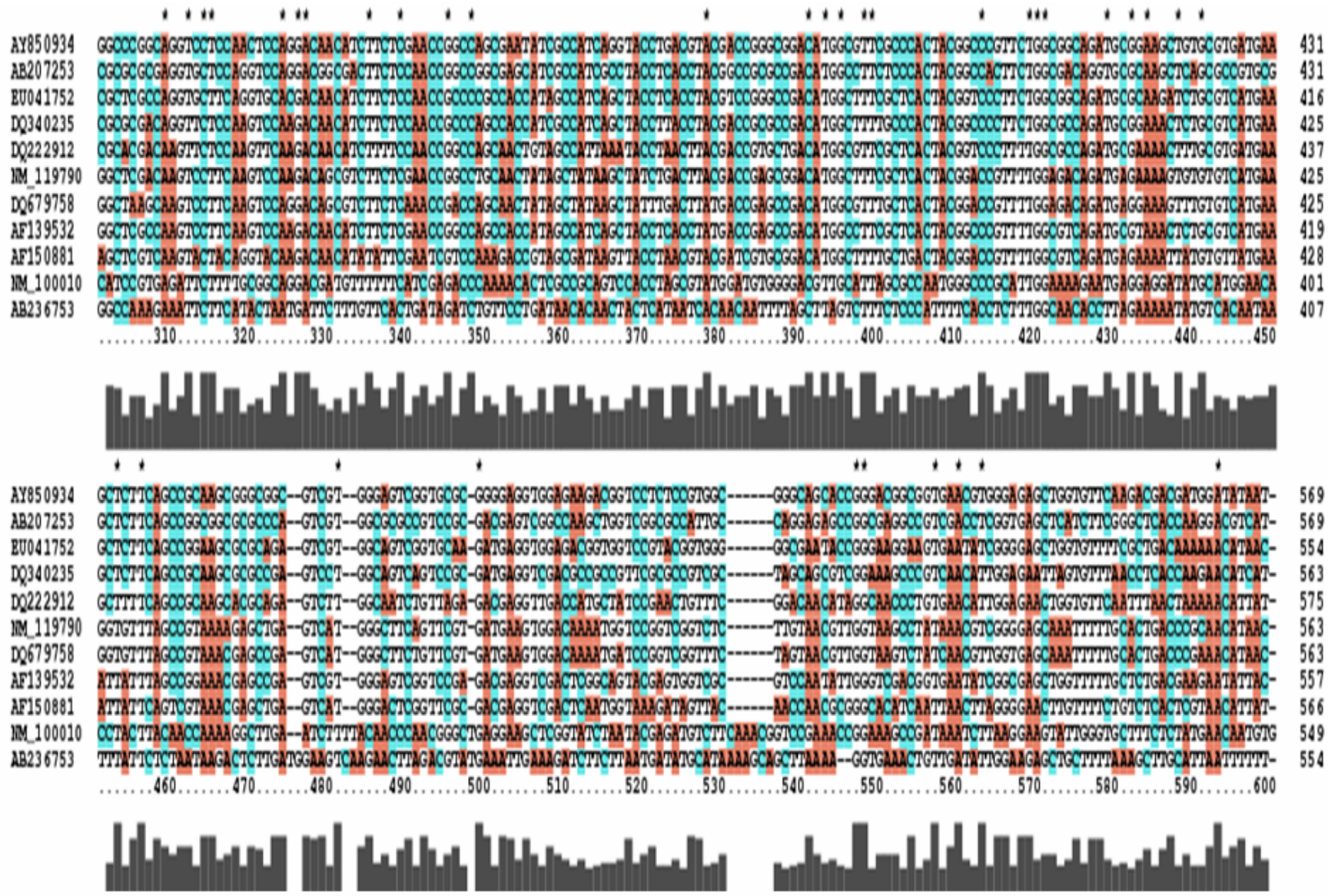


Fig 3.18: continued

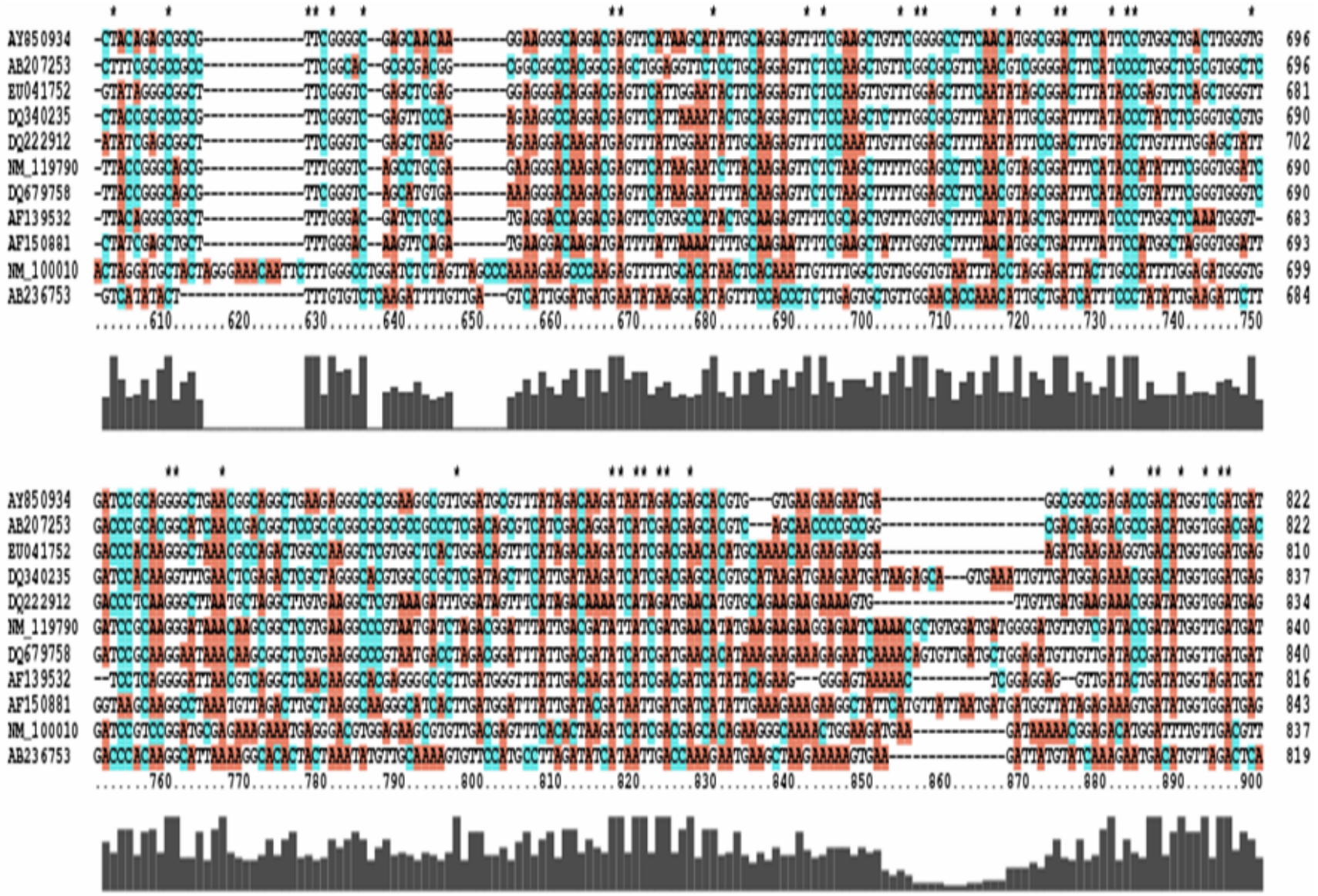


Fig 3.18: continued

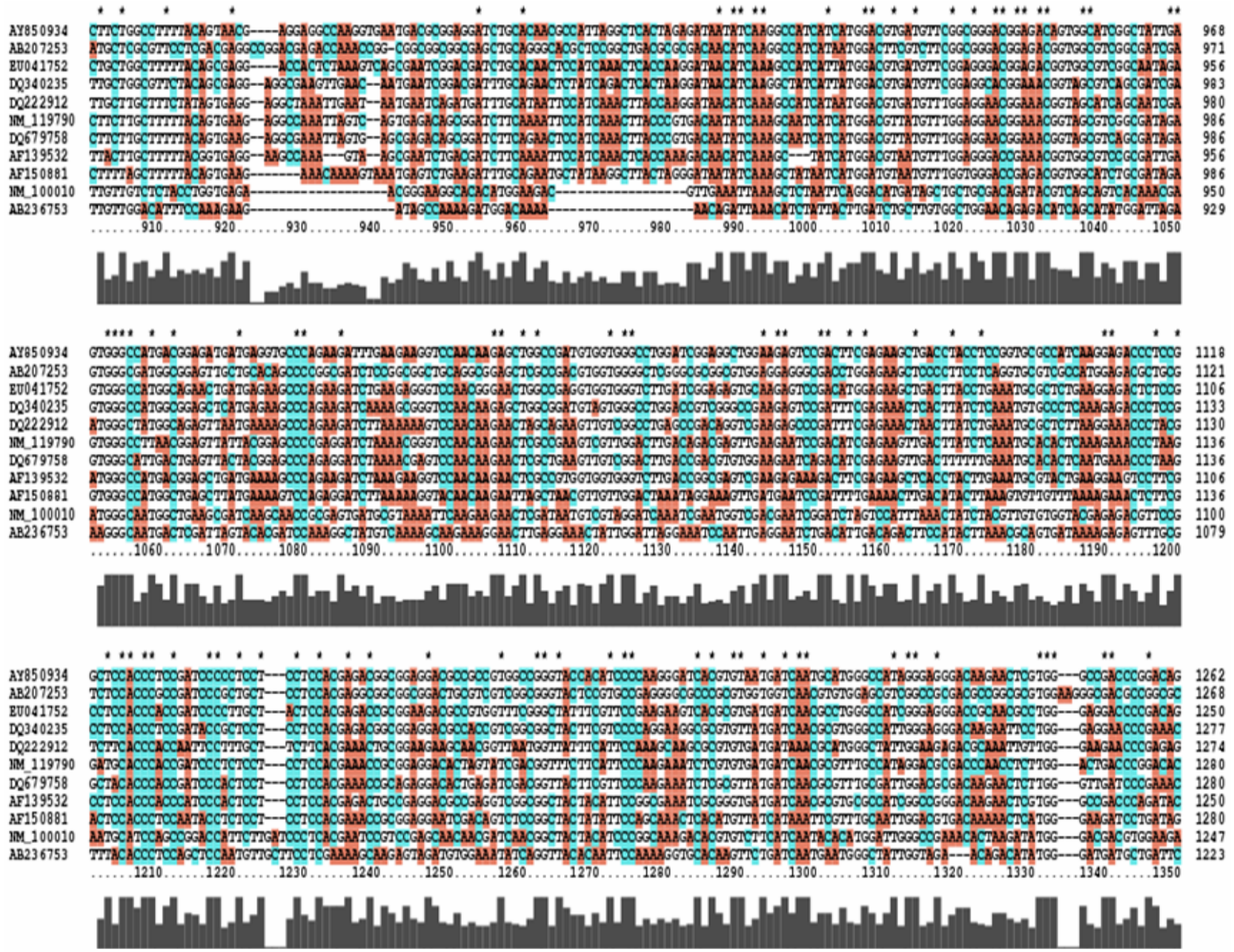


Fig 3.18: continued

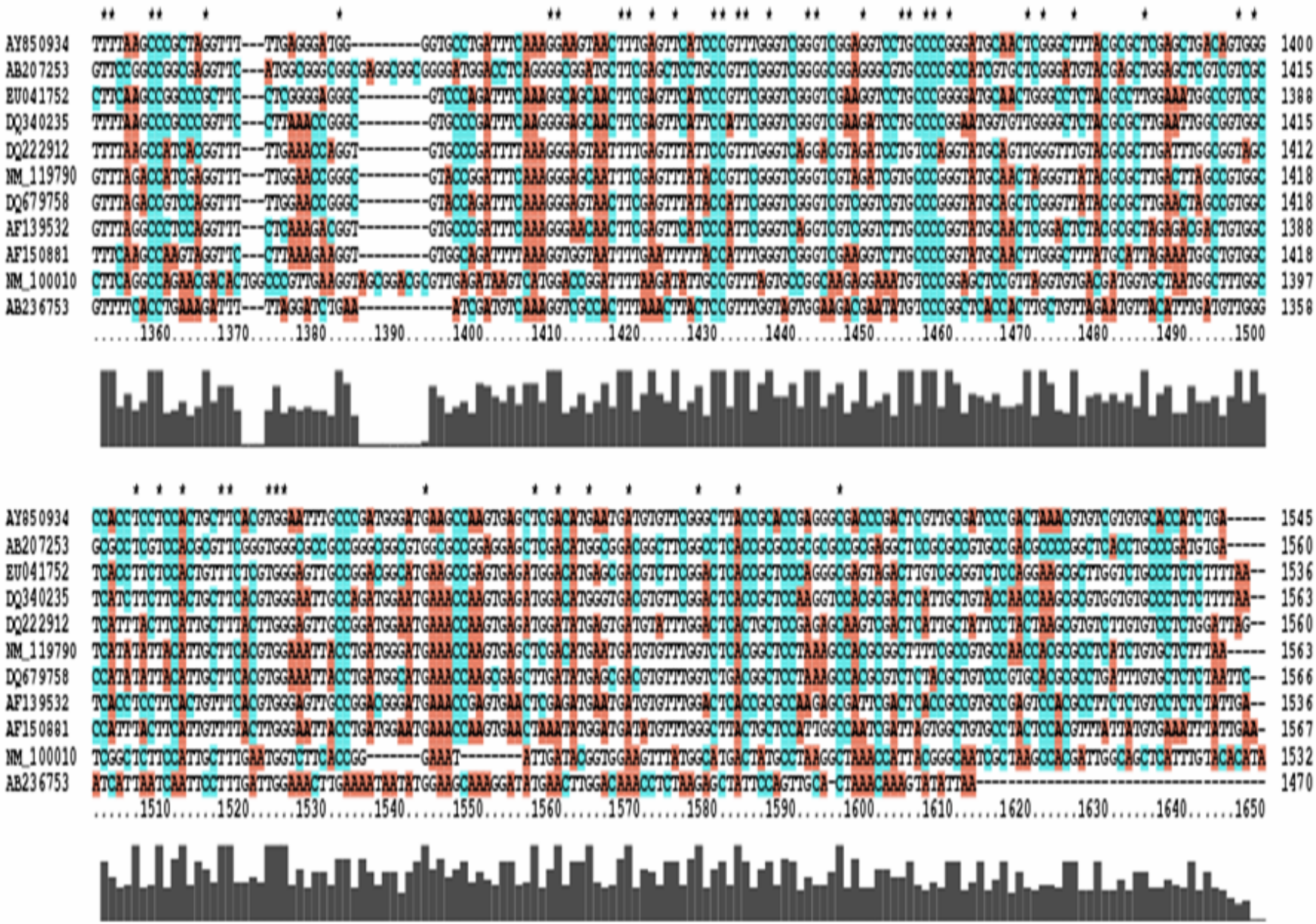


Fig 3.18: Multiple sequence alignment of selected plant P450s with LICALD5H gene.

3.3.7.3 Characterization of amino acid sequence

Deduced amino acid sequence (Figure 3.19) has shown 70-95% sequence similarity with the reported CAld5H/F5H protein sequences. Moreover, the peptide sequence exhibits characteristic features of a typical cytochrome P450 protein (Figure 3.20). The calculated *p*I value of the protein is 6.55. The signal peptide is of 29 amino acid residues, as predicted by signal-P bioinformatics tool. Approximately 50% of the signal peptide, 14 out of 29 is comprised of leucine residues only. This signal peptide sequence is immediately followed by a putative stop transfer sequence RFRKR (Meyer *et al*., 1999).

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

Figure 3.19. Deduced amino-acid sequence of LlCAld5H. The unbroken underline sequence highlights the N-terminal signal sequence. The broken line below the sequence shows the stop transfer sequence. Pro-rich consensus sequence has depicted as blue letters whereas the most conserved heme- binding ligand sequence characteristics of P450 proteins has shown as red and underlined letters.

Following the stop transfer sequence near the amino terminus is the sequence PPGPKGLP, which obeys the consensus for the proline rich sequence found in many P450s. A notable stretch between Pro-440 to Gly-450 is present. This region contains eight residues that comprise the heme-binding domain and are highly conserved among most P450s.

The cys-448 in LlCAld5H is the most likely heme-binding ligand in this enzyme (?). This part forms an active center for catalysis in iron-protoporphyrin IX (heme) with the thiolate of the conserved cysteine residue as fifth ligand. The resting P450 is in the ferric form and partially six-coordinated with a molecule of solvent. P450s are usually monoxygenases, catalyzing the insertion of one of the atoms of molecular oxygen into a substrate, the second atom of oxygen being reduced to water. The most frequently catalyzed reaction is hydroxylation (O insertion) using the very reactive and electrophilic iron-oxo intermediate (Daniele Werck-Reichhart and Rene Feyereisen, 2000).

3.3.7.4 LlCAld5H belongs to the CYP450s family and classified as CYP84A40

The LlCAld5H amino-acid sequence was characterized as in the previous section. The LlCAld5H protein belongs to cytochrome P450 family and it was classified as **CYP84A40**, by Prof. David Nelson upon request. (Prof. Nelson maintains the all P450s protein database and keeps an update of the CYP450s of all the organisms. He does classification of the P450s protein upon request. He is associated with the "P450 nomenclature committee''). The LlCAld5H has been added to Nelson home page [html http://drnelson.utmem.edu/CytochromeP450].

3.3.7.5 Analysis of amino acid sequence of the *Leucaena* coniferaldehyde 5-hydroxylase for conserved domain

Amino- acid sequence of LlCAld5H 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 domain, cytochrome P450 has been detected which belongs to CypX super-family. It involves in secondary metabolites biosynthesis, transport and catabolism (Figure 3.20).

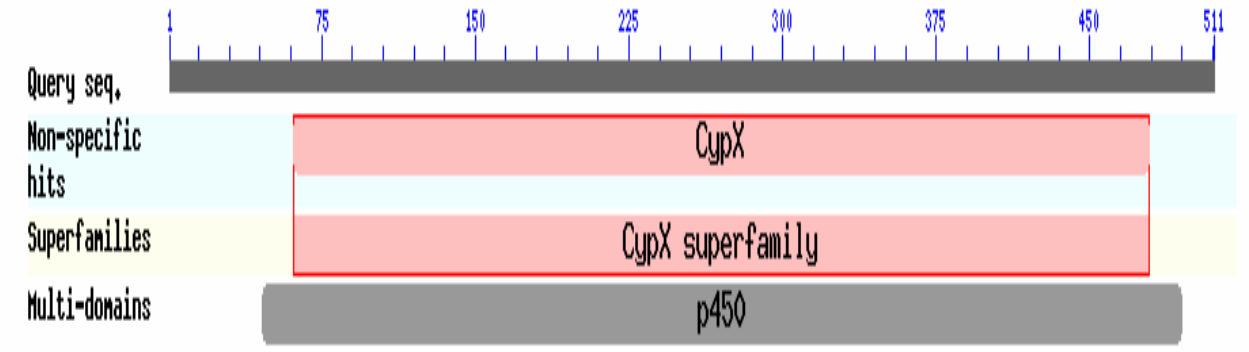


Figure 3.20. Image showing the putative conserved domain of LlCAld5H protein.

3.3.7.6 Multiple sequence alignment of *Leucaena leucocephala* CAld5H amino acids with plant CYP450s proteins.

Multiple sequence alignment of the coniferaldehyde 5-hydroxylase amino acid sequences was done using, *Glycine max* (DQ340235), *Medicago sativa* (DQ222912), *Liquidambar styraciflua* (AF139532), *Brassica napus* (DQ679758), *Trifolium pratense* (AB236753), *Arabidopsis thaliana* (NM_119790), *Broussonetia papyrifera* (AY850934), *Solanum lycopersicum x Solanum peruvianum* (AF150881), *Oryza sativa Japonica Group* (AB207253), and *Arabidopsis thaliana* (NM_100010) amino-acids sequences (Figure 3.21). GenBank accession numbers were given in bracket.

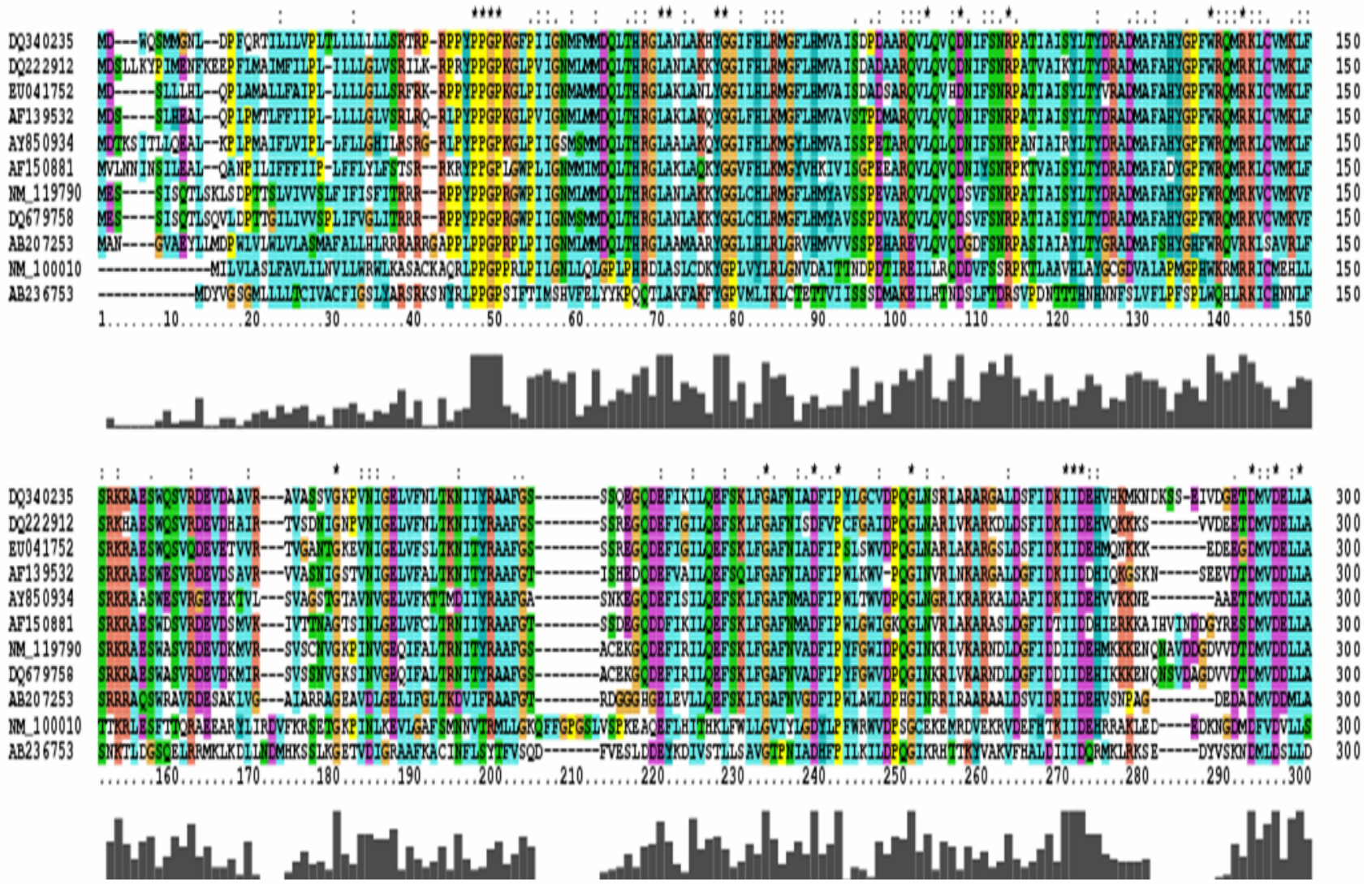


Fig. 3.21: Continued

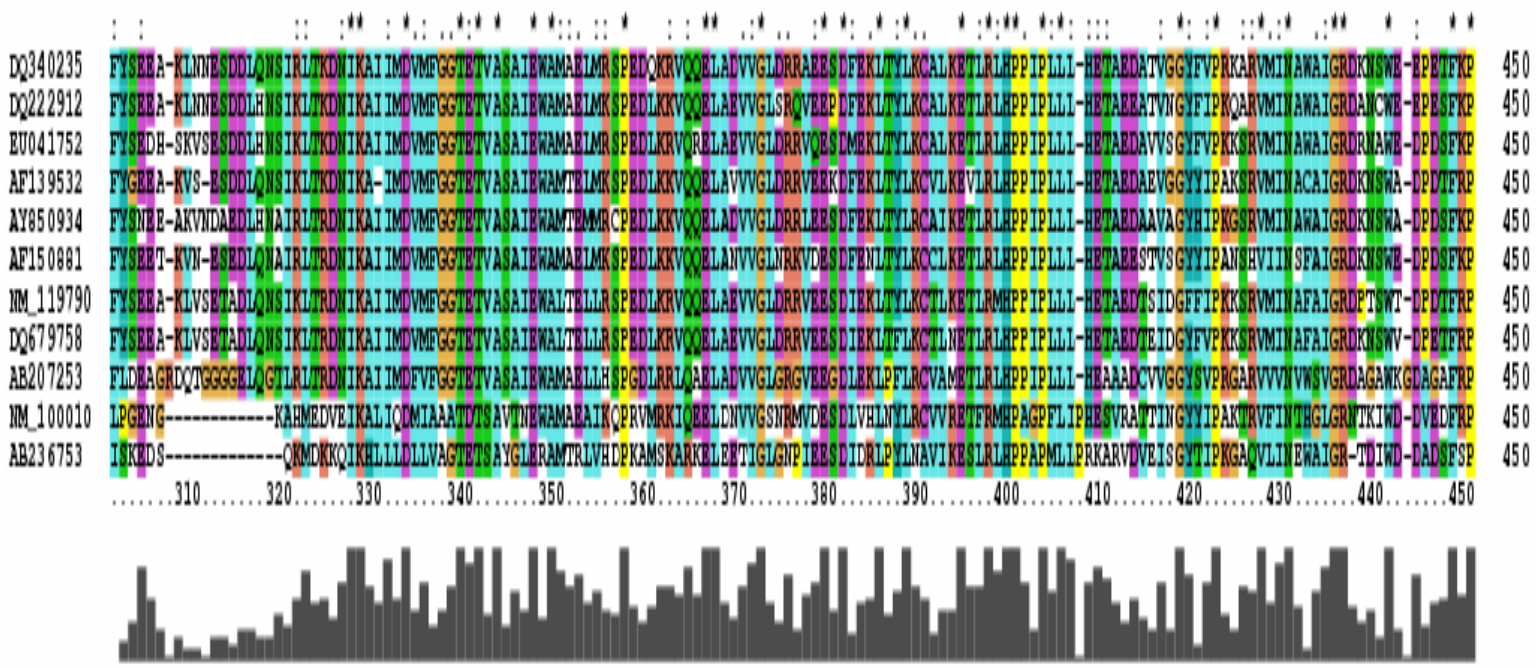


Fig. 3. 21: Multiple sequence alignment of selected plant P450s with LICALD5H protein

3.3.7.7 Theoretical molecular weight (MW) and Isoelectric Point of the coniferaldehyde 5-hydroxylase.

Theoretical molecular weight (MW) and Isoelectric Point of the LlCAld5H were analyzed at http://www.expasy.ch/cgi-bin/protparam and found to be 57426.5 Daltons and 6.52 (theoretical pI) respectively.

3.3.7.8 Amino acid composition of the *Leucaena* coniferaldehyde 5-hydroxylase

Amino acid composition of the *Leucaena* coniferaldehyde 5-hydroxylase was analyzed at http://www.expasy.ch/cgi-bin/pi_tool. Amino acid, number and the percentage composition of individual amino acid is given in table 3.1.

Table 3.1 Amino acid Composition of LlCAld5H

Amino acids	Total No.	%	Amino acids	Total No.	%
Ala (A)	43	8.4	Leu (L)	60	11.7
Arg (R)	33	6.5	Lys (K)	28	5.5
Asn (N)	15	2.9	Met (M)	23	4.5
Asp (D)	32	6.3	Phe (F)	27	5.3
Cys (C)	5	1.0	Pro (P)	26	5.1
Gln (Q)	14	2.7	Ser (S)	37	7.2
Glu (E)	32	6.3	Thr (T)	15	2.9
Gly (G)	33	6.5	Trp (W)	7	1.4
His (H)	13	2.5	Tyr (Y)	10	2.0
Ile (I)	27	5.3	Val (V)	31	6.1

Total number of negatively charged residues (Asp + Glu): 64

Total number of positively charged residues (Arg + Lys): 61

Atomic composition of LlCAld5H:

Carbon	C	2575
Hydrogen	H	4068
Nitrogen	N	700
Oxygen	O	731
Sulfur	S	28

Formula of LlCAld5H protein**:** $\mathbf{C_{2575}H_{4068}N_{700}O_{731}S_{28}}$

Total number of atoms**: 8102.**

3.3.7.9 Hydropathy index of the LlCAld5H amino acids

The hydropathy index of an amino acid is a number representing the hydrophobic or hydrophilic properties of its side-chain (Jack Kyte and Russell Doolittle, 1982). The larger the number is, the more hydrophobic the amino acid. The most hydrophobic amino acids are isoleucine (4.5) and valine (4.2). The most hydrophilic ones are arginine (-4.5) and lysine (-3.9). This is very important in protein structure; hydrophobic amino acids tend to be internal (with regard to the protein's 3- dimensional shape) while hydrophilic amino acids are more commonly found towards the protein surface. The LlCAld5H amino acid sequence was analyzed using Kyte-Doolittle Hydropathy plot at http://gcat.davidson.edu/rakarnik/kd.cgi. This plot clearly indicates that the hydrophobic amino acids outnumbers the hydrophilic one and are equally distributed throughout of the sequence. This finding is in agreement with the nature of other plant and animal P450 proteins as they are rich in hydrophobic amino-acids and localized in microsomes.

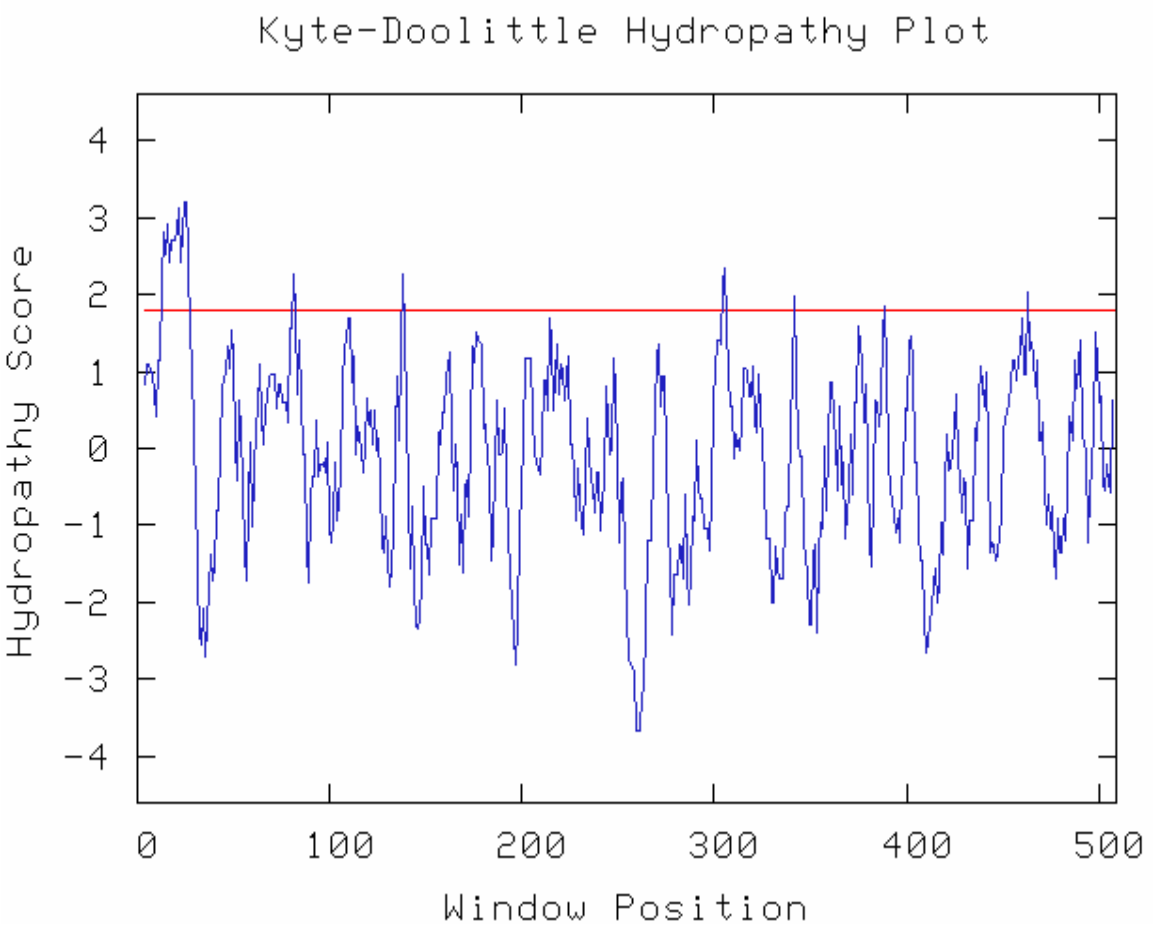


Figure 3.22. Hydropathy index of LlCAld5H amino acids (Kyte-Doolittle plot). Regions with values above 0 are hydrophobic in character. Window size refers to the number of amino acids examined at a time to determine a point of hydrophobic character (default window size is 7).

3.3.8 Phylogenetic relationship of *L. leucocephala* coniferaldehyde 5-hydroxylase protein with plant P450s families.

A phylogenetic tree was constructed after multiple sequence alignment of plant P450 proteins involved in secondary metabolites biosynthesis along with *Leucaena leucocephala* coniferaldehyde 5-hydroxylase.

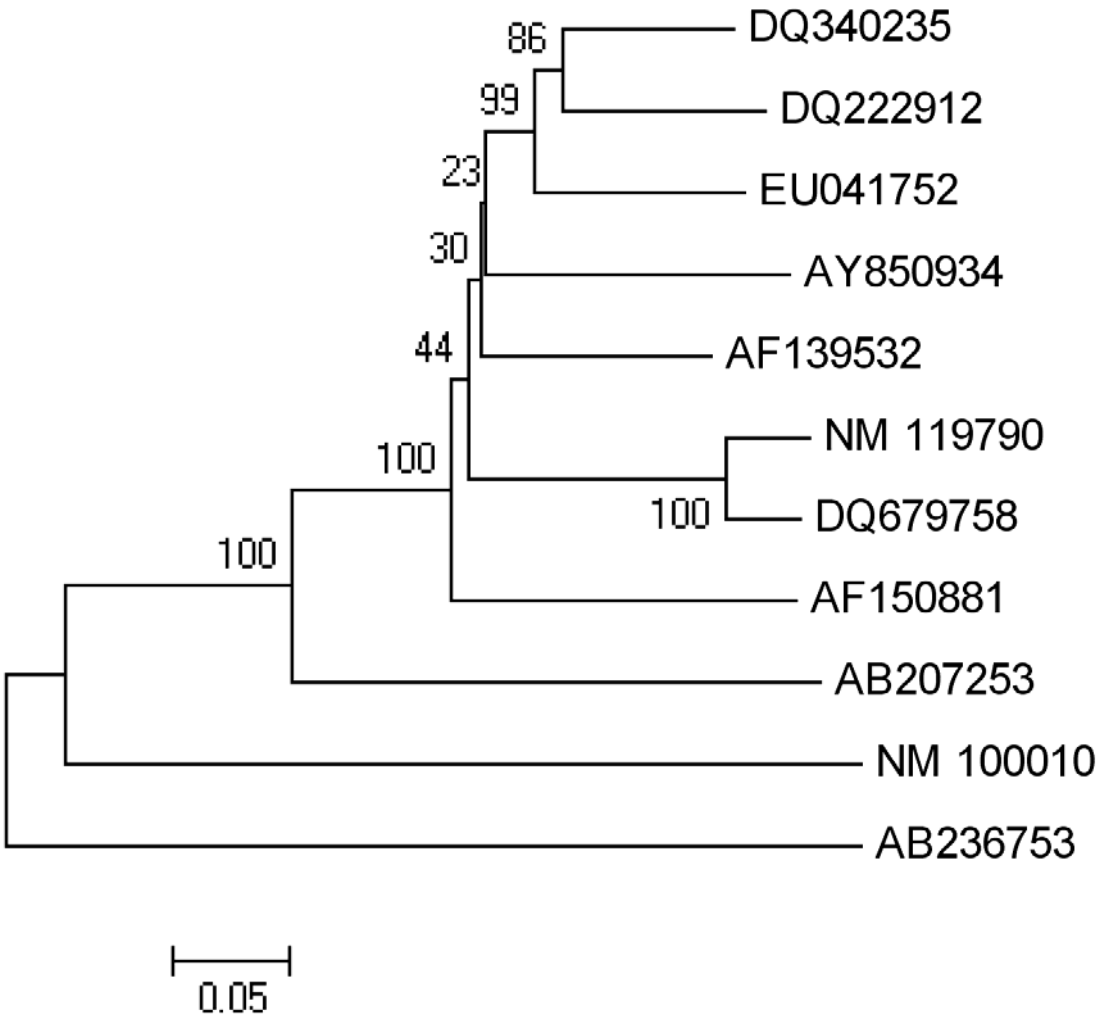


Figure 3.23. Phylogenetic tree of 11 selected plant P450s proteins, constructed by neighbor-joining methods using Mega 4.0 software program after alignment of amino acids sequences by Clustal X, followed by manual editing. The numbers shown at each node represent support values (percent) obtained by bootstrap analysis (1000 replicates). The tree shows the relatedness between amino acid sequences of CAld5H/F5H, the other plant P450s of known function and the *Leucaena leucocephala* CAld5H. The closest homologues of *Leucaena* CAld5H protein are F5H of *Medicago sativa* (DQ222912) and CYP84A16 of *Glycine max* (DQ340235). All branches are drawn to scale as indicated by the scale bar (= 0.05 substitutions/site). Details of the accession numbers has been given in the table 3.2.

Phylogenetic analysis reveals that *Leucaena leucocephala* CAld5H protein clustered with Ferulate 5-hydroxylase of *Medicago sativa* (DQ222912), *Broussonetia papyrifera* (AY850934) and CYP84A16 of *Glycine max* (DQ340235). These enzymes are involves in lignification process. This finding was tempted us to speculate that the higher degree of identity among CYP84A16 of *Glycine*, F5H of *Medicago* and *Broussonetia* and CAld5H of *Leucaena* is due to the fact that all of them catalyze the *meta*-hydroxylation of *para*–hydroxylated aromatic ring. The degree of identity among these proteins is more than the cutoff value of 40% that delineates P450 families (Nelson, D.R. and Strobel, H.W., 1988; Gotoh, O., 1992), thus CAld5H from *Leucaena leucocephala* identifies a P450 family that can be categorized as CYP84A40.

Table 3.2. Details of GenBank accession numbers used in Figure 3.23 and 3.24.

Accession No.	**Proteins/nucleotides**	**Plant species**
DQ340235	CYP 84A16	*Glycine max*
DQ222912	F5H	*Medicago sativa*
EU041752	CAld5H	*Leucaena leucocephala*
AY850934	F5H	*Broussonetia papyrifera*
AF139532	Aldehyde-5-hydroxylase (Ald5H) mRNA	*Liquidambar styraciflua*
NM119790	F5H (FAH1)	*Arabidopsis thaliana*
DQ679758	F5H	*Brassica napus*
AF150881	F5H (CYP84)	*Solanum lycopersicum x Solanum peruvianum*
NM_100010	CYP450	*Arabidopsis thaliana*
AB236753	CYP84	*Trifolium pratense*
AB207253	CYP450	*Oryza sativa Japonica Group*

3.3.8.1 Phylogenetic relationship of *L. leucocephala* coniferaldehyde 5-hydroxylase gene with plant P450s gene families.

To determine the relatedness of *L. leucocephala* CAld5H gene with other P450s gene families of plants, phylogenetic analysis was undertaken. A neighbor-joining phylogenetic tree was constructed after nucleotide sequence alignment using Clustal X. The CYP84A16 of *Glycine* and the CYP450 gene of *Oryza sativa japonica cultivar* are clustered together with *L. leucocephala* CAld5H gene. The ferulate-5- hydroxylase nucleotide shows very distant relation with *Leucaena* CAld5H gene in contrast with the previous tree constructed based on protein sequence alignment where the both shares much closer homology.

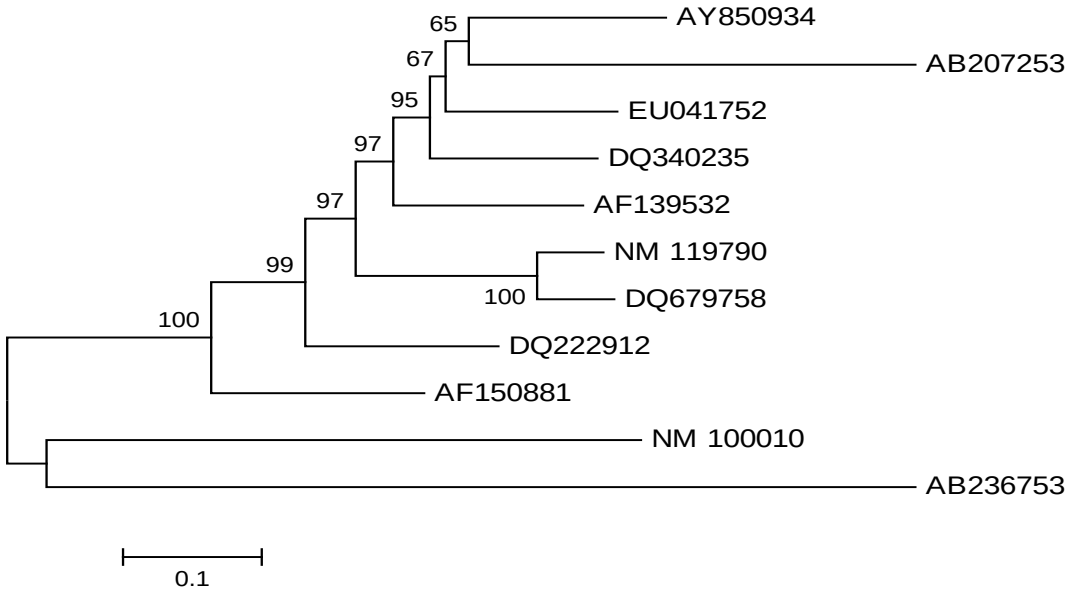


Figure 3.24. Phylogenetic tree of 11 plant P450s genes constructed by neighbor-joining methods using Mega 4.0 software program. The numbers shown at each node represent support values (percent) obtained by bootstrap analysis (1000 replicates). The tree shows the relatedness between nucleotide sequences of CAld5H/F5H of different plant P450s of known function and the *Leucaena leucocephala* CAld5H. All branches are drawn to scale as indicated by the scale bar (= 0.1 substitutions/site). Details of the accession numbers has given in the table 3.2.

It is also inferred from the study that the CYP84A16 of *Glycine max* is phylogenetically very close to the *Leucaena* CAld5H with reference to protein as well as nucleotides. The study further strengthens the hypothesis we made in the previous section 3.3.8, that the CAld5H from *Leucaena leucocephala* belongs to a P450 family that can be categorized as CYP84A40. We further hypothesize that the closely clustered members exhibit similar, if not identical, substrate specificities.

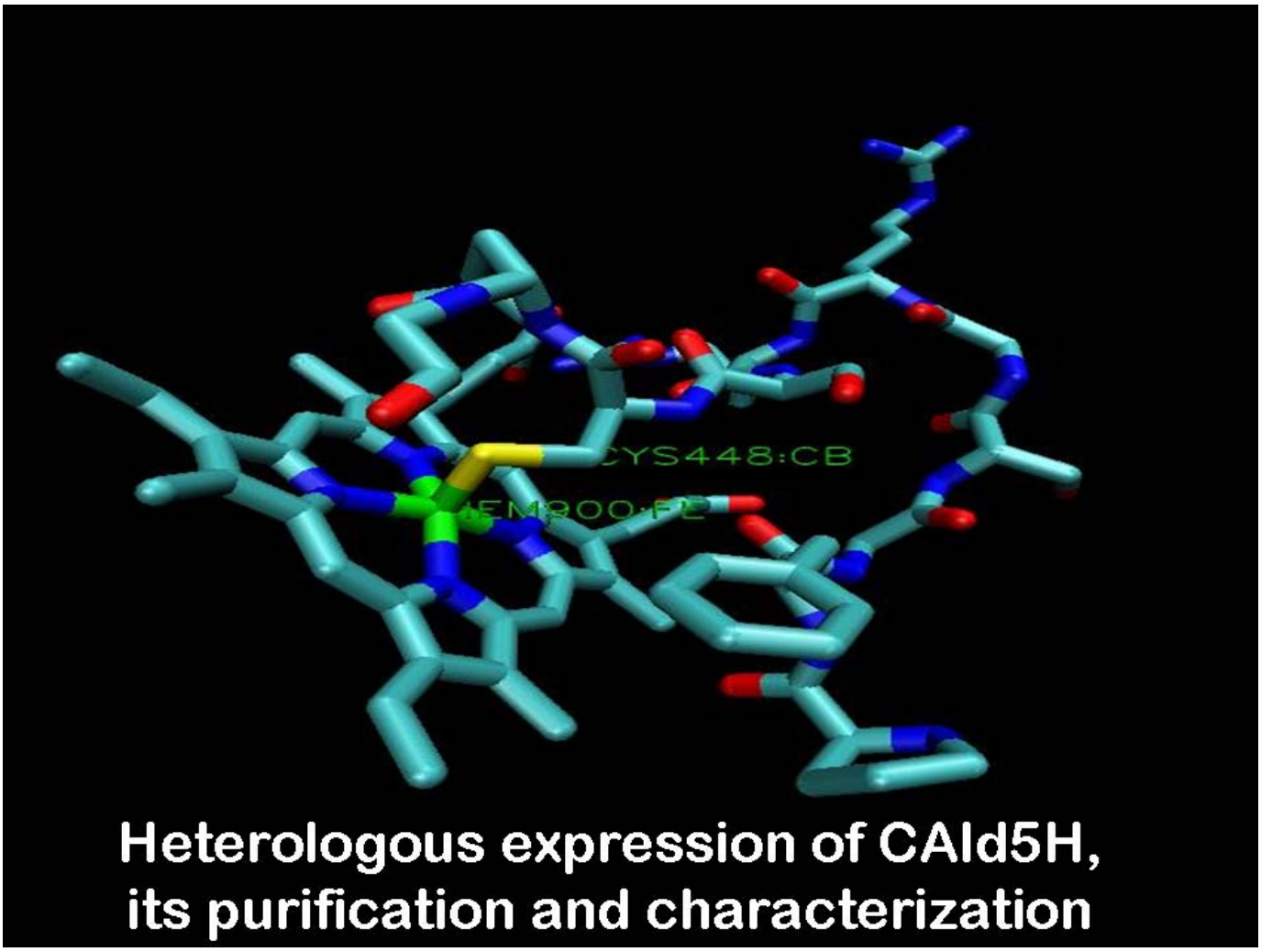
3.3.9 Discussions

Adopting the PCR based approach, a full length cDNA clone of coniferaldehyde 5-hydroxylase from *Leucaena leucocephala* was isolated. Nevertheless, a plethora of optimization has been done to get amplification from gDNA and/or cDNA template. Even several sets of primers were used in amplification. Ultimately a 621 bp fragment was obtained. The sequence analysis shows that it is a partial-length CAld5H gene. Based on this sequence, few sets of forward and reverse primers were designed. 5'and 3'rapid amplification of cDNA ends (RACE) was performed. The full length cDNA clone thus obtained has shown significant similarity with the published CAld5H/CYP450 gene sequences. The sequences was submitted to NCBI with accession no. EU041752. Along with the complete open reading frame of 1536 bp, a partial 5' UTR and complete 3' UTR was also isolated. Southern hybridization and slot blot analysis with homologous probes reveals that the CAld5H gene in *L. leucocephala* exists as a gene family of possibly 2-3 members. The deduced amino acid sequences were characterized and the calculated molecular weight was found to be 57.4 kDa. This protein contains several characteristics features of cytochrome P450s protein. The Phylogenetic analysis reveals that the protein is more closely related to *Glycine max* (both the plants belong to *Fabaceae* family. Finally, it was concluded that this LlCAld5H protein belongs to CYP 450 family and it was classified as CYP84A40 (with curtsy of Prof. David Nelson, USA).

3.3.10 Conclusion

- A partial cDNA clone, LlCAld5H was isolated and sequence of the clone was submitted to NCBI GenBank with accession no. DQ986906.
- Southern hybridization and slot blot analysis with homologous probes reveals that the CAld5H gene in *L. leucocephala* exists as a gene family of possibly 2-3 members.
- Rapid amplification of both 5' and 3' ends of the cDNA ends was done and the sequences corresponding to 5' and 3' regions of the gene were isolated.
- A complete open reading frame of 1536 bp of the gene was isolated, cloned and characterized. The GenBank accession no. of full length gene is EU041752.
- Approximately 300 bp of 3' untranslated region (3'UTR) and a small 5' UTR were isolated and characterized.
- The clone containing complete ORF of the CAld5H from *Leucaena leucocephala* was referred as LlCAld5H. The LlCAld5H nucleotide shows 77% sequence homology with CAld5H gene of sweetgum.
- The LlCAld5H encodes a protein of 511 amino acids with molecular weight of 57.24 KDa and the theoretical pI (isoelectric point) of the protein is 6.52.
- The amino acids sequence shows 83% sequence identity with CYP84A16 of *Glycine max*. The CAld5H protein sequence displays all the characteristic features of a plant P450 protein including Heme-binding ligands, Stop transfer sequence etc.
- LlCAld5H protein and the LlCAld5H gene are phylogenetically closer to CYP84A16 of *Glycine max* (both the plants belongs to *Fabaceae* family).

Chapter 4A


CYS448:CB
Heterologous expression of CAld5H,
its purification and characterization

This section comprises heterologous expression of *LlCAld5H* gene in *E. coli* BL21 (DE3) and purification of recombinant protein from inclusion bodies using Ni-chelated affinity column. In addition to that, raising polyclonal antibody in rabbit against the purified LlCAld5H and the activity studies from the crude enzyme preparation have also been discussed. Finally, the homology modeling for structural information of the LlCAld5H protein based on amino-acid sequences has been done.

4A.1. Introduction

From last several decades it has been thought that syringyl monolignol biosynthesis in angiosperms occurs through ferulate pathway from caffeate to sinapate via ferulate and 5-hydroxyferulate would lead to such a biosynthesis (Grisebach 1981; Grand 1984; Higuchi 1985). But in a revolutionary study based on HPLC/MS characterization of products from reactions of microsomal proteins from lignifying stem xylem of sweetgum (*Liquidambar styraciflua*) with a mixture of four potential 5-hydroxylation substrates, ferulate, feruloyl-CoA, coniferaldehyde and coniferyl alcohol, Osakabe et al., (1999) discovered that 5-hydroxyferulate was not synthesized. Instead, the exclusive product from this mixed substrate reaction was 5-hydroxyconiferaldehyde, demonstrating for the first time that a coniferaldehyde 5-hydroxylase (*CAld5H*) is involved in monolignol biosynthesis, and F5H may not be (Osakabe et al., 1999). To address the central question in lignin biosynthesis, how syringyl monomers synthesized from guaiacyl intermediates Osakabe et al., (1999) have suggested this conversion takes place with involvement of a CYP450 and COMT mediated hydroxylation and methylation respectively, in sweetgum (*Liquidambar styraciflua*) xylem. Based on these findings, in an attempt to get an insight of the central objective of the thesis, a cDNA encoding cytochrome P450 monooxygenase from *Leucaena leucocephala* was cloned and characterized as described in the previous chapter. To study its functional expression the cDNA has been cloned into the expression vector, protein expression, purification, generation of polyclonal antibody, Western blotting and homology modelling has been described in the subsequent sections of results and discussion.

4A.1.1 Expression of *LlCAld5H* gene and protein purification

Over-expression of the recombinant protein can be achieved by use of the appropriate and efficient expression vector. Expression systems are designed to produce many copies of a desired protein within a host cell. In order to accomplish this, an expression vector is inserted into a host cell. This vector contains all of the genetic coding necessary to produce the protein, including a promoter appropriate to the host cell, a sequence which terminates transcription, and a sequence which codes for ribosome binding (Purves et al., 2001). Usually, vectors for cloning and expressing target DNA are derived from medium-copy plasmids such as pET. *E. coli* expression systems should meet several criteria including: a) minimal basal expression of the gene to be expressed under repressed conditions, b) Fast and uncomplicated induction of a wide variety of genes to a high level of expression and c) easy cloning and DNA manipulation features.

4A.1.2 General overview of pET expression system

One expression system was developed in 1986 by W. F. Studier and B. A. Moffatt, who created an RNA polymerase expression system which was highly selective for bacteriophage T7 RNA polymerase. The initial system involved two different methods of maintaining T7 RNA polymerase into the cell - in one method, a lambda bacteriophage was used to insert the gene which codes for T7 RNA polymerase, and in the other, the gene for T7 RNA polymerase was inserted into the host chromosome (Studier et al, 1986). This expression system has become known as the pET Expression System, and is now widely used because of its ability to mass-produce proteins, the specificity involved in the T7 promoter which only binds T7 RNA polymerase, and also the design of the system which allows for the easy manipulation of how much of the desired protein is expressed and when that expression occurs (Unger, 1997). A pET vector is a bacterial plasmid designed to enable the quick production of a large quantity of any desired protein when activated. This plasmid (Fig. 4A.2) contains several important elements - a *lacI* gene which codes for the *lac* repressor protein, a T7 promoter which is specific to only T7 RNA polymerase (not bacterial RNA polymerase) and also does not occur anywhere in the prokaryotic genome, a *lac* operator which can block transcription, a polylinker, an f1 origin of replication (so that a single-stranded plasmid can be produced when co-

infected with M13 helper phage), an ampicillin/kanamycine resistance gene, and a ColE1 origin of replication (Blaber, 1998). To start the process, the desired gene is cloned into a pET plasmid at the polylinker site. Both the T7 promoter and the *lac* operator are located 5' to the cloned gene. When the T7 RNA polymerase is present and the *lac* operator is not repressed, the transcription candidate gene proceeds rapidly. Because T7 is a viral promoter, it transcribes rapidly and profusely for as long as the T7 RNA polymerase is present (Campbell, 2003). The expression of the corresponding protein increases rapidly as the amount of mRNA transcribed from the candidate gene increases. Within a few hours, desired protein is one of the most prevalent components of the cell (Unger, 1997).

4A.1.3 Introducing T7 RNA Polymerase to the host cell

One of the most important parts of the pET expression system involves the fact that the cloned gene is not transcribed unless the T7 RNA polymerase is present. Prokaryotic cells do not produce this type of RNA, and therefore the T7 RNA polymerase must be added. Usually, the host cell for this expression system is a bacteria which has been genetically engineered to incorporate the gene for T7 RNA polymerase, the *lac* promoter and the *lac* operator in it's genome. When lactose or a molecule similar to lactose is present inside the cell, transcription of the T7 RNA polymerase is activated. Typically, the host cell used is *E. coli* strain BL 21(DE3) (Blaber, 1998). T7 RNA polymerase can be introduced to the cell through methods other than incorporation a gene in the host chromosome. It can be introduced through infection of the original host cell with lambda CE6 (Novagen, 2003).

4A.1.4 How and Why IPTG control the pET Expression system?

Control of the pET expression system is accomplished through the *lac* promoter and operator. Before cloned gene can be transcribed, T7 polymerase must be present. The gene on the host cell chromosome usually has an inducible promoter which is activated by IPTG (isopropyl-β-D-thiogalactopyranoside). This molecule, IPTG, displaces the repressor from the *lac* operator. Since there are *lac* operators on both the gene encoding T7 polymerase and the desired cloned gene, IPTG activates both genes (Science Advisory Board, 2003; Campbell, 2003). Therefore, when IPTG is added to the cell, the

T7 polymerase is expressed, and quickly begins to transcribe the cloned gene which is then translated. IPTG works to displace a *lac* repressor since IPTG is an analogue of lactose (Blaber, 1998). The *lac* genes express enzymes which are involved in the breaking down of lactose, and therefore, the presence of lactose (or its analogue) would trigger the initiation of transcription of *lac* genes.

4A.1.5 Molecular modeling of coniferaldehyde 5 – hydroxylase.

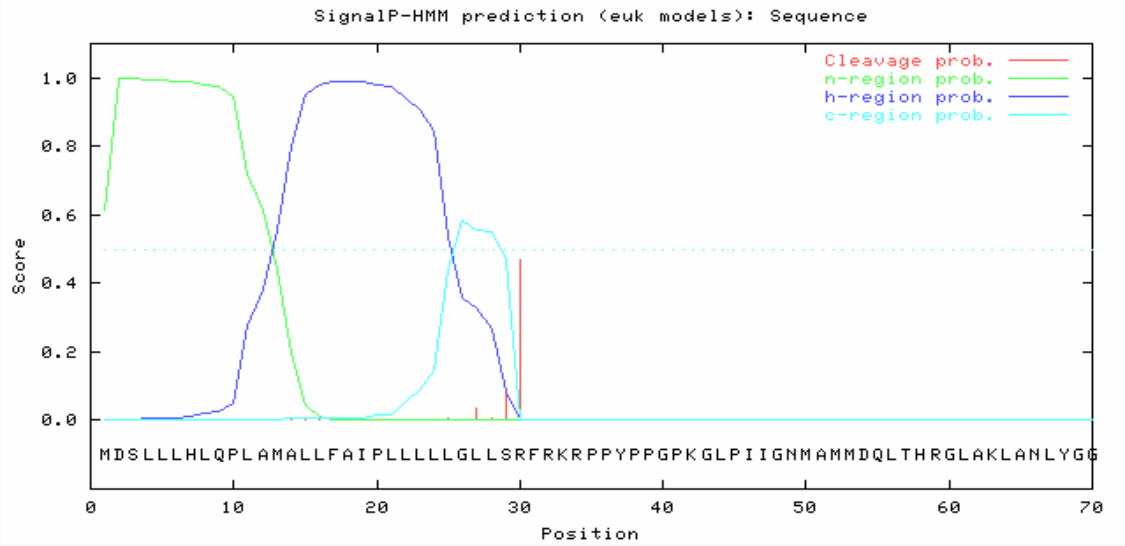
The structural analysis of proteins is a crucial step in understanding their mechanism of action and function. The three-dimensional structure of the CAld5H protein is not available till date from any plant; its protein sequence deduced from the LlCAld5H gene sequence was used to carry out a homology modeling. The software MODELLER9v7 (Eswar *et al.,* 2000) was used to generate 3D model of the protein. The quality and reliability of the suggested models were primarily assessed using the MODELLER9v7 parameters and then using PROCHECK (Laskowski *et al.*, 1993, Morris *et al.*, 1992), WHATCHECK (Hooft *et al.*, 1996) and Verify3D (Bowie *et al.*, 1991; Lüthy *et al.*, 1992). Since the heme-binding domain is crucial for the cytochrome P450 class of proteins, the class to which this protein belongs as well, a superposition of model with known structures was carried out in order to understand the nature of the binding site. The bound heme prosthetic group and the surrounding residues were analyzed for the interactions present in the binding site.

4A.2 Material and Methods

4A.2.1 Signal Peptide prediction

The signal peptide was predicted using Signal-P bioinformatics tool (www. cbs.dtu.dk/services/signalp/) and the mature protein was deduced by considering the output of the prediction. The output data (Fig. 4A.1) suggests the presence of a signal peptide of 29 amino-acids at N-terminal. The cleavage site was predicted between 29 and 30 amino acids.

SignalP-HMM result:



data

>Sequence
Prediction: Signal peptide
Signal peptide probability: 0.613
Signal anchor probability: 0.386
Max cleavage site probability: 0.467 between pos. 29 and 30

Figure 4A.1. Picture showing the LlCAld5H signal peptide predicted by Signal–P tools. The most probable signal peptide cleavage site has shown by red vertical line on x-axis.

4A.2.2 Cloning of LlCAld5H cDNA in expression vector

The pET-28a (+) vector (Novagen, USA) was used for expression of LlCAld5H gene in *E.coli* BL 21 (DE3). This vector carries a N-terminal His•Tag®/thrombin/Ser•Tag™/ enterokinase configuration with an additional C-terminal His•Tag sequence. This vector has unique restriction sites for cloning (Fig. 4A.2). The cloned gene is expressed under the T7 RNA polymerase promoter.

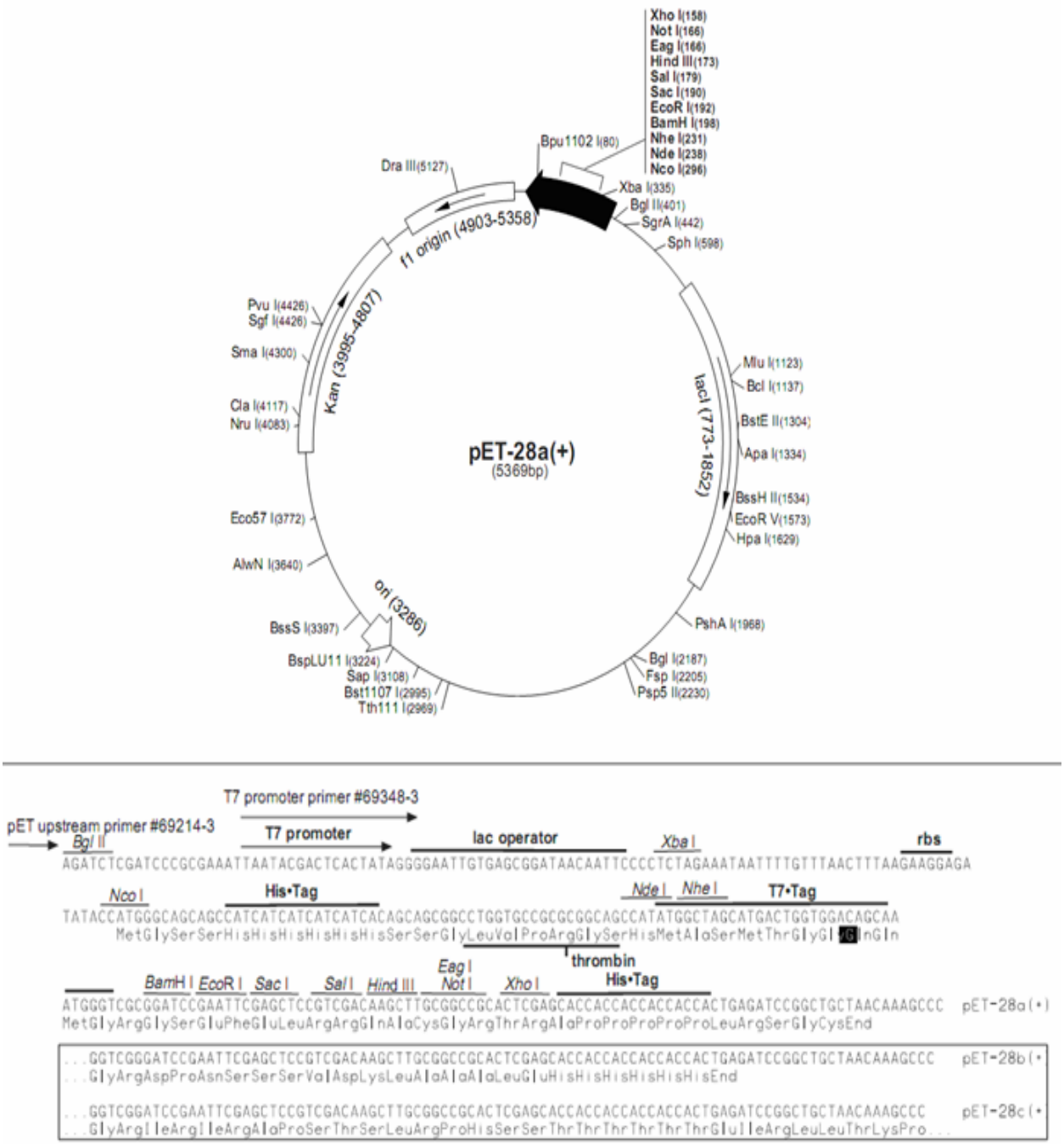


Figure 4A.2. Vector map of pET 28a (+) and the lower panel shows MCS of the vector.

4A.2.2.1 Digestions and Ligation

Digestion: 1 µg purified plasmid DNA : 3.0 µl
10 x restrictions Buffer : 3.0 µl
Restriction Enzyme/s : 1.0 µl (10 U/µl)
H2O up to : 23.0 µl
Total Volume : 30.0 µl
the reaction was incubated at 37 °C for 2-3 h for complete digestion.

Ligation: 100 ng plasmid vector : 0.5 µl
300 ng of Insert : 2-3 µl
2 x Ligation Buffer : 5.0 µl
T4-DNA ligase : 1.0 µl (3 U/µl) H2O up to : 1.5-2.5 µl Total Volume : 10.0 µl the reaction was incubated at 4 °C for 24 h or 16 °C for 16 h.

4A.2.3 Preservation of bacterial culture

Transformed bacterial cells (*E. coli*) containing plasmids with target insert were stored in glycerol. In an Eppendorf tube, 800 µL of bacterial culture were mixed with 200 µL of glycerol (sterilized by autoclaving). The mixture was vortexed to ensure that the glycerol is dispersed. The tube was stored at -80 °C for long-term storage. To recover the bacteria, the frozen surface of the culture was scraped with a sterile inoculating needle and the bacteria adhered to the needle were immediately streaked on the surface of a LB agar plate containing the appropriate antibiotic. The plates were incubated O/N at 37 °C

4A.2.4 Expression and purification of recombinant proteins

Many natural proteins have metal binding sites which can be used for purification. The concept of this type of purification tool is rather simple. A gel bead is covalently modified to display a chelator group for binding a heavy metal ion like Nickel. Affinity

chromatography is viewed as a group-selective tool for purifying the metal-binding class of proteins. His-tagged recombinant protein can be purified by metal chelate affinity chromatography. The initial stage of His-tagged protein purification is based on the remarkable selectivity and high affinity of patented Ni^{2+}- NTA (nickel-nitrilo-triacetic acid) resin for proteins containing an affinity tag of six consecutive histidine residues, the 6xHis-tag. NTA, which has four chelating sites for nickel ions, binds nickel more tightly than metal-chelating purification systems that have only three sites available for interaction with metal ions. The extra chelation site prevents nickel ion leaching, providing a greater binding capacity and high-purity protein preparations. Isolation and purification of proteins were conducted using Ni^{2+}- NTA agarose beads (Qiagen). Fusion proteins were purified as described earlier (Chapter 2; section 2.8.7.2). Purity of protein was checked on 10 % SDS PAGE (Chapter 2; section 2.8.7.6).

4A.2.4.1 Expression of active LlCAld5H and enzyme assay

The protein expression in native form was done according to the protocol described in chapter 2; section 2.8.8.3.

4A.2.5 Primary and secondary antibodies

Purified LlCAld5H protein was used for raising rabbit immune-serum (National Toxicology Centre, Badgaon Khurd, India). Polyclonal antibodies were purified from the immune-serum. Secondary antibodies *i.e.* goat anti-rabbit IgG conjugated with alkaline phosphatase were purchased from Bangalore Genei, Bangalore, India.

4A.2.6 Western blotting

Purified recombinant protein was used to raise antibody in rabbit as described previously (National Toxicology centre, Badgaon Khurd, Pune,India). Anti-LlCAld5H polyclonal antibodies were purified and used for Western blotting. The protocol was optimized in the laboratory as follow;

Semi-dry western blot transfer procedure: The protocol describes the procedure for transferring protein from polyacrylamide gel onto the nylon membrane using semi-dry transfer methods.

1. Cut 6-10 sheets of Whatman 3MM, the same size as the mini PAGE gels.

2. Cut membrane (nylon or immobilon-p PVDF) to the size of the gel.
3. Membrane activation: Wet PVDF membrane in methanol for 15 seconds (don't need this step for nylon membrane) and then transfer to a container filled with ddH_2O for 5 min. Soak membrane in 1x transfer buffer for 10 min while preparing transfer sandwich.
4. Soak Whatman filter paper in transfer buffer up to 10 min to avoid trapping of air bubble.
7. Assemble sandwich: (Bottom up on the surface of wet semi-transfer plate)
a) 3-5 sheets of wet whatman filter paper, b) PVDF membrane,
c) Polyacrylamide gel and d) 3-5 sheets of wet whatman filter paper.
8. Use a roller to roll over the sandwich gently to remove trapped air bubbles. Apply few ml 1x transfer buffer on top of the sandwich to avoid drying out during membrane transfer.
9. Transfer at 50 voltage for 2 h at 4 °C.

1x transfer buffer recipe: 25 mM Tris (pH 8.0), 250 mM Glycine and 15% methanol.

The blot was developed by immunodetection essentially as described by Mierendorf, R.C. *et al*., (1987). After blocking non-specific protein binding sites with 3% BSA for 1 h, the blot was incubated for 2 hr with 1μg ml^{-1} (1:10,000 dilution of the serum obtained) of anti-LlCAld5H. Alkaline phosphatase conjugated goat anti-rabbit IgG (Bangalore Genei,Bangalore, India) was incubated with the blot for 45 min. The blot was exposed to X-ray film (Kodak) with intensifying screen followed by color development by chemiluminescent substrate (Invitrogen).

4A.2.7 Homology modeling of LlCAld5H protein

The 3D model of the protein was built using homology modeling based on the high resolution homologous crystal structures obtained from Brookhaven Protein Database

(PDB). Initially a BLAST search was carried out to identify the related structures that could be used as template. Several structures of the cytochrome P450 family showed homology with the CAld5H protein. Crucial among them were the 2HI4 (Cytochrome P450 1A2), 3DL9 (CYTOCHROME P450 2R1), 3G5N (CYTOCHROME P450 2B4), 2P85 (Cytochrome P450 2A13), 2V0M (CYTOCHROME P450 3A4) and 2FDU (MICROSOMAL P450 2A). The homology modeling software MODELLER9v7 was used to build the 3D models in a workstation running Linux operating system. The software is used for comparative protein structure modeling that optimally satisfies spatial restraints (Sali and Blundell, 1993) which include:

- Homology derived restraints on the distances and dihedral angles in the target sequence extracted from its alignment with the template structures.
- Stereo chemical restraints such as bond length and bond angle preferences, obtained from CHARMM-27 molecular force-field. (MacKerell *et al.,* 1998).
- Statistical preferences for dihedral angle and non-bonded inter-atomic distances, obtained from a representative set of known well refined protein structures (Sali and Overington, 1994).

The spatial restraints are expressed as probability density functions (pdfs) for the features on which restraints are applied. The pdfs restrain C^{α}- C^{α} distances, main-chain N-O distances, main-chain and side-chain dihedral angles. Further, the search module of MODELLER9v7 based on pir_95 structure sequence database was used to refine the template search. The align2D module was used to build the alignment for the template 2HI4 on which the model was built. The model built was assessed by the DOPE score and GA341 potential which indicate the quality of the model generated. To separately assess the geometry of the model, it was submitted to PROCHECK and WHATCHECK. The PROCHECK suite of programs provides a detailed check of the stereochemistry of a protein structure. This analysis provides an assessment of the overall quality of the structure as compared with the database of well refined structures determined at same resolution and also highlights regions that may need further investigation. The PROCHECK programs are useful for assessing the quality of existing structures and of

those being modelled based on known structures. Finally the model was submitted to the Verify-3D server (http://nihserver.mbi.ucla.edu/Verify_3D/) to assess the quality of the structure. Verify3D analyzes the compatibility of an atomic model (three-dimensional) with its own amino acid sequence (one-dimensional). Each residue is assigned a structural class based on its location and environment (alpha helix, beta sheet, loop segment, polar or non-polar environment, etc). A collection of good structures is used as a reference to obtain a reliability score for each of the 20 amino acids in this structural class. The scores of a sliding 21-residue window (from -10 to +10) are added and plotted for individual residues.

4A.3 Results and Discussion

4A.3.1 Construction of LlCAld5H protein expression vector

LlCAld5H cDNA sequence encoding mature protein was amplified using forward primer (5'-GAATTCCGATTCCGTAAAAGGCCACCCTACC-3') and reverse primer (5'-GCGGCCGCAAAGAGAGGGCAGACCAAGCGC-3') containing *Eco*RI and *Not*I restriction sites respectively. The PCR reaction was performed (As described in chapter 2; section 2.8.4.9) with a proof-reading Taq DNA polymerase (LD-Taq DNA polymerase, Sigma-Aldrich, USA) to avoid any mis-incorporation of nucleotide and the consequent frame-shift in the open reading frame. An amplification product of ~1.4 bp was amplified. PCR product was cloned into pGEMT-easy vector. The positive recombinant clones were selected by colony PCR (Fig. 4A.3) and used in subsequent sub-cloning. The recombinant clones thus obtained were digested with *Eco*RI and *Not*I and sub-cloned into expression vector pET 28a (+). This expression plasmid, in which the LlCAld5H cDNA was sequenced and confirmed to have no PCR errors, was used to transform *E.coli* strain BL21 (DE) for expression. The recombinant plasmid thus obtained was referred as pET-28a-LlCAld5H. The plasmid DNA from host cells was isolated and confirmed to inhabiting the expression plasmid construct (recombinant clones) by restriction digestion analysis (Fig. 4A.4).

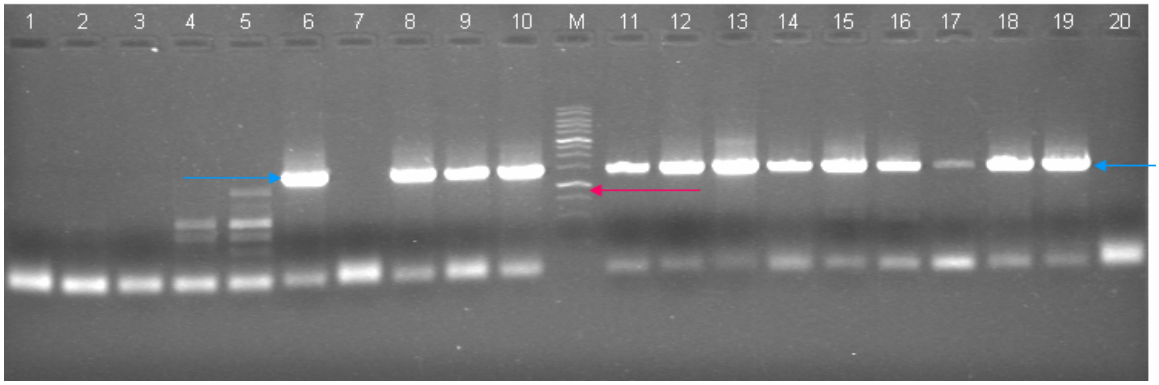


Figure 4A.3. Screening of the recombinant clones by colony PCR. The 1450 bp band (shown with blue arrow) was amplified from most of the clones. Lane M contains DNA size marker (Bioenzyme) resolved on the 1% agarose (red arrows locate 1 Kb band in marker lane).

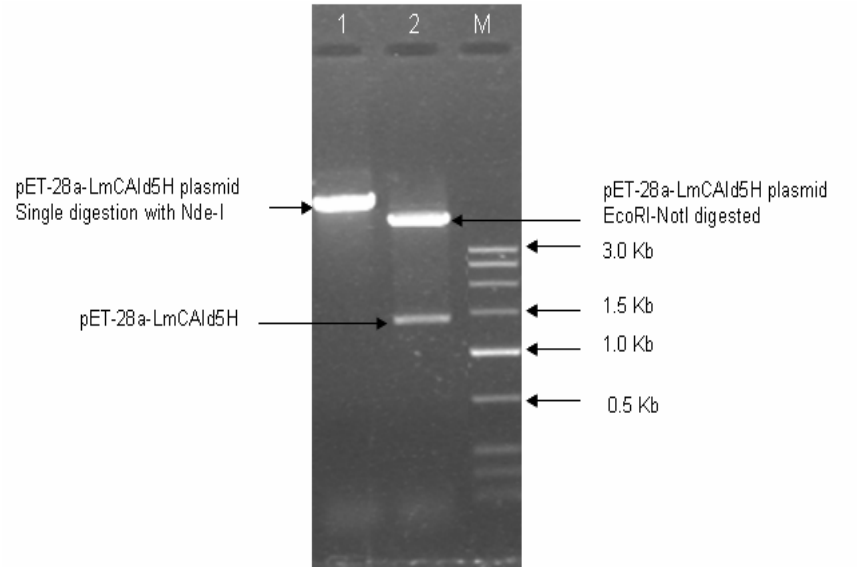


Figure 4A.4. Agarose (1%) gel showing the ~1.4 Kb fragment yielded from double digestion of pET28a-LmCAld5H clone in lane 2. The undigested but recombinant clone has shown in lane 1. Lane 3 is DNA size marker (Bangalore Genei, India). LmCAld5H clone contains coding sequence corresponding to mature protein (LmCAld5H=LlCAld5H-signal peptide).

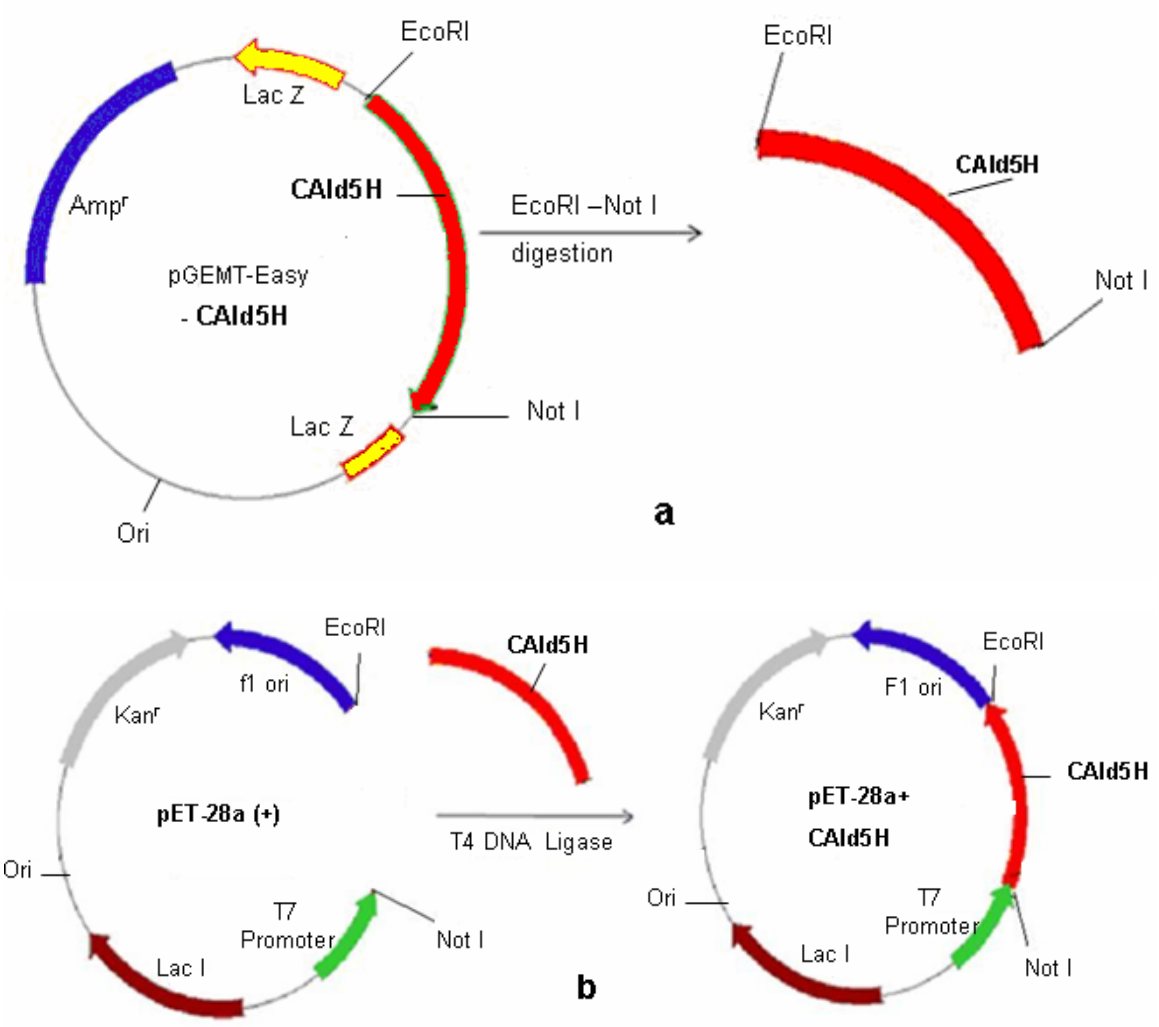


Figure 4A. 5. Diagrammatic representation of pET28a-LlCAld5H vector construction.

4A.3.2 Heterologous expression of LlCAld5H in *E. coli* BL21 (DE)

E.coli BL21 was used as a host for the expression of mature LlCAld5H protein. The protein expression (chapter 2; section 2.8.7.1) was induced by the addition of 1mM IPTG. Since, all the P450s protein contains the heme domain, 1mM δ-aminolevulinic acid (ALA) and 0.5 mM $FeCl_3$ were added to the culture as the precursor to the heme ligand (Haitham A. Hussain and John M. Ward, 2003). For protein visualization cells were harvested at different time interval (1-5 hrs) after IPTG induction and isolated by centrifugation. The pelleted cells were resuspended in sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) sample buffer and heated at 90˚C.

Proteins were separated by SDS-PAGE on 10% gel and stained with Coomassie blue. Molecular mass of the expressed protein was 57 kDa as confirmed by SDS- PAGE and consistent with the predicted sequence. The expression level of the LlCAld5H protein was increased gradually with time as reflected in figure 4A.6; lane 2-6. In uninduced bacterial cells the LlCAld5H level was merely detectable (Fig. 4A.6 lane 1). However, this study doesn't show whether the protein is getting expressed as in the soluble form or it is embodied in as inclusion body. To clarify the obvious doubt, cytosolic as well as insoluble fractions were prepared and analyzed for the LCAld5H protein expression. The study revealed the presence of LlCAld5H protein in inclusion bodies. This finding has shown an agreement with the Kyte-Doolittle plot constructed against the deduced LlCAld5H protein (chapter 3; section 3.3.7.8).

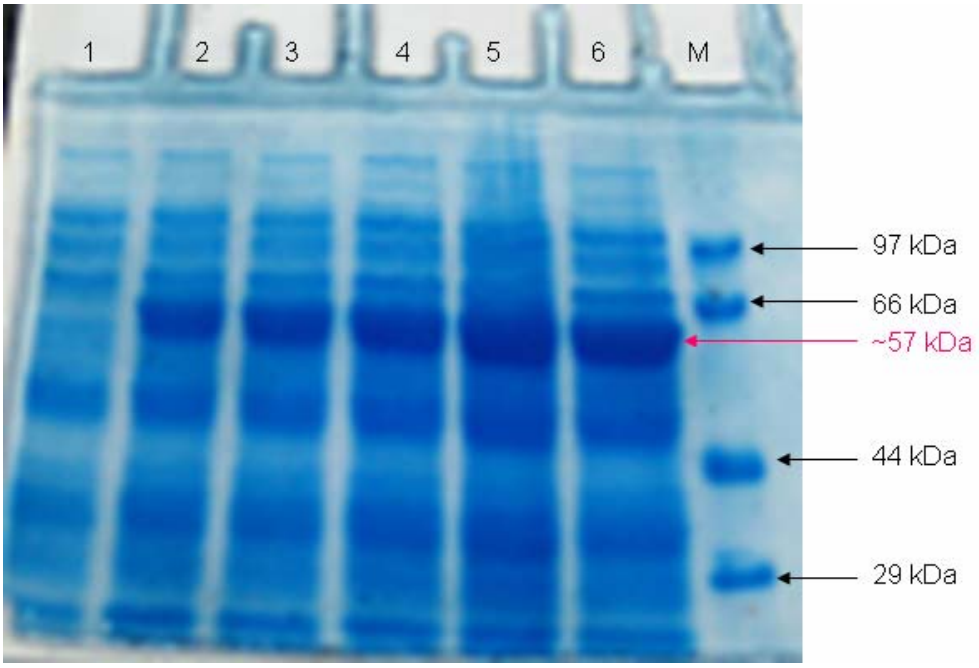


Figure 4A.6. 10% SDS-PAGE depicting gradual increase in expression level with time. Lane1 shows un-induced recombinant crude protein, lane 2-6 shows induced recombinant crude protein isolated at different time interval e.g. 1, 2, 3, 4 and 5 hrs respectively after IPTG induction. The last lane marked as M, contains low molecular weight protein marker ranging from 99, 66, 44, 30 and 14 kDa from top to bottom (the last band in the marker lane has migrated out due to the over running and thus not visible).

Once confirmed the protein expression in pilot experiments, the induction of protein expression was done at large scale. The protein isolation from inclusion bodies was done in denaturing conditions as the protocol prescribed in chapter 2; section 2.8.8.1. The protein was purified to homogeneity with Ni-NTA affinity chromatography as described in chapter 2; section 2.8.8.2. The LlCAld5H protein of ~57 kDa was separated under denaturing condition by 10% SDS-PAGE and visualized with Coomassie blue staining (Fig. 4A.7).

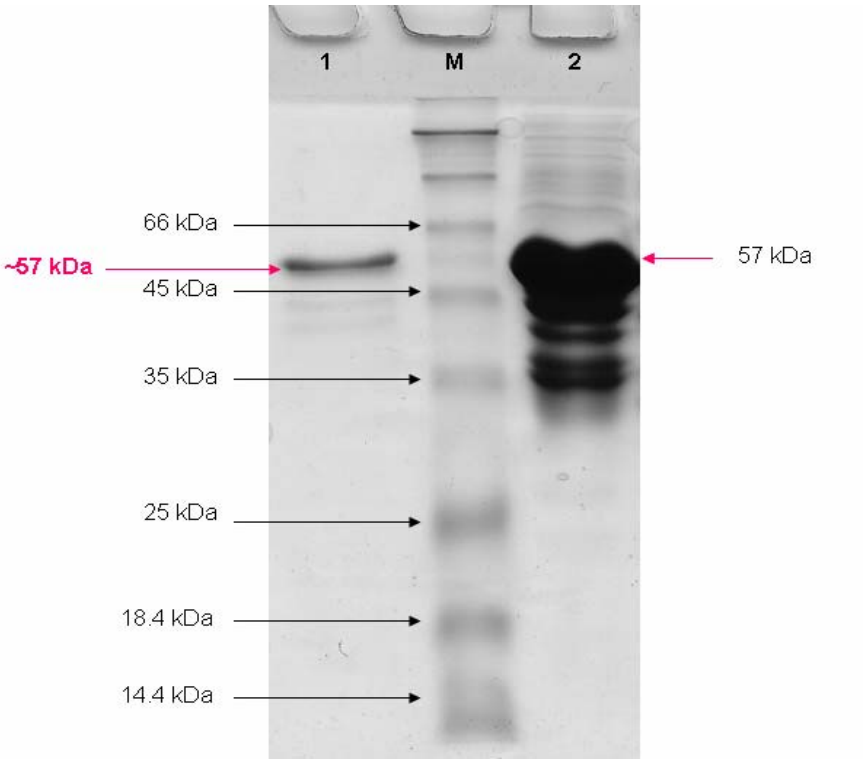


Figure 4A.7. The purified LlCAld5H protein migrated at ~57 kDa (lane 1) by 10% SDS-PAGE. The crude protein was separated in lane 2 and the protein marker in lane M.

4A.3.3 Generation of polyclonal antibody and IgG purification

200 µg of purified LlCAld5H protein (denatured conditions) was emulsified with an equal volume of Freund's complete adjuvant and injected into multiple subcutaneous sites on the back of New Zealand white rabbit. The injection was repeated 3 weeks later with 30 µg of protein. One week later the first bleed was taken out and the resulting

serum was tested to inhabit LlCAld5H activity by ELISA. The IgG fraction was purified by chromatography on recombinant protein A agarose (Sigma-Aldrich, USA).

4A.3.4 Western blot analysis

Different quantity of crude plant protein and recombinant protein samples were separated on 10% SDS-PAGE and transferred on nitrocellulose membrane (Fig. 4A.8) by semi-dried electro-blotting method as described previously in this chapter. Membrane was incubated with primary antibody and then with secondary antibody conjugated to chemiluminescent substrate. The blot was exposed to X-Ray film and developed in dark.

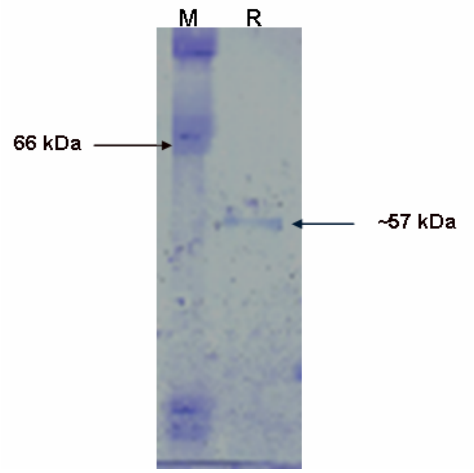


Figure 4A.8 The 57 kDa LlCAld5H protein blotted on nitrocellulose membrane (lane R).

The polyclonal anti-LlCAld5H antibodies reacted strongly with crude and purified recombinant LlCAld5H protein. Relatively small amount of CAld5H protein was detected in the crude extract of plant stem (Fig. 4A.9b lane S). However, the presence of one more protein of molecular weight ~80 kDa in *Leucaena* stem tissue was detected , which can be justified as the antibody used in the present study is polyclonal and the epitope of conserved domain of the anti-LlCAld5H could be able to cross react with another cytochrome-P450. However, a P450 protein of the molecular weight of ~80 kDa has been reported in Douglas-fir (Timothy J. Tranbarger et al., 2000), which strengthen

the above possibility. The pre-immune IgG had neither reacted with recombinant protein nor with plant crude extract.

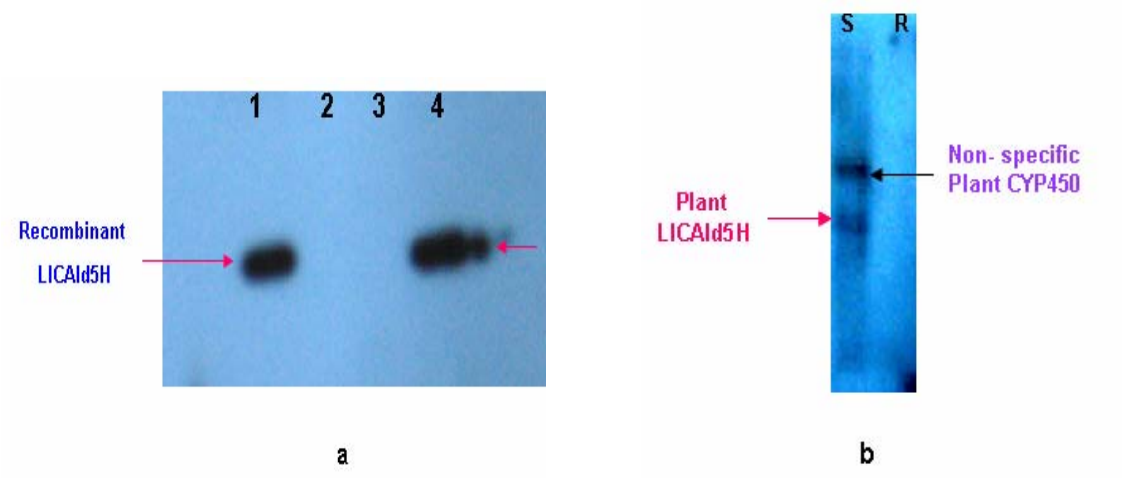


Figure 4A. 9 a) Western blot analysis of 1 µg of purified *E. coli*-expressed LlCAld5H recombinant protein (lane 1) and 20 µg of crude recombinant protein (lane 4). As a negative control, 10 and 20 µg of crude protein extracts of non-recombinant *E.coli* prepared in similar experimental methods (lane 2 and 3 respectively). **b**) 40 µg of crude protein extract of stem (xylem tissue) of *Leucaena leucocephala* in lane marked as S. Lane marked as R contain crude protein extract of root of *Leucaena*.

4A.3.5 Optimizing expression of active LlCAld5H protein

To obtain correctly folded recombinant protein, the bacterial cells harboring the LlCAld5H gene were grown in Terrific broth supplemented with heme precursor, ALA (1 mM) and $Fecl_3$ (0.5 mM). Different conditions were used in an attempt to optimize the amount of active LlCAld5H in *E. coli*. Initial studies were carried out with *E.coli* BL21 (DE3) and were aimed at determining the optimal IPTG concentration for induction of LlCAld5H. The optimal level of IPTG was 0.3 mM and effective for induction when added at mid-log phase (optical density at 600 nm, 0.6-0.8). Time course studies revealed that the maximum yield of LlCAld5H in soluble-fraction as determined by SDS-PAGE and ELISA profile performed with Anti-LlCAld5H was obtained after 24-30 h of induction at 25 °C. The activity was detected at 20-25˚C and at 0.03 mM IPTG but for

24-30 hrs of incubation. The specific activity of crude protein is found to be 109 U / mg (Unit definition- amount of enzyme required to oxidize 1nmol of NADPH per min, under standard assay condition). Prior to initialization of enzyme assay the auto-oxidation of NADPH was checked. The enzyme assay was set up as shown in chapter 2; table 2.14. However, the effort to purify the protein in native form through Ni-NTA agarose chromatography had not been successfully accomplished that limits the evaluation of specific activity and kinetic studies.

4A.3.6 Molecular modeling of LlCAld5H

The deduced protein sequence was submitted to the NCBI BLAST tool (http://blast.ncbi.nlm.nih.gov/Blast.cgi). This resulted in identification of 19 structures showing varying levels of homology with the LlCAld5H protein. Significant among them were 2HI4 (Cytochrome P450 1A2), 3DL9 (CYTOCHROME P450 2R1), 3G5N (CYTOCHROME P450 2B4), 2P85 (Cytochrome P450 2A13), 2V0M (CYTOCHROME P450 3A4) and 2FDU (MICROSOMAL P450 2A) showing 27%, 26%, 24%, 24%, 25%, 24% sequence identity. These structures were selected for further analysis on the basis of their sequence identity, e-score, source organism and sequence length.

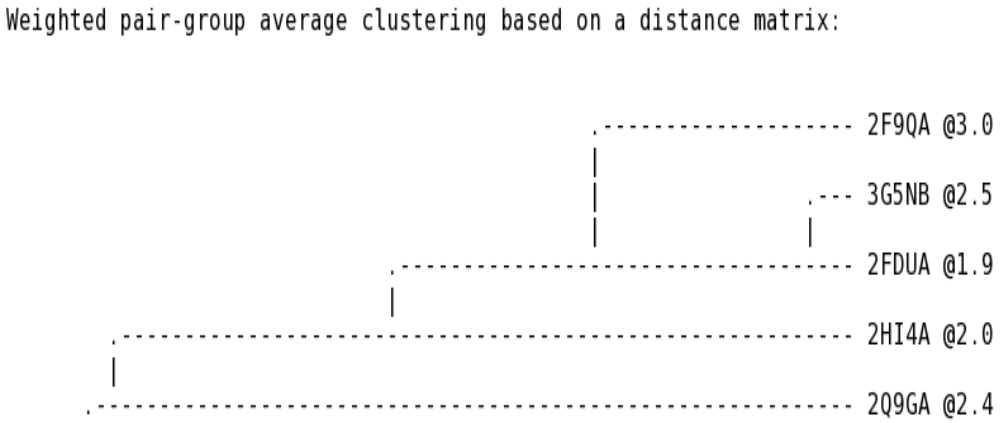


Figure 4A.10. Distance Matrix used for selecting the templates

```
| _aln.pos         10        20        30        40        50        60
2HI4A     ----------------RVPK----GL-----KSPPEP---WGWPLLGHVLTLGKNPHLALSRMSQRY
cal5d     MDSLLLHLQPLAMALLFAIPLLLLLGLLSRFRKRPPYPPGPKGLPIIGNMAMMDQLTHRGLAKLANLY
 _consrvd                     *     **     * ** *    * *  *         *  *      *

 _aln.p   70        80        90       100       110       120       130
2HI4A     GDVLQIRIGSTPVLVLSRLDTIRQALVRQGDDFKGRPD-LYTSTLITDGQSLTFSTDSGPVWAARRRL
cal5d     GGILHLRMGFLHMVAISDADSARQVLQVHDNIFSNRPATIAISYLTYVRADMAFA-HYGPFWRQMRKI
 _consrvd *  *  * *       *  *  ** *      *  **     * *        *    ** *   *

 _aln.pos  140       150       160       170       180       190       200
2HI4A     AQNALNTFSIASDPASSSSCYLEEHVSKEAKALISRLQELMAGPGHFDPYNQVVVSVA-NVIGAMCFG
cal5d     CVMKLFSRKRAE------S---WQSVQDEVETVVRTVG---ANTGKEVNIGELVFSLTKNITYRAAFG
 _consrvd     *     *       *      *  *            *  *        * *   *      **

 _aln.pos    210       220       230       240       250       260       270
2HI4A     QHFPESSDEMLSLVKNTHEFVETASSGNPLDFFPILRYL-PNPALQRFKAFNQRFLWFLQKTVQEHYQ
cal5d     SSSREGQDEFIGI---LQEFSKLFGAFNIADFIPSLSWVDPQGLNARLAKARGSLDSFIDKIIDEHMQ
 _consrvd     *  **          **       *  ** * *    *     *          *  *   ** *

 _aln.pos      280       290       300       310       320       330       340
2HI4A     DFDK-NSVRDITGALFK-HSKKGPRAS--GNL-----IPQEKIVNLVNDIFGAGFDTVTTAISWSLMY
cal5d     NKKKEDEEGDMVDELLAFYSEDHSKVSESDDLHNSIKLTKDNIKAIIMDVMFGGTETVASAIEWAMAE
 _consrvd    *     *    *    *      *    *          *     *    *  **  ** *

 _aln.pos        350       360       370       380       390       400
2HI4A     LVTKPEIQRKIQKELDTVIGRERRPRLSDRPQLPYLEAFILETFRHSSFLPFTIPHSTTRDTTLNGFY
cal5d     LMRSPEDLKRVQRELAEVVGLDRRVQESDMEKLTYLKCALKETLRLHPPIPLLL-HETAEDAVVSGYF
 _consrvd *   **     * **  * *  **   **   * **     ** *     *    * *  *    *

 _aln.p  410       420       430       440       450       460       470
2HI4A     IPKKCCVFVNQWQVNHDPELWEDPSEFRPERFLTADGTAINKPLSEKMMLFGMGKRRCIGEVLAKWEI
cal5d     VPKKSRVMINAWAIGRDRNAWEDPDSFKPARFL-GEGVPDFKGSNFEFIPFGSGRRSCPGMQLGLYAL
 _consrvd  ***  *  * *    *   ****  * * ***   *    *        ** * * * *  *

 _aln.pos  480       490       500       510       520
2HI4A     FLFLAILLQQLEFSVPPGVK---VDLTPIYGLTMKHA-RCEHVQARRFS----
cal5d     EMAVAHLLHCFSWELPDGMKPSEMDMSDVFGLTAPRASRLVAVSRKRLVCPLF
 _consrvd     * **       * * *    *     ***   * *   *   *
```

Figure 4A.11. Alignment of the target sequence of coniferaldehyde 5 -hydroxylase with the sequence of the template 2 HI4 (Human Microsomal Cytochrome P450 IA2).

The BLAST search thus gave a precise idea of the type of templates which could be used for building the model. Next the search module of the MODELLER9v7 was used to select the final templates to be used in model building. The module uses pir_95 as the structure sequence database which contains sequences of pdb structures clustered at 95% identity. This resulted in refinement of the initial search giving more templates. Based on the distance matrix 2FDU, 2F9Q, 2HI4, 2Q9G, 3G5N were selected for further analysis (Fig. 4A.10). Using ClustalW a multiple alignment of these templates with the target protein was generated (Figure 4A.11). The phylogenetic tree generated was used as a

guide to identify the final template which would be employed to build the model. Based on the preliminary analysis the structure of the protein Human Microsomal Cytochrome P450 1A2 (2HI4) was selected for use as the reference structure in modelling coniferaldehyde 5 – hydroxylase. The co-ordinates of the reference structure were assigned to the target sequence based on satisfactory spatial restraints.

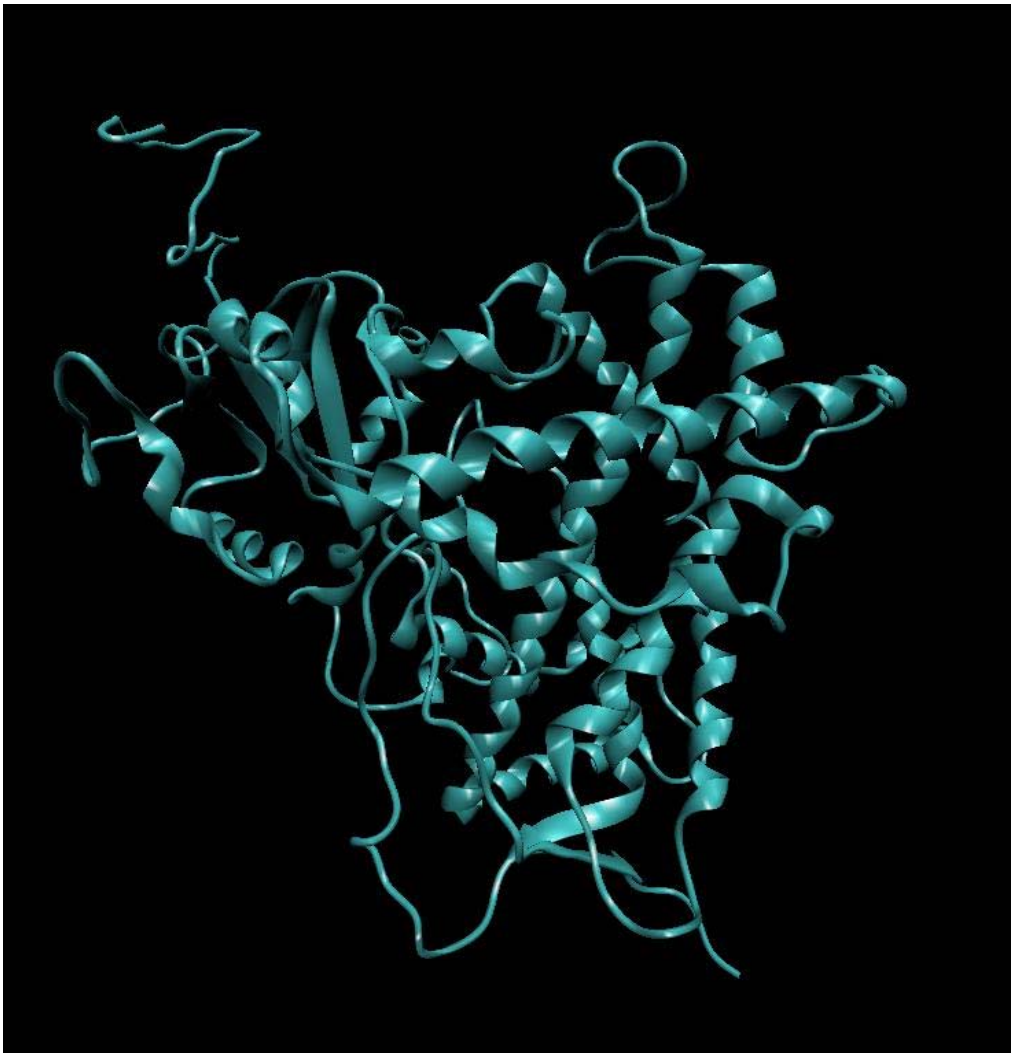


Figure 4A.12. Model of the coniferaldehyde 5-hydroxylase

From the five models generated, the model with the least value of MODELLER objective function was selected for further analysis. The model was first superimposed on the template and the RMSD was calculated as 0.64. The RMSD value less than 1.00 showed that although the identity of the template was less than 30%, the model built (Fig. 4A.12) was good in all the domains modeled in the predicted structure (Fig.4A.13).

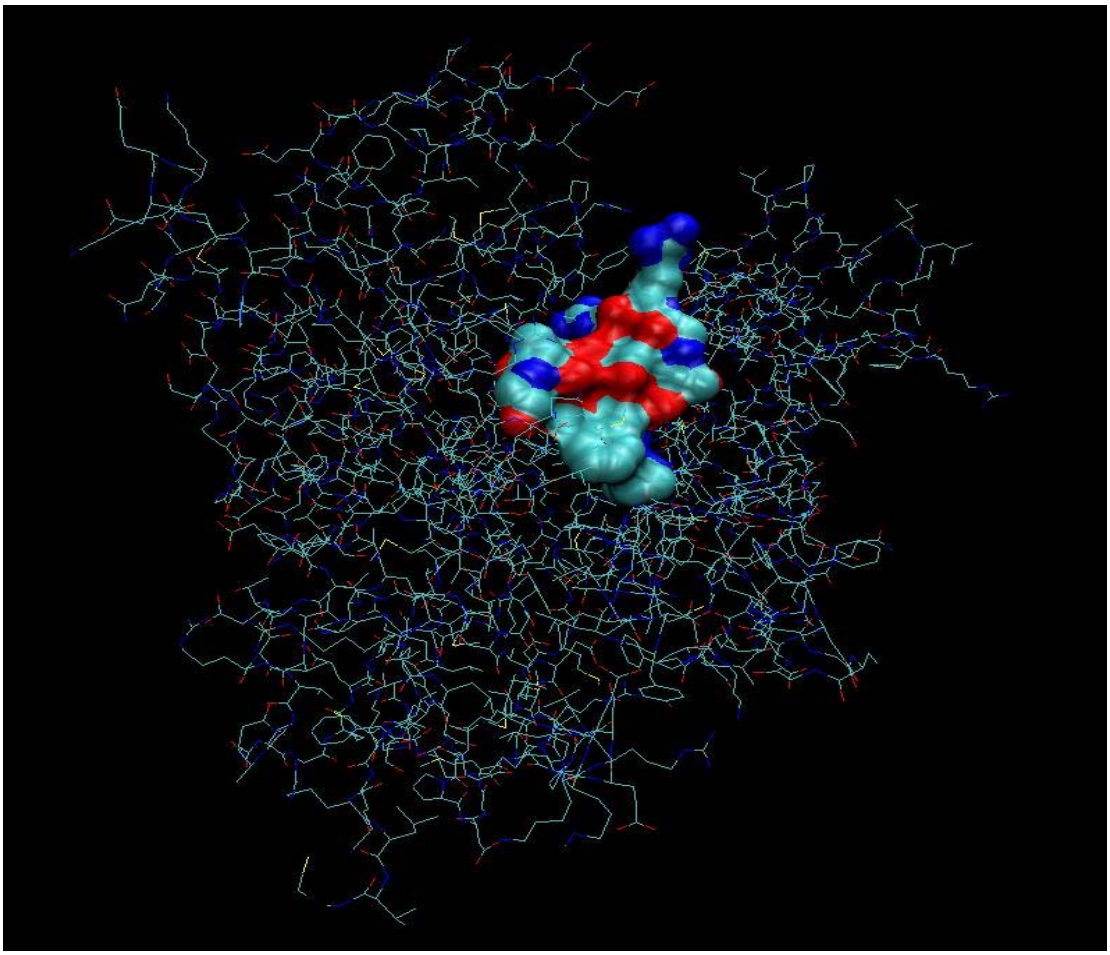


Figure 4A.13. Structure of coniferaldehyde 5 -hydroxylase highlighting the heme binding domain.

The model was then submitted to PROCHECK to assess the geometry. The PROCHECK analysis showed 88.7%, 9.1%, 2.0% residues in most favorable, additionally allowed and generously allowed regions, respectively, of the Ramachandran map (Fig. 4A.14). Only one residue was found to lie in the disallowed region. The comparatively good

stereochemical quality of the predicted model highlights the fact that the overall structure in case of the cytochrome P450 related proteins remain conserved, and as a result the structure built could be correctly modeled on the basis of the template chosen. The overall PROCHECK G-factor was -0.11. The environment profile of Verify-3D was found to be greater than 1 for the residues which were actually template-based modeled (i.e. residues 27-499). Further evaluation of the model by WHATCHECK program predicted the RMS Z score of backbone contacts as -1.23, backbone-side chain contacts as -2.50, side chain- backbone contacts as -3.45 and side chain-side chain contacts as -2.85.

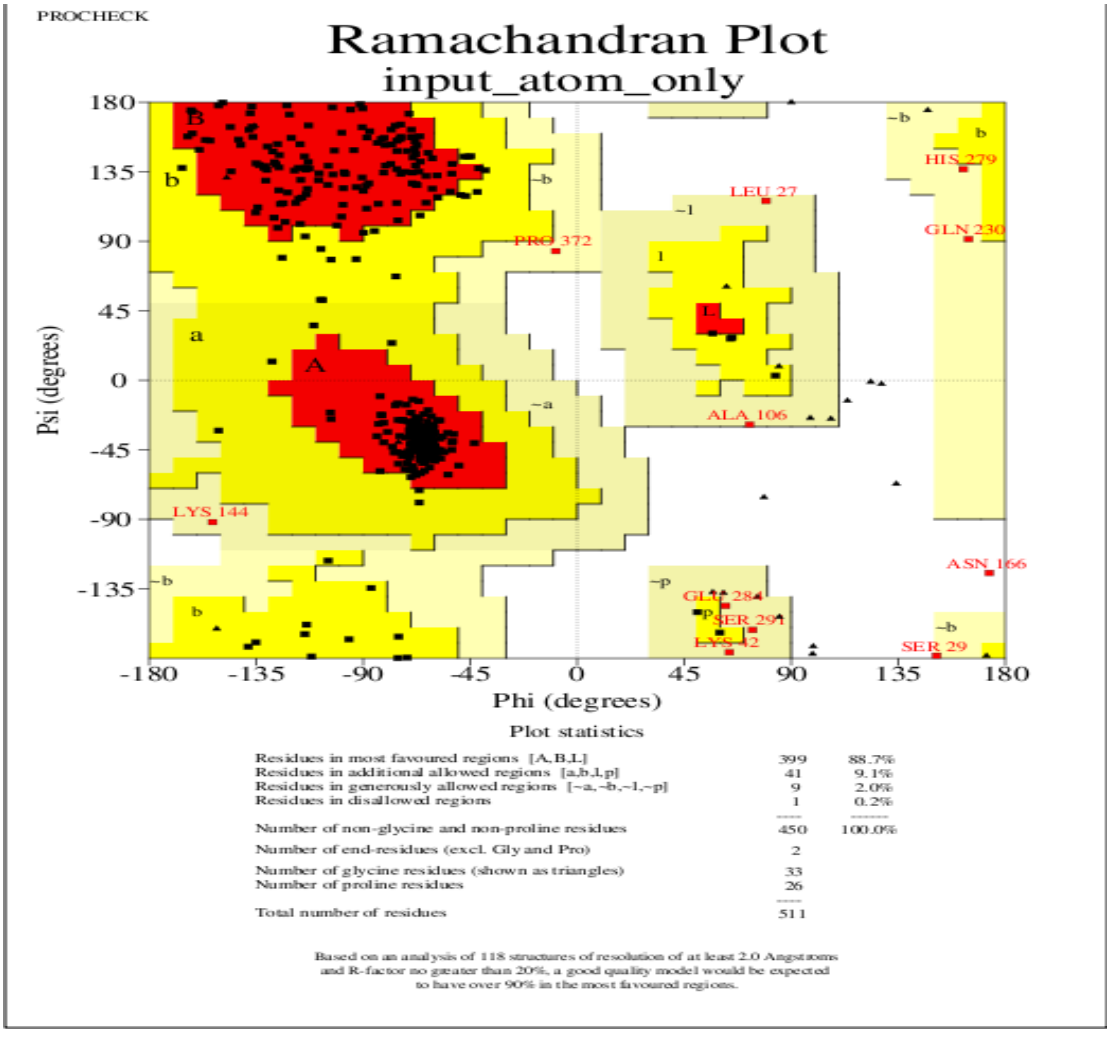


Figure 4A.14. All residues plotted in the Ramachandran map for the model built.

Next the analysis was focused on the heme-binding domain. The heme binding domain of the model was superimposed on the template (Fig.4A.14) and the variations of the surrounding residues were identified. Also the binding sites of other cytochrome P450 proteins of other organisms were superimposed and the residue variations were recorded. The residues Phe, Gly, Arg, Cys, Gly in 2^{nd}, 5^{th}, 7^{th}, 9^{th}, and 11^{th} position of heme binding domain are conserved (Fig.4A.15). The analysis of the interactions at the binding site has been done and this has been provided a plethora of information (described in the discussion section).

Table 4A.1. Variation of residues in the heme binding domain of cytochrome P450 from various organisms

Protein	Residue Position in Sequence	Organism	Residues in the Heme Binding Site										
Coniferaldehyde – 5 - Hydroxylase	440-450	LEUCEANA LEUCOCEPHALA	P	F	G	S	G	R	R	S	C	P	G
CYTOCHROME P450 1A2	450-460	HOMO SAPIENS	L	F	G	M	G	K	R	R	C	I	G
CYTOCHROME P450 2B4	444-454	ORYCTOLAGUS CUNICULUS (Rabbit)	P	F	S	L	G	K	R	I	C	L	G
CYTOCHROME P450 102	392-402	BACILLUS MEGATERIUM	P	F	G	N	G	N	R	A	C	I	G

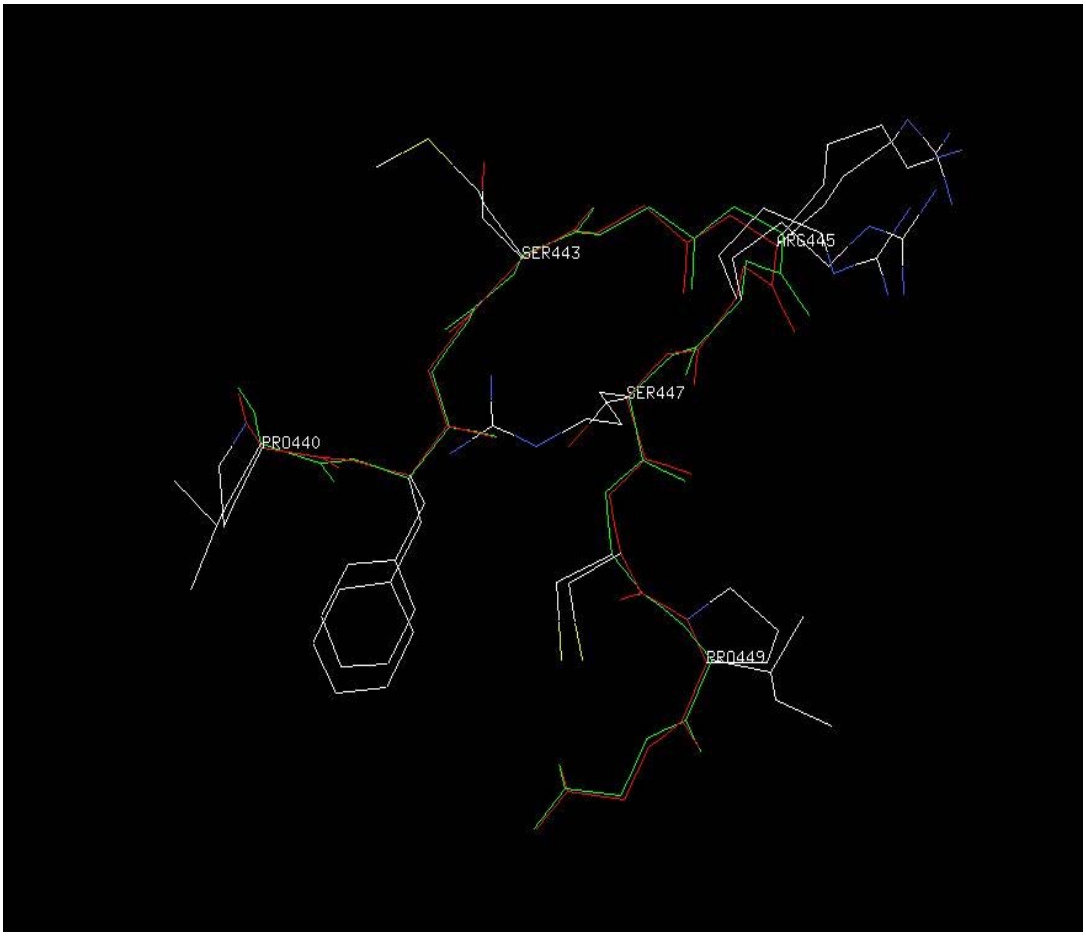


Figure 4A.15. Superposition of the residues at the binding site of modeled protein with the template 2HI4

4A.3.6 Discussion

Leucaena leucocephala contains a coniferaldehyde-5-hydroxylase protein as confirmed by Western blotting. This enzyme belongs to plant cytochrome P450s and it is categorized as **CYP84A40** (nomenclature was done by David Nelson, on request as and when amino acid sequence was sent to him. He keeps an update of P450s and he is also been associated to P450 nomenclature committee).

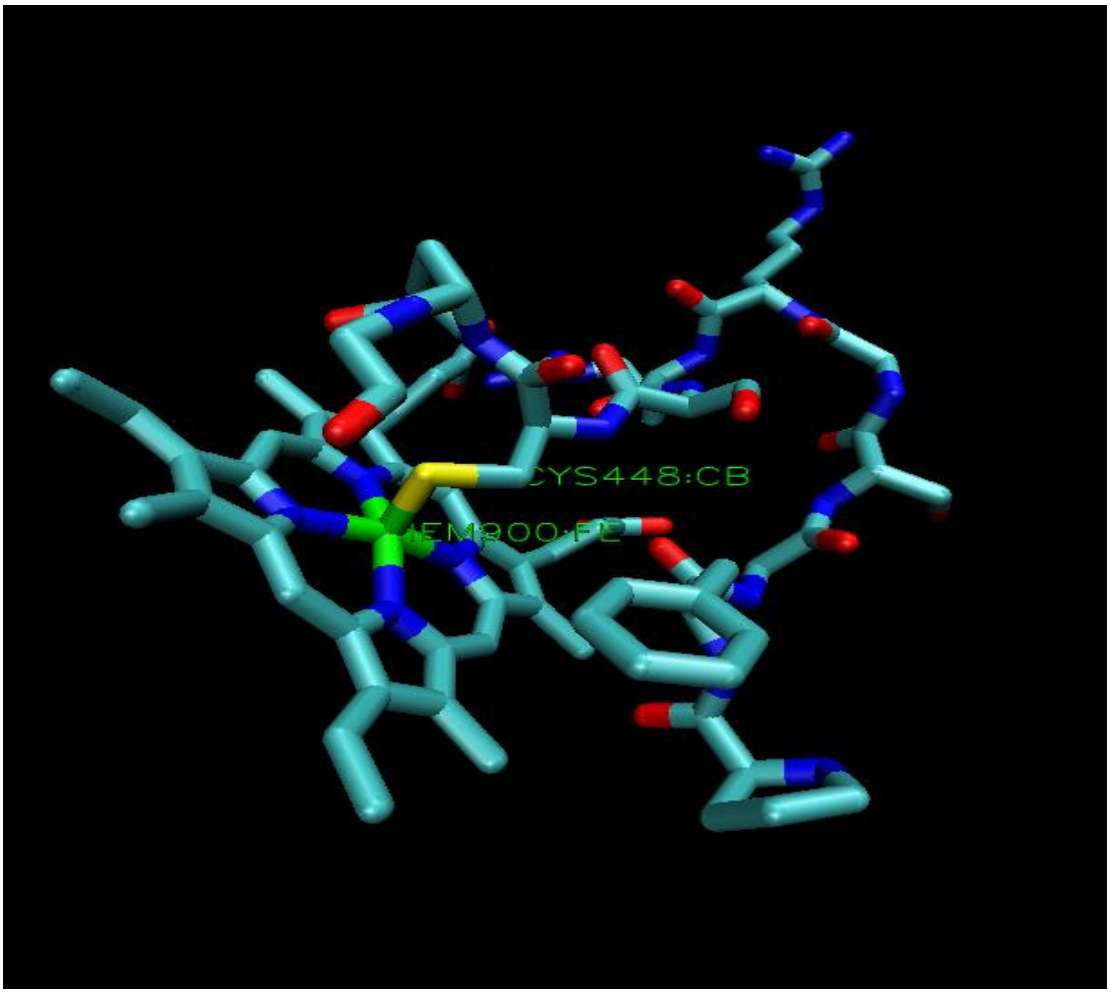


Figure 4A.16. Heme Binding Domain of the model showing the heme group bound to cysteine. Some of the conserved residues near the binding site are also shown.

A series of modifications were required in growth conditions and the choice of expression vector. Adding δ-aminolevulinic acid (ALA) and 0.5 mM $FeCl_3$ to the growth medium (Terrific broth) presumably increases the expression level of the protein as the addition of the ALA to the culture medium stimulates heme biosynthesis to match the increase in P450 polypeptide synthesis (Imai, T. *et al*., 1993). It has been observed previously that such an increase of P450 suggest that the rate limiting step in the synthesis of hemoprotein lies in the synthesis of the heme prosthetic group (Richardson, T.H., 1995; Jansson, I. *et al*., 2000). Therefore, it will be helpful in the formation of correctly folded holoenzyme with the heme group in its proper protein environment

which increases when ALA and $Fecl_3$ was added. The maximum expression level was at 24-30 h after IPTG induction.

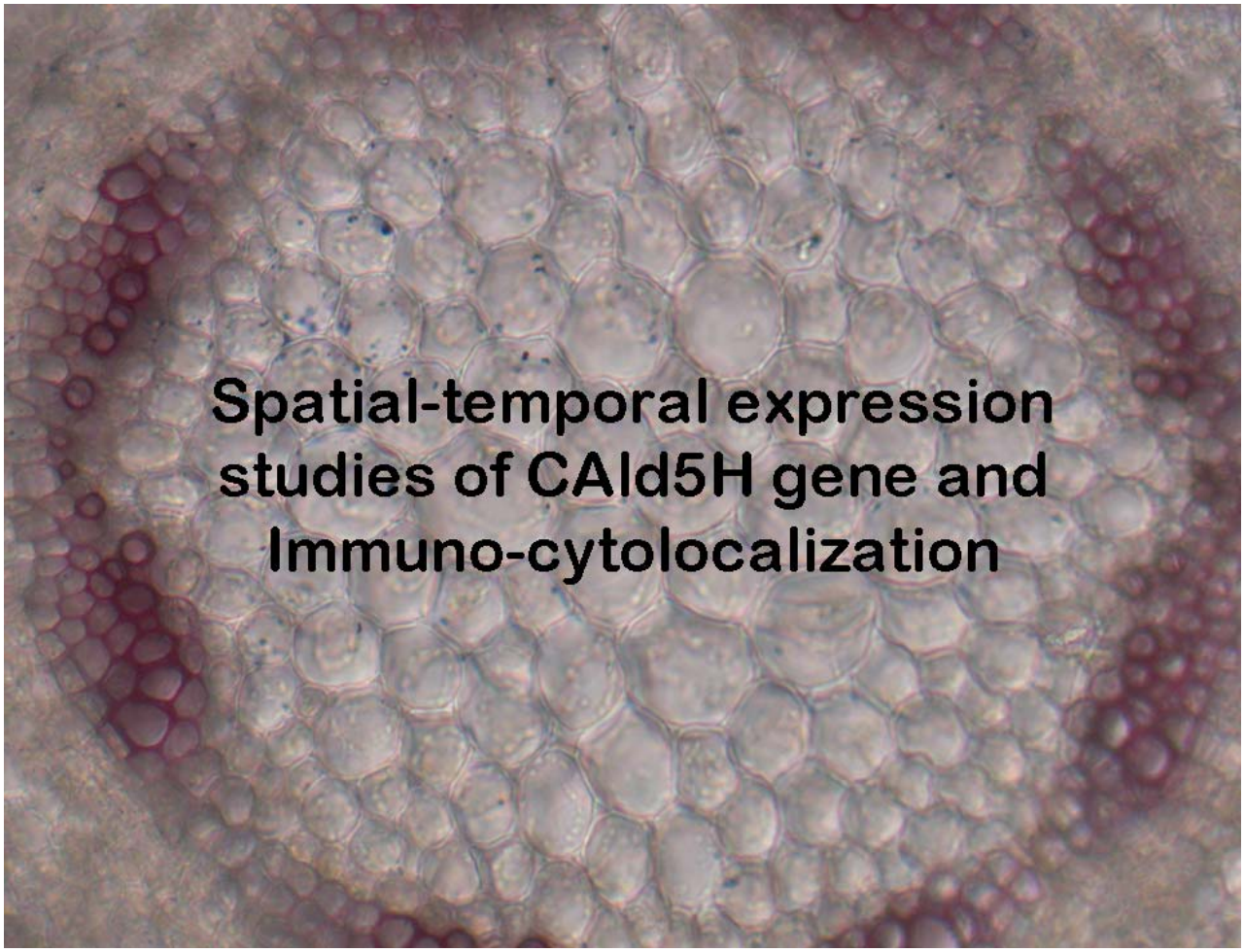
Though, the abundance in hydrophobic amino acids in the sequence makes it difficult to obtain the correctly folded protein in prokaryotic system, expression of the protein in native form could be obtained by optimizing the above mentioned parameters. The LlCAld5H activity was detected in the crude extract. The enzyme assay was done by monitoring the reduction of NADPH to NADP. NADPH is the cofactor of this enzyme and acts as an electron donor to catalyze the molecular oxygen, leading to the oxidative attack on coniferaldehyde (substrate of this enzyme) and coverts it into 5-hydroxyconiferaldehyde. The reaction catalyzed by this enzyme is basically hydroxylation reaction but it can utilize the other substrates like coniferyl alcohol and ferulic acid as well.

Based on the Blast search and the search module of the MODELLERv7 the final template was selected that can be used to build a model for LlCAld5H protein. The model was first superimposed on the template and the RMSD was calculated as 0.64. The RMSD value less than 1.00 showed that although the identity of the template was less than 30%, the model built was good in all the domains modeled in the predicted structure. The PROCHECK analysis showed 88.7%, 9.1%, 2.0% residues in most favorable, additionally allowed and generously allowed regions, respectively, of the Ramachandran map. Only one residue was found to lie in the disallowed region. The residues Phe, Gly, Arg, Cys, Gly in 2^{nd}, 5^{th}, 7^{th}, 9^{th}, and 11^{th} position of heme binding domain are conserved. The superposition of the model with the template also revealed that along with the residues in the heme binding site other residues namely, Arg108, Trp129, Arg133 and His379 interacting with the heme group also might play a crucial role in the binding of the substrate at the heme moity. The analysis of the interactions at the ligand binding site revealed that the hydrogen bond between Phe441 and Cys 448 is crucial in maintaining the structure of the domain (Fig.4A.16). Also it was observed that Arg 446 forms a hydrogen bond with the heme group which could be important for the substrate binding. The H-bonds observed to form with the Cys448 residue highlight the fact that it is a crucial residue at the heme binding site.

4A.4 Conclusion

- LlCAld5H sequence encoding the mature coniferaldehyde-5-hydroxylase protein was cloned in pET 28a (+) expression vector.
- The recombinant protein was purified in denatured condition and used to raise polyclonal antibody against it.
- The candidate protein, coniferaldehyde-5-hydroxylase was detected in *Leucaena leucocephala* by Western blotting.
- Optimization of expression of native protein was done and the best expression was achieved when bacterial cultures were grown in Terrific broth for 24-30 h after induction with 0.03 mM IPTG at 25°C. Specific activity with crude extract was estimated.
- The homology modeling of LlCAld5H protein sequence was done based on the template Human Microsomal Cytochrome P450 1A2 (2HI4).
- The RMSD value less than 1.00 showed that although the identity of the template was less than 30%, the model built was good in all the domains modeled in the predicted structure.
- The PROCHECK analysis showed 88.7%, 9.1%, 2.0% residues in most favorable, additionally allowed and generously allowed regions, respectively, of the Ramachandran map. Only one residue was found to lie in the disallowed region.
- The analysis of the interactions at the ligand binding site revealed that the Hydrogen bond between Phe441 and Cys 448 is crucial in maintaining the structure of the domain.

Chapter 4B


Spatial-temporal expression
studies of CAld5H gene and
Immuno-cytolocalization

To get an insight of developmental regulation of coniferaldehyde 5-hydroxylase gene expression, the spatio-temporal expression and immunocytolocalization studies in different age group of *in-vitro* grown seedlings of *Leucaena leucocephala* was done. This section describes importance of this study in *Leucaena*, methods and techniques used, followed by results obtained and finally the analysis of the results and conclusions.

4B.1 Introduction

Lignin in gymnosperms is polymerized from the guaiacyl monolignols. In angiosperms, the syringyl monolignol branches out from the guaiacyl pathway and polymerizes with guaiacyl monolignols to form a heterologous guaiacyl-syringyl lignin. However, the regulation of this significant evolutionary event has remained largely unclear for the past five decades mainly because of a lack of fundamental biochemical and genetic knowledge of syringyl monolignol biosynthesis. But the recent biochemical studies revealed that the branch pathway for syringyl monolignol biosynthesis is regulated in sequence by three genes encoding coniferaldehyde 5-hydroxylase (CAld5H), 5-hydroxyconiferladehyde O-methyltransferase (AldOMT), and sinapyl-alcohol-dehydrogenase. The regulation of phenylpropanoid biosynthesis genes, particularly the first and last step of the ‘‘general phenylpropanoid’’ pathway, catalyzed by phenylalanine ammonia-lyase (PAL) and coumaryl-CoA ligase (4CL), have been extensively studied. The considerable amount of data accumulated on the regulation of these genes. Moreover, the studies based on RT-PCR revealed that *PopF5H* is expressed in xylem, bark and young leaves and, to a lower extent, in mature leaves of *Populus tricocharpa* (Richard Sibout *et al.,* 2002). This pattern of expression has also been reported in *A. thaliana* and *Brassica napus*. However, the developmental regulation of this gene responsible for syringyl lignin synthesis has been very poorly studied.

Ruggers *et al.,* (1999) have observed a high AtF5H/CAld5H expression level in *A. thaliana* roots without any significant content of the sinapate esters. This result reveals that F5H/CAld5H activity is limiting for the production of S-units in the roots. This discrepancy suggests that F5H/CAld5H might be involved in other biosynthetic routes than the one leading to sinapate esters, S lignin units. This observation suggests a role of other molecules such as lignans. Therefore, the tissue specific expression studies of the

gene with reference to root, stem and leaf would reflect an clear image of the expression of the gene and that would be attributed to the functional expression and the ultimate effect of the gene on lignin monomer biosynthesis. Moreover, no such study has been reported in *Leucaena leucocephala*. Therefore, the spatio-temporal studies in the *Leucaena leucocephala* will definitely provide a strategy of interest to increase the S-unit in lignin that makes lignin more vulnerable to the kraft industrial delignification process.

4B.1.1 Introduction of Real-Time quantitative PCR

The use of real-time PCR has increased over last one decade as a fast and precise method of quantifying nucleic acids in nearly every area of biological research connected to all organism kingdoms (Brazeau, 2004; Bubner and Baldwin, 2004; Bustin and Nolan, 2004; Kurrasch *et al.,* 2004; Mackay *et al.,* 2002). However, the measurement of transcripts can be done with Northern blot analysis, but is not the most suitable experimental approach to semi-quantify the RNA especially when very low concentrations are expected. Therefore, real-time PCR is the method of choice when exact quantification is sought, in conjunction with high sensitivity (Bustin, 2000; Bustin 2002; Mackay, 2004). Nevertheless, the combined features of a slightly high G/C content found in LlCAld5H gene (57%) and the contamination of polyphenolics and polysaccharides in the nucleic acid preparation from plants in general demand an adjustment of plethora of protocols published elsewhere.

4B.2 Materials and method

Plant materials: *in- vitro* grown *Leucaena leucocephala* cultivar K-636 of different age group was used for the spatio-temporal as well as immunocytolocalization studies.

4B.2.1 RNA extraction and reverse transcription (RT)

Total RNA was extracted from root, shoot and leaf of different age group i.e. 0, 5, 10, 20 and 30 days old *in-vitro* grown *Leucaena leucocephala* seedlings by TRIzol (Invitrogen)/Biozol (Bioworld) method. Deviations from standard protocol according to the manufacturer's recommendations were found to be necessary. A heating step was introduced which enhanced the net RNA yield, and it deemed necessary to precipitate the

RNA with a high salt solution to discriminate against polysaccharides. The details of the protocol have been described in chapter 2; section 2.8.4.5. RNA quality was analyzed by formaldehyde-denatured gel electrophoresis as described in chapter 3; section 3.3.3.1. DNase I digestion of all RNA samples was executed as follows: 2 µg RNA was digested with 2 U of DNase I (Sigma–Aldrich) for 15 min at 25 °C using the supplied buffer. Following the DNase I digest, 1 µl of 50 mM EDTA was added and heat inactivation was performed for 10 min at 65 °C. The lack of DNA traces in the RNA sample was checked by using 0.4 µl of the DNase I digest as template in a PCR reaction, in comparison to a positive control PCR. The RNA was analyzed for its quantity and quality by Bioanalyser 2100 (Agilent technology, USA) system. Equal quantity (1 µg) of RNA extracts from each samples were used individually to synthesize the first strand cDNA as explained in chapter 2; section 2.8.4.8. The random hexamers primers were used as it produces the least bias cDNA. Reverse transcription reaction was done in a final volume of 20 µl using first strand cDNA synthesis kit from Invitrogen/Promega.

4B.2.2 Real-time PCR

The cDNA first strands were used as template for semi and absolute quantification of CAld5H transcripts. 'Brilliant II™ SYBR® Green QPCR master mix (Stratagene, Agilent Technology, TX, USA) was employed, unless stated otherwise, for real-time PCR. The reaction mixture was prepared as described in chapter 2; section 2.8.4.11.2. The reference dye was diluted (1:500) as recommended, using nuclease-free PCR grade water resulting in a final reference dye concentration of 30 nM in each reaction. Cycling was performed as described in chapter 2; section 2.8.4.11.3.

4B.2.3 Relative and absolute quantification methods

Relative (comparative) quantification relates the PCR signal of the target transcript in a treatment group to that of another such as untreated control.

The comparative C_t method is also known as the $2^{-\Delta\Delta Ct}$ method, where $\Delta\Delta C_t = \Delta C_{t,\,sample} - \Delta C_{t,\,reference}$

Here, $\Delta C_{T,\ sample}$ is the C_t value for any sample normalized to the endogenous housekeeping gene and $\Delta C_{t,\ reference}$ is the C_t value for the calibrator also normalized to the endogenous housekeeping gene.

Means,

$\Delta C_{t,\ sample} = Ct_{\ sample} - Ct_{\ endogenous\ control\ (18s\ rRNA\ gene\ or\ any\ other\ housekeeping\ gene)}$

$\Delta C_{t,\ reference} = Ct_{\ reference\ or\ Normal/untreated\ sample} - Ct_{\ endogenous\ control}$

For the $\Delta\Delta C_t$ calculation to be valid, the amplification efficiencies of the target and the endogenous reference must be approximately equal. This can be established by looking at how ΔC_t varies with template dilution. If the plot of cDNA dilution versus delta C_t is close to zero, it implies that the efficiencies of the target and housekeeping genes are very similar. If a housekeeping gene cannot be found whose amplification efficiency is similar to the target, then the standard curve method is preferred.

4B.2.4 SYBR® Green real-time PCR

All real-time PCR measurements were run on Stratagene Mx3005P thermal cycler (Stratagene, Agilent Technology, TX, USA). Real-time PCR efficiency was measured in between 0.75 to 0.9 which corresponds to a successful amplification of 75-90% of the template. (In the case of 100% efficiency during the PCR reaction, there will be a doubling of the amount of DNA at each cycle corresponding to 1.0 (or 100%). As a normalizing signal for real-time PCR measurements, 18S rRNA representing a housekeeping gene was used (Dhar, A.K. *et al.,* 2009).

The lists of primers and probes sequences that were designed for *Leucaena CAld5H* gene and for the 18S rRNA gene as an internal standard are given in Table 4B.1. Optimal numbers of PCR cycles within the linear range of amplification for each gene were determined in preliminary experiments.

4B.3 Results and Discussion

4B.3.1 Optimization of molar concentration of primers

Each primer pair was tested before being used in a real-time measurement for ideal molar concentration of forward and reverse primers to each other. The test parameters were defined as the influence of the primer concentration on the reaction efficiency given that by-products (primer-dimer) did not occur. Used primers sequences along with their annealing temperature, GC content, amplicon length and optimum molar concentrations are listed in table 4B.1.

Table 4B.1 Details of the primers used in qRT-PCR

Primer Name	Primer sequence (5'-> 3')	G/C content	Molar conc. (nM)	Annealing temperature (°C)
CAld5H_RTF	CTCCACATGGTAGCCATCTCC	57	75	57.2
CAld5H_RTR	GCCCGGTCGTAGGTGAGGTAG	66	75	57.3
CAld5H_RTFo	CTCCGGATGGGCTTCCTCCAC	66	75	59
CAld5H_RTRo	ATGTCGGCCCGGACGTAG GT	65	75	59
CAld5H_RTFn	ACGGCGGCATACTCCACCTC	65	75	61
CAld5H_RTRn	GTGAGGTAGCTGATGGCTAT	50	75	55
18s_RTF	GCACGCGCGCTACAATGAAA GTAT	50	75	57
18s_RTR	TGTGTACAAAGGGCAGGGAC GTAA	50	75	57.1

4B.3.2 Spatio-temporal expression studies of CAld5H gene in *Leucaena*

The standard quantitative RT-PCR was performed on Stratagene Mx3005P thermal cycler. The mean Ct value was calculated from three technical replicates of each sample (target as well as internal control gene) run individually. The mean Ct values obtained from all the samples along with the standard deviation have listed in the table 4B.2. Based on the Ct values, the expression patterns in different tissue types of different age groups were analyzed.

Table 4B.2 Mean Ct values obtained from different samples by qRT-PCR.

Tissues	Ct of untreated tissue	Ct of treated tissue
Rad- 0 d	25.96±0.98†	23.88±0.09
R-5 d	30.83±1.25	27.46±0.99
R- 10 d	30.72±1.75	27.92±0.30
R- 20 d	30.46±0.66	26.05±0.88
R- 30 d	24.76±0.9	22.25±0.41
S- 5 d	27.77±1.44	25.68±0.92
S- 10 d	25.30±2.65	23.90±0.98
S- 20 d	29.54±0.4	26.28±0.10
S- 30 d	24.79±1.05	22.35±1.70
L- 5 d	31.05±0.48	29.73±0.27
L- 10 d	30.28±1	28.00±0.63
L- 20 d	32.50±2.65	28.68±0.51
L- 30 d	29.54±0.96	26.89±0.25
18s rRNA from all samples	17.99±0.25	17.99±0.25
Anova *	Sig 1%	Sig 1%

† Standard deviation; *Anova =Analysis of variance; Sig =significant figures.

The RT-PCR studies based on the Ct values, revealed that LlCAld5H is expressed in radical (0 day), root, stem and young leaves. The extent of expression varies amongst different tissue type. Among the age group of 5 to 20 days, stem shows maximum CAld5H expression followed by root. However, the expression level of the candidate gene in root of 30 days old *Leucaena* seedlings matches with stem tissue of the same age. A unique pattern of expression is exhibited by stem and leaves as in both the tissue the expression is increasing from 5-10 days but the 20 days old plants showed the decline in expression followed by further increase in expression of 30 days samples. Amongst all the tissues selected for studies, the CAld5H expression is found to be comparatively less in leaves. An exquisite expression level of LlCAld5H was observed in the radical (considered as 0 day) even more than root and leaves.

4B.3.3 Relative quantification of CAld5H transcripts of *Leucaena* grown on salt stress

The *Leucaena leucocephala* seeds were kept for growth on ½ MS medium as described previously with additional 100 mM NaCl supplement. The RNA extraction, cDNA preparation as well as the RT-qPCR was performed in similar way as the untreated plants. The relative quantification method was employed because it relates the PCR signal of the target transcript in a treatment group to that of another such as untreated control. This method involves comparing the C_t values of the samples of interest with a control or calibrator such as a non-treated sample or RNA from normal tissue. The C_t values of both the calibrator and the samples of interest are normalized to an appropriate endogenous housekeeping gene (18 s rRNA). The $2^{-\Delta\Delta Ct}$ method is a convenient way to analyze the relative changes in gene expression from real-time quantitative PCR experiments (Pfaffl MW. 2001). The $2^{-\Delta\Delta Ct}$ for all the samples were calculated as described earlier in this chapter and listed in the table. The $2^{-\Delta\Delta Ct}$ actually reflects the fold expression. The relative quantification analysis shows an increase in the LlCAld5H transcript in the salt treated samples without any exception. The fold changes in expression are varying among the different tissues types of different age groups as depicted in table 4B.3. The salt stress is capable to enhance the LlCAld5H expression by ~14 fold in 20 days root and 15 days old leaves. The minimum increase in expression has

been sought in 10 days stem with 2.63 fold as compared of untreated samples of same age group.

Table 4B.3 Relative Quantification of the transcripts abundance

Tissue and Day	*C_t Treated A	C_t reference B	ΔC_t treated sample C=(A-B)	C_t untreated (baseline) D	ΔC_t Treated E=(D-B)	$\Delta\Delta C_t$ F=(C-E)	$2^{-\Delta\Delta Ct}$ G=2^{-F} (Fold change)
Rad- 0 d	23.87	18.32 †	5.55	25.95	7.66	-2.11	4.31
R- 5 d	27.47	18.24	9.23	30.83	12.59	-3.36	10.26
R-10 d	27.92	18.19	9.73	30.74	12.55	-2.28	4.85
R- 20 d	26.28	18.23	8.05	30.12	11.89	-3.84	14.32
R- 30 d	22.30	17.95	4.35	24.79	6.84	-2.49	5.61
S- 5 d	25.66	17.78	7.88	27.79	10.01	-2.13	4.37
S-10 d	23.90	17.72	6.18	25.30	7.58	-1.4	2.63
S-20 d	26.18	17.45	8.73	29.54	12.09	-3.36	10.26
S-30 d	22.32	17.10	5.22	24.79	7.69	-2.47	5.54
L-5 d	29.09	18.68	10.41	31.00	12.32	-1.91	3.75
L-10 d	28.00	18.01	9.99	30.28	12.27	-2.28	4.85
L-15 d	28.62	18.34	10.28	32.50	14.16	-3.88	14.72
L-30 d	26.87	18.21	8.66	29.58	11.37	-2.71	6.54

† It was assumed that the chosen reference gene (18S rRNA) doesn't vary in copy number or expression level amongst the samples of study (as evident by the very little fluctuations in the Ct values obtained from amplification). Each Ct values are the average of three biological replicates run in separate PCR tubes. * The Ct values obtained from the plant samples grown on 100 mM salt stress. Rad, Radical; R, Root; S, Stem; L, leaf, d, day; 18s rRNA reference gene; Normal means untreated plant samples.

4B.3.4 Histology and Immunocytolocalization

The histology and immunocytolocalization study was done as described in chapter 2.8.9.

The stem and root sections stained with anti-LlCAld5H antibody and visualized under Axiocam vision microscope and analyzed for functional expression of the LlCAld5H and its developmental regulation.

4B.3.4.1 Immunocytolocalization

CAld5H protein was immuno-cytolocalized in tissues undergoing active lignification *i.e.* vascular bundle and xylem fibers as described in chapter 2; section 2.9. The section were incubated with antibodies followed by incubating with BCIP/NBT mix confirms the presence of CAld5H protein near the sites of lignification. In an attempt to stain the entire lignified tissues in the section the phloroglucinol staining was done. The antibody stained section could be used further to stain with phloroglucinol. The stained sections were visualized on AxioCam visison microscope and pictures were captured with the digital camera (Sony) attached with the microscope. The snaps were captured with different magnifications such as 2.5x, 5x, 10x, 20x, 40x and 100x as shown in figures 4B. The LlCAld5H was immunolocalized in the transverse sections of root and stem tissue of *Leucaena leucocephala* seedlings of age group varies from 5 days to 20 days. The LlCAld5H expression has definitely been increased in 5 days to 20 days plant as evident from the figure 4B.1 and 4B.2. However, the extent of increase could not be accurately determined with this experiment. The pictures of the histochemical stained sections from 20 days old plants have not been shown. In an attempt to study the lignification pattern in stem and root, the transverse sections from these tissues were stained with phloroglucinol. This will stain the total lignin in general. The intensity of the stain in the vascular tissues has consistently been increased with age (Fig. 4B.1 and 2) that reflects the active lignification in these tissues. At 100x magnification the lignification in the S1, S2 and S3 layers of the cell wall could be clearly visualized (Figure 4B.1 g and 4B.2 j).

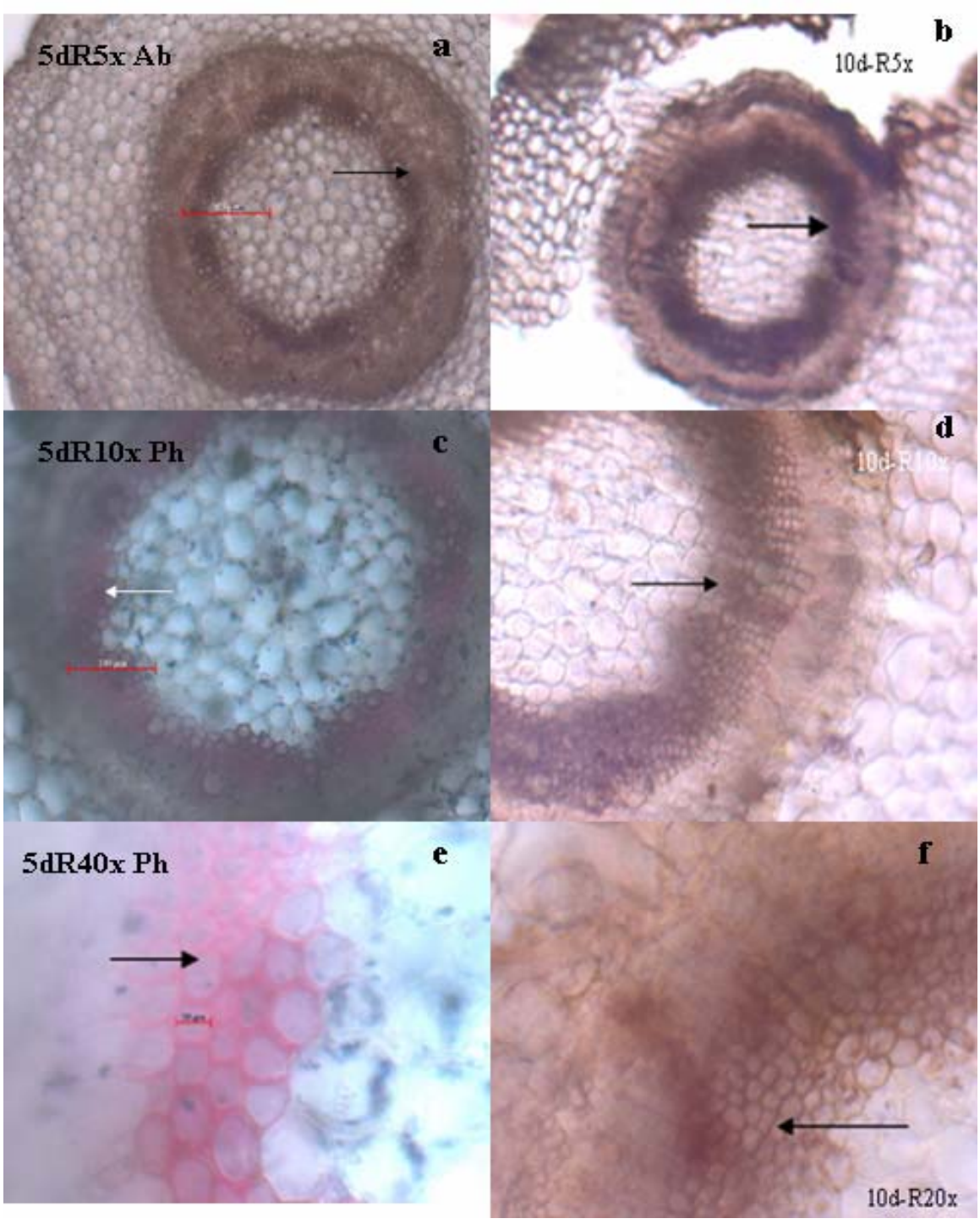


Figure 4B.1. Immunodetection of LlCAld5H protein levels and cell wall histochemical staining in the 5 days old *Leucaena* seedlings. Light micrographs of transverse sections of root showing LlCAld5H protein localization (a, b, d and f) and cell wall lignin staining with phloroglucinol (c and e).

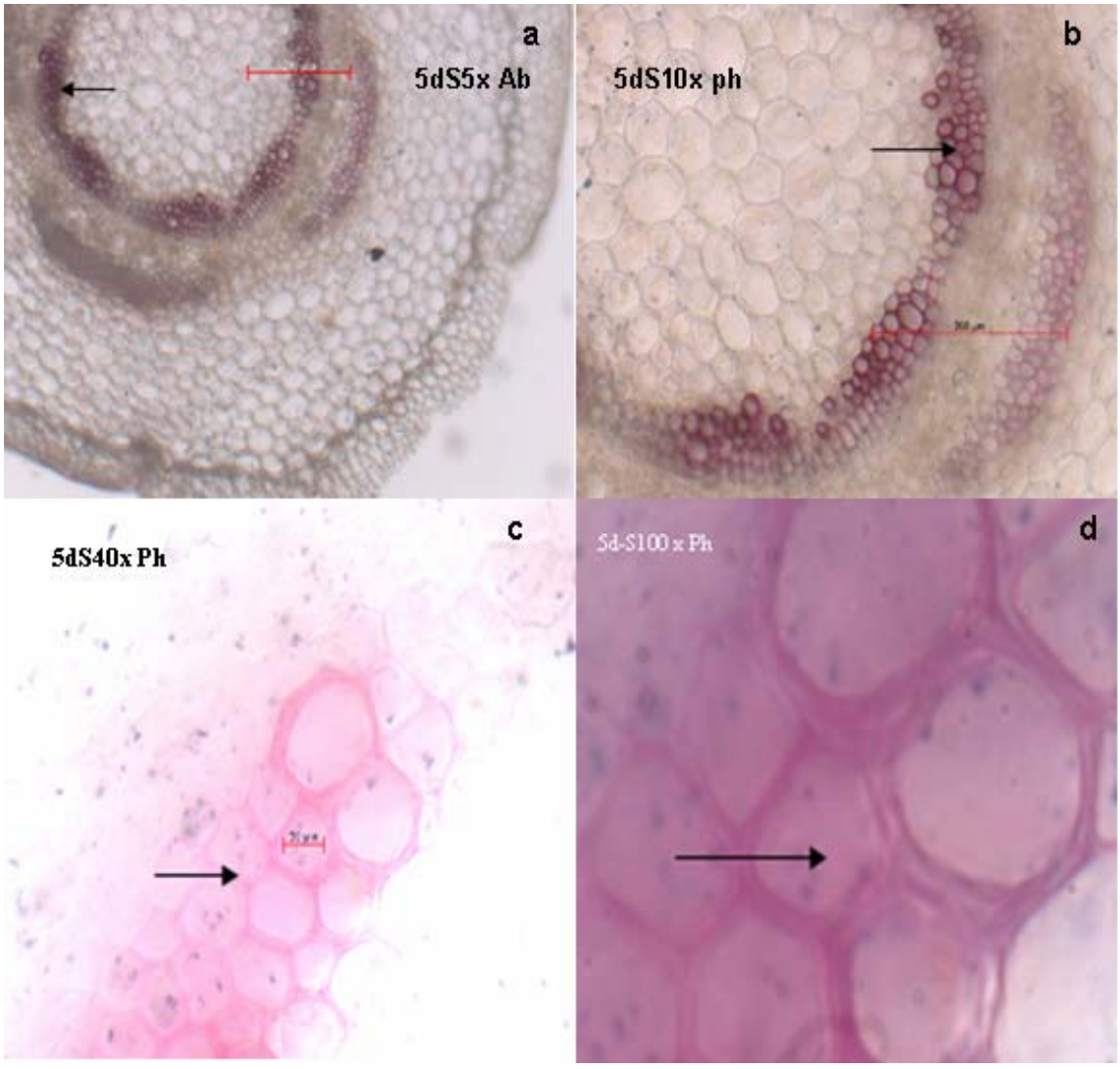


Figure 4B.2. Immunodetection of LlCAld5H protein levels and cell wall histochemical staining in the 5 days old *Leucaena* seedlings. Light micrographs of transverse sections of stem showing CAld5H protein localization (a) and cell wall lignin staining with phloroglucinol and snaps were captured at different magnifictions (b, c and d).

The abbreviation used in the pictures can be described as follows;

5dR5x Ab= Root sections from 5 day old plants stained with AntiCAld5H antibody and visualized at 5x of magnification.

5dR5x Ph= Root sections of 5 day old plants stained with phloroglucinol and visualized at 5x of magnification.

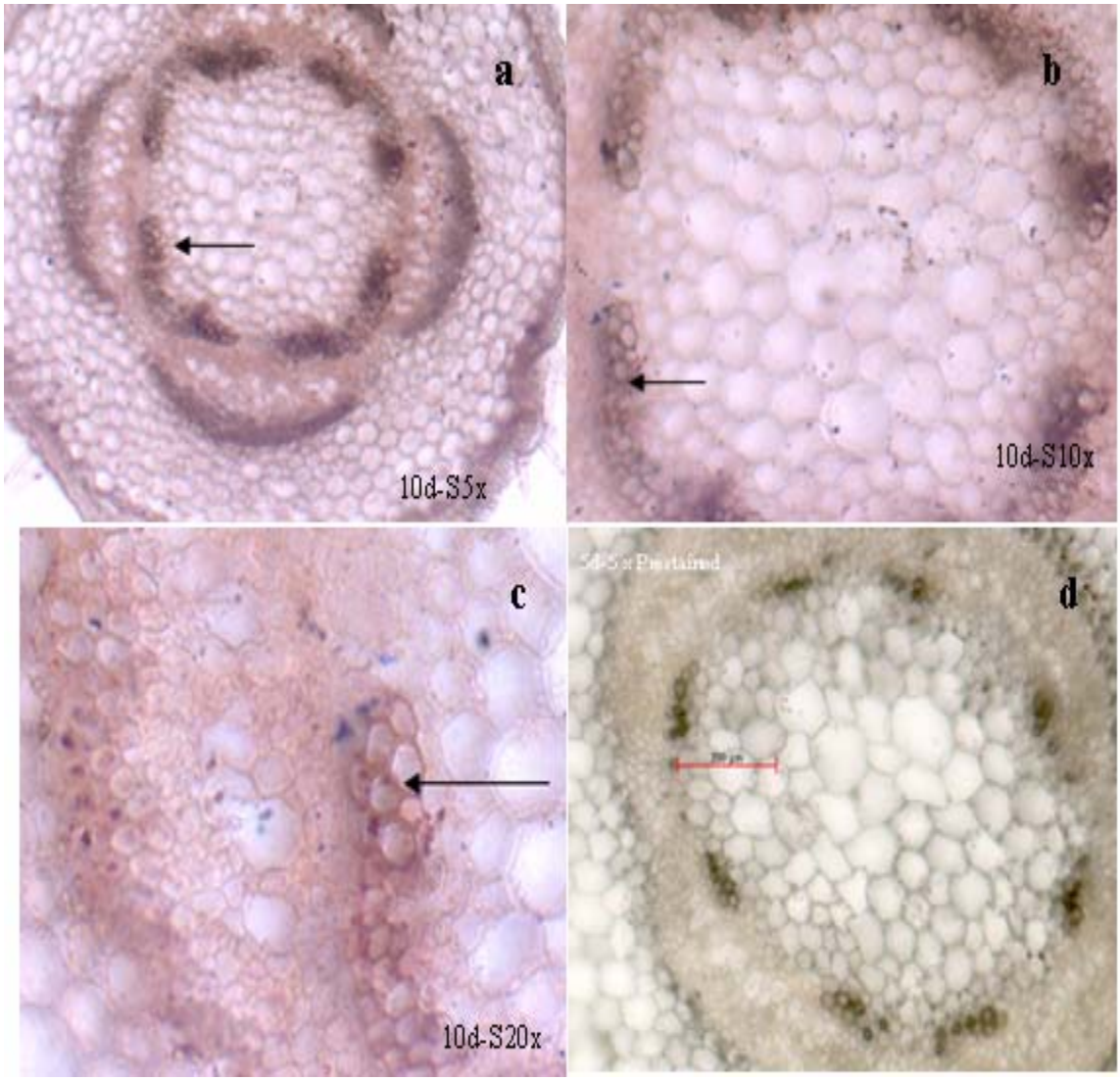


Figure 4B.3. Immunodetection of LlCAld5H protein levels in the 10 days old *Leucaena* seedlings. Light micrographs of transverse sections of stem showing LlCAld5H protein localization (a-c) at different magnifications. The pre-stained sections do not have any background color.

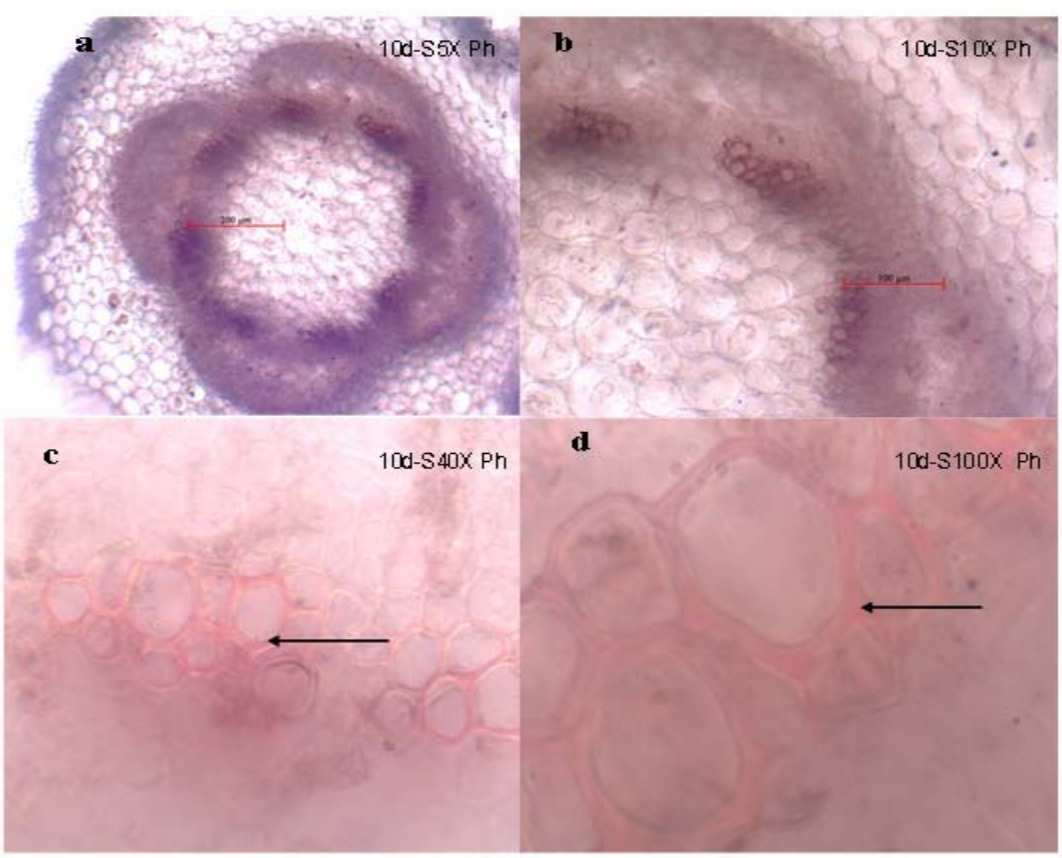


Figure 4B.3. Cell wall histochemical staining in the 10 days old *Leucaena* seedlings. Light micrographs of transverse sections of stem showing lignin localization in vascular tissue (a-d) at different magnifications. At higher magnification the apparent lignifications could be seen in the different layers of the cell wall i.e. S1, S2, and S3 layers (d). The pre-stained sections do not have any background color.

4B.3.5 Discussion

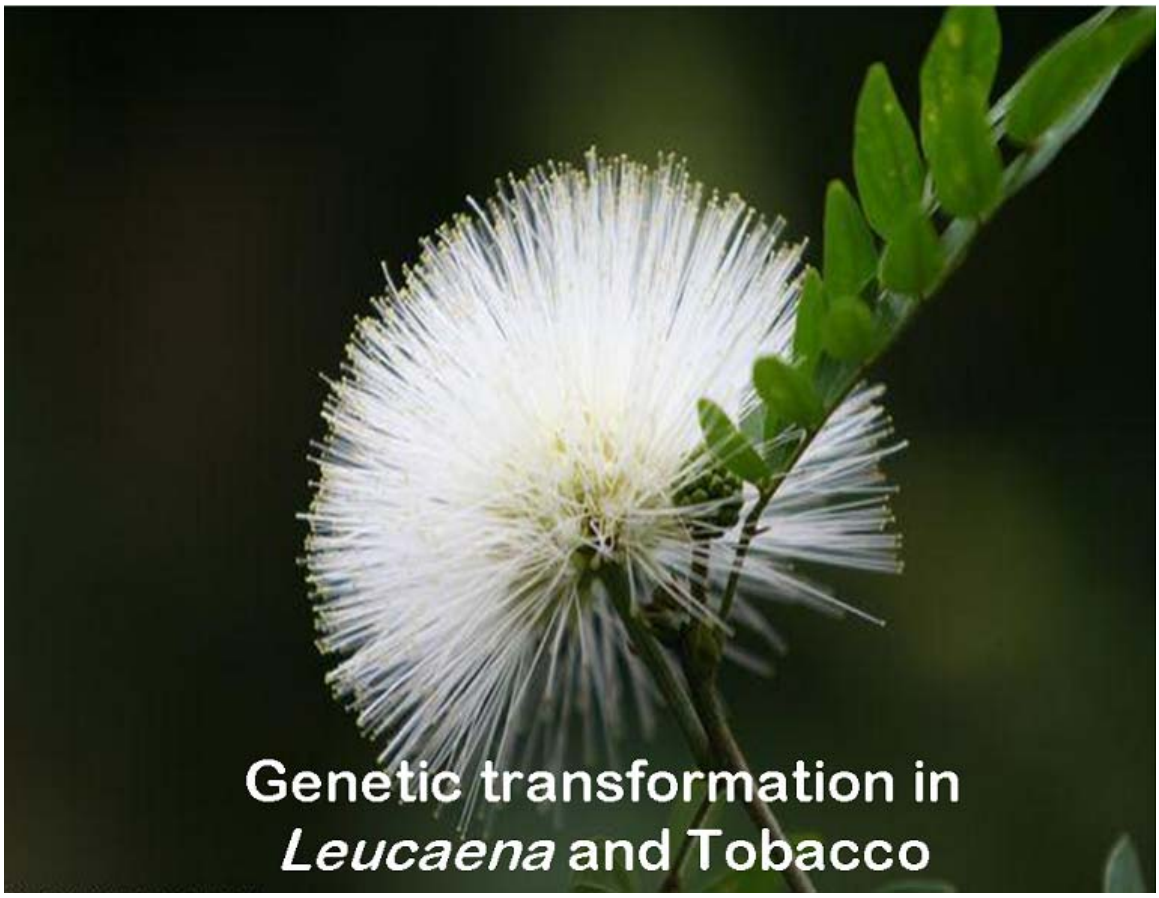
In an attempt to study the spatial and temporal expression of LlCAld5H gene the qRT-PCR was done. The pre-requisite for the RT-PCR is a good quality and accurately quantified RNA extracted from the tissue being used for this study. RNA extraction was done by TRIzol method with little deviations from standard protocol according to the manufacturer's recommendations were found to be necessary. A heating step was introduced which enhanced the net RNA yield, and it deemed necessary to precipitate the RNA with a high salt solution to discriminate against polysaccharides. The RNA was quantified by a very sensitive and robust 2100 Bioanalyzer system, just to make sure that

the equal quantity must be utilized in the subsequent steps i.e. cDNA synthesis. First strand cDNA was prepared by using random hexamers primers were used in the first strand cDNA synthesis as it produces the least bias cDNA. On the contrary, the use of oligo-dT primers can generate a biased RT product due to RNA secondary structure and length of polyA tail, and it preclude the use of 18S ribosomal gene as a reference (normaliser) gene with the relative quantification methods (David G. Ginzinger, 2002). In this chapter, a series of steps were carried out to verify the data with respect to their significance. First, each primer pair was tested for its optimal concentration and the molar ratio of individual primer pair was optimized to prevent primer-dimer build-up. The amplification efficiency of each primer was assumed to be optimum (between 75-100 %). The Ct values obtained from real-time PCR reflects the highest expression of the LlCAld5H gene in stem followed by radical and then in root. The expression of the gene in the young leaves has found to be less as compared to the stem and root. However, the increase in the expression level with age of the plant is obvious in root, but non homogenous increase has been observed in the stem and leaves as evident by the Ct values obtained. The expression of the LlCAld5H gene was increased in almost all the tissue types of 0-30 days of seedlings used in the present studies without any exception. Induction in LlCAld5H expression in response to salt stress (100 mM supplemented with ½-MS) was observed in the real-time qPCR studies done on *Leucaena.* The relative abundance in the transcript level was measured in different tissue types of different age. There are differences in the extent of increase in expression (fold expression) as compared to the untreated (normal) plant samples. This preliminary study tempted us to hypothesize the fact that, LlCAld5H expression is developmentally regulated in the different tissue types especially in the tissues undergoing lignification. Moreover, the expression level deviates towards positive side in response to the stress condition. The report of stress induced lignification is available (S, Koutaniemi *et al*., 2007), in extension to that we can speculate that there is also an increase in S lignin in response to salt stress. Finally, the LlCAld5H was immunolocalized in the transverse sections of root and stem tissue of *Leucaena leucocephala* seedlings of age groups that reflects the active lignification in vascular tissues.

4B.4. Conclusion

Spatio-temporal expression profile was investigated using real-time PCR analysis after optimizing a plethora of conditions such as primer concentration, RNA extraction, cDNA preparations and PCR cycling conditions. It was found that LlCAld5H gene transcript was expressed in radical, root, stem and to a slightly lower extent in leaves. However, a unique pattern of expression is exhibited by stem and leaves as in both the tissue the expression is increasing from 5-10 days but the 20 days old plants showed the decline in expression followed by further increase in expression of 30 days samples. LlCAld5H protein was immuno-cytolocalized in tissues undergoing lignification.

Chapter 5



5.1. Introduction

In angiosperm trees, lignin is polymerized from the guaiacyl and syringyl monolignols. Usually, the syringyl to guaiacyl ratio in angiosperm tree is 2-2.5 (Chang, H.M. and Sarkaren, 1973; Trotter, P.C., 1986). Lignin limits the degradability (reactivity) of the wood and thus is a potential barrier in pulpwood production. Paper and pulp industries across the world mostly depend on the lignin removal from wood through chemical reaction. Therefore, the current tree biotechnology emphasizes to develop genetically engineered trees, with altered lignin quality and quantity (reduced lignin quantity and increased S/G ratio). Wood pulp kinetics revealed that every unit increase in the lignin S/G ratio would approximately double the rate of lignin removal (Chang, H.M. and Sarkaren, 1973); supporting the idea that combination of S/G ratio augmentation and lignin reductions may offer far-reaching opportunities for maximizing wood-pulp production efficiency. However, in trees genetic reduction of lignin, which has been achieved through the suppression of the monolignol pathway genes encoding either 4-coumarate–CoA ligase (4CL) [Hu, W. J. *et al*., 1999] or caffeoyl CoA *O*-methyltransferase [Zhong, R. *et al*., 2000], had no significant effect on the S/G ratio. Attempts to modify the S/G ratio in trees could not succeed in lignin reduction (Van Doorssaelaere, 1996; Frank, R., 2000). These results argue that lignin quantity and the S/G ratio are regulated independently during lignin biosynthesis in tree.

5.1.1 Potential of CAld5H in transgenic plant with phenotype of altered lignin composition

In an extensive study based on CAld5H and COMT expression and kinetic studies from *Liquidambar styraciflua*, Osakabe *et al*., (1999) challenge the conventional concept of a ferulate 5- hydroxylation/methylation –regulated biosynthesis of syringyl lignin in angiosperms. Although a putative *F5H* cDNA was cloned from *Arabidopsis*, its biochemical function remains unknown. The over-expression of this cDNA in *Arabidopsis* mutant (fah1) lines deficient in syringyl lignin, but did not result in detectable F5H activity (Knut Meyer *et al*., 1998). In fact, 5-hydroxyferulate (the presumed product of F5H) as an intermediate for monolignol biosynthesis has not been reported in plants. Based on these facts and on the general lack of evidence that ferulate -

5-hydroxylation is involved in syringyl monolignol biosynthesis, Osakabe *et al.,* (1999) hypothesized that F5H may encode an enzyme that catalyzes 5-hydroxylation of guaiacyl intermediates other than ferulate to iniate the biosynthesis of the syringyl monolignol. In another study, Humphrey et al, (1999) has shown that coniferaldehyde but not the ferulate is the preferred substrate of the F5H enzyme as the *K*m of the enzyme for coniferaldehyde is 1.0 μM as compared to 1.0 mM for ferulic acid. Thus, F5H was renamed as CAld5H (Osakabe *et al*., 1998; Humphrey *et al*., 1999).

Considerable evidence is now available that shows that in angiosperm trees, the syringyl monolignol pathway (Figure 5.1) branches out from the guaiacyl pathway through coniferaldehyde and is regulated in sequence by three genes encoding hydroxylase (CAld5H), 5-hydroxyconiferaldehyde *O*-methyltransferase and sinapylaldehyde dehydrogenase. Enzyme kinetics further demonstrated that CAld5H has a 6-to 50-times-slower turnover rate than other two enzymes, thus CAld5H is a rate limiting enzyme and consequently regulate the S/G ratio.

Several studies have been done on genes of lignin biosynthesis like *CCR* (Piquemal *et al*., 1998), *COMT* (Pincon *et al*., 2001), *CAD* (Halpin *et al*., 1994; Chabannes *et al*., 2001) to produce transgenic plants using antisense and siRNA technology with low lignin content. Down-regulation of various key genes have been achieved to obtain transgenic plants with low or modified lignin content. The simultaneously up-regulating *CAld5H* and down- regulating *4CL* gene expression may lead to a concurrent lignin-reactivity augmentation and quantity reduction in plants (Li *et al.,* 2003).

Hence, work was initiated with the objective of cloning and characterizing *CAld5H* cDNA of *Leucaena* and also for its genetic transformation and consequent effect on lignin quality in transgenic plants. Significant amount of study has been done on various lignin biosynthesis genes in last ten years around the world. Prime objectives were to increase forage value, proper utilization of biomass and to increase resistance against pathogens. To meet the increasing demand of high quality wood for paper industry, it is essential to provide transgenic plant species.

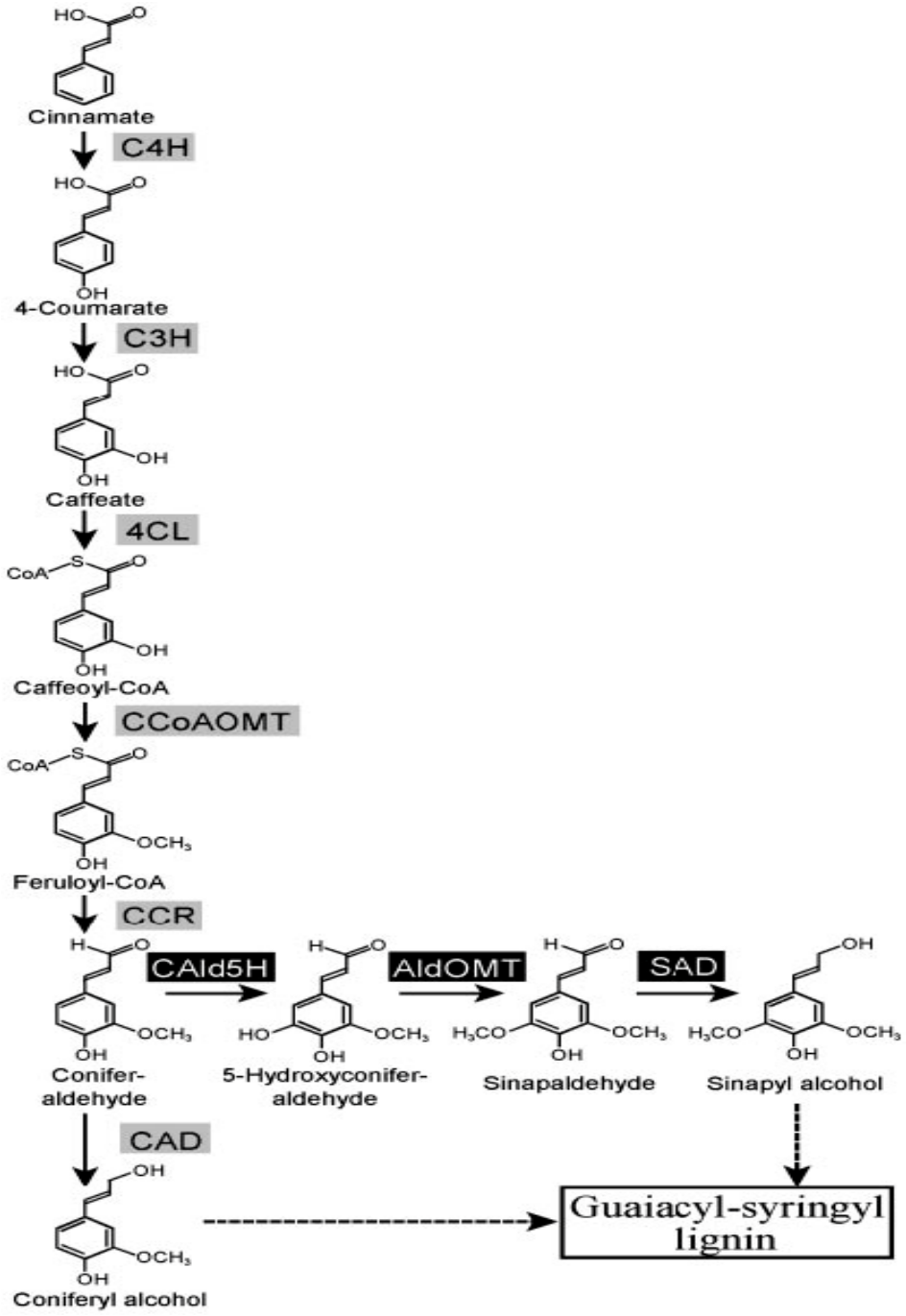


Figure 5.1. Biosynthetic pathway for the formation of guaiacyl (coniferyl alcohol) and syringyl (sinapyl alcohol) monolignols. C4H, cinnamate 4-hydroxylase; C3H, 4-coumarate 3-hydroxylase; 4CL, 4-coumarate:CoA ligase; CCoAOMT, caffeoyl-CoA O-methyltransferase; CCR, cinnamoyl-CoA reductase; CAD, cinnamyl alcohol dehydrogenase; CAld5H, coniferaldehyde 5-hydroxylase; AldOMT, 5-hydroxyconiferaldehyde O-methyltransferase; SAD, sinapyl alcohol dehydrogenase.

To meet the proposed objectives, isolation, cloning, and characterization of *CAld5H* cDNA was done. The heterologous expression of *CAld5H* gene in *E. coli* and the preliminary studies on the enzyme catalysis where it utilizes the coniferaldehyde as the substrates reveals its role in lignification. The involvement of the candidate gene in the monolignol biosynthesis has also been confirmed by the spatio- temporal expression studies and by immunocytolocalization.

To study the effect of *CAld5H* gene on lignification, sense *CAld5H* gene construct was introduced in *Leucaena* using particle bombardment followed by *Agrobacterium* co-cultivation methods, whereas the tobacco were transformed with the antisense construct of the partial length LlCAld5H gene by *Agro-infusion* method. Dual purpose will be served with this study in *L. leucocephala*: a transgenic could be obtained with altered lignin content, which can be utilized in paper and pulp industry and secondly, it will be proved that cDNA isolated from *Leucaena* encodes *CAld5H* gene, which is actively involved in constitutive lignification.

Different strategies used in genetic transformation of *Leucaena* and *tobacco* plants and the screening and analysis of transformants are the major contents of this chapter. The plant transformation vector pCAMBIA1301, harboring the *CAld5H* gene in sense and antisense orientation and genes for GUS (reporter), was used. Three different strategies i.e. *Agrobacterium* mediated, particle bombardment and particle bombardment followed by co-cultivation are described in details in this chapter. The chapter also includes the evaluation and analysis of putative transformants and confirmation of integration of the reporter and selectable marker gene in *L. leucocephala* genome by GUS assay and by molecular techniques like PCR.

5.2 An introduction of genetic transformation of grasses and trees

Agrobacterium-mediated and particle bombardment method of transformation are the two most widely used methods for plant genetic modification. The process of particle bombardment involves direct mechanical integration of the DNA into the host genome using metal beads as carriers. Particle bombardment has been used for many years in many graminaceous species such as perennial ryegrass (*Lolium perenne* L.)

(Spangenberg *et al*., 1995a; Altpeter *et al*., 2000; Cho *et al.,* 2000; Jensen *et al.,* 2004), Kentucky bluegrass (*Poa pratensis* L.) (Gao *et al.,* 2006), tall fescue (*Festuca arundinacea* Schreb.) (Wang *et al.,* 1992; Spangenberg et al. 1995b; Cho *et al.,* 2000), red fescue (*Festuca rubra* L.) (Spangenberg *et al*., 1994, 1995b; Altpeter and Xu. 2000), creeping bentgrass (*Agrostis palustris* Huds.) (Hartman *et al*., 1994), zoysiagrass (*Zoysia japonica* Steud.) (Inokuma *et al*., 1997) and triploid bermudagrass (*Cynodon dactylon* × *C. transvaalensis*) (Zhang *et al*. 2003). The *Agrobacterium* method has been successfully practiced in dicots for many years, but only recently have effective protocols been developed for grasses (Yu *et al.,* 2000; Somleva *et al*., 2002; Luo *et al*., 2004; Altpeter *et al*., 2004; Li *et al*., 2005; Wu *et al*., 2005; Ge *et al.,* 2006; Bajaj *et al*., 2006). In general, the *Agrobacterium* method is considered superior to particle bombardment because of a greater frequency of low copy transgene insertions and higher and more stable transgene expression over generations (Dai *et al*., 2001; Shou *et al*., 2004; Travella *et al*., 2005).

In the case of woody tree the genetic transformation protocols based on *Agrobacterium*-mediated and/or direct gene transfers by biolistic bombardment have been successfully applied for numerous woody angiosperm species (Merkle & Nairn, 2005), including *Populus* and *Betula.* Plants are genetically engineered by introducing gene(s) into plant cells that are growing *in vitro* or *ex vitro*. The development of transgenic plants is based on the stable insertion of foreign DNA into the plant genome, regeneration of these transformants to produce the whole plant and expression of the introduced gene(s). *Agrobacterium*-mediated transformation has provided a reliable means of producing transgenics in a wide variety of plant species that can be cultured and regenerated *in vitro*. Recently, some plants such as *Arabidopsis thaliana* have also been transformed by *Agrobacterium*-mediated transformation by dipping the young buds of flowers of *ex vitro* grown plants (Rakoczy-Trojanowska, 2002). This method is known as infiltration in plant transformation. Other methods of gene-transfer systems include particle bombardment, electroporation and membrane permeabilization using chemicals. Of these, particle bombardment has proved to be successful with plants that are less sensitive to *Agrobacterium* infection, such as cereals and legumes (Walden & Wingender, 1995). However, recently, *Agrobacterium*-mediated transformation has

become the method of choice for these plants (Nadolska-Orcyzyk, Orczyk & Przetakiewicz, 2000). The development and optimization of several regeneration protocols, efficient vector constructs and availability of defined selectable marker genes and different methods of transformation have resulted in the production of transgenic plants in more than 100 plant species (Babu *et al.*, 2003; Wimmer, 2003). These transgenic plants include many important crops, fruits and forest plants. The plant transformation technology is not only used to improve plants but also a versatile platform for studying gene function in plants. Plant genetic transformation technology has a great potential in increasing productivity through enhancing resistance to diseases, pests and environmental stresses and by qualitative changes such as chemical composition of the plants. Plants can also be used for high volume production of pharmaceuticals, nutraceuticals and other beneficial chemicals. Transgenic plants might be used as drug delivery devices, with vaccines being synthesized in plants (Hansen & Wright, 1999). Many plant species previously considered to be recalcitrant to transformation, with advances in tissue culture combined with improvements in transformation technology, have now been transformed.

5. 2.1 *Agrobacterium* mediated plant transformation

The natural ability of the soil microorganism *Agrobacterium* to transform plants is exploited in the *Agrobacterium*-mediated transformation method. During infection process, a specific segment of the plasmid vector, T-DNA, is transferred from the bacterium to the host plant cells and integrates into the nuclear genome.

5.2.2 Biology and life cycle of *Agrobacterium tumefaciens*

Agrobacterium tumefaciens is a gram negative soil inhabiting bacteria that causes crown gall disease in a wide range of dicotyledonous plants, especially in members of the rose family such as apple, pear, peach, cherry, almond, raspberry and roses. The strain, biovar 3, causes crown gall of grapevine. Although this disease reduces the marketability of nursery stock, it usually does not cause serious damage to older plants. Smith and Townsend first described *Agrobacterium infection* in 1907. The bacterium transfers part of its DNA to the plant, and this DNA integrates into the plant's genome, causing the

production of tumors and associated changes in plant metabolism. The unique mode of action of *A. tumefaciens* has enabled this bacterium to be used as a tool in plant transformation. Desired genes, such as insecticidal or fungicidal toxin genes or herbicide-resistance genes, can be engineered into the bacterial T-DNA and thereby inserted into a plant. The use of *Agrobacterium* allows entirely new genes to be engineered into crop plants. *Agrobacterium*-mediated gene transfer is known to be a method of choice for the production of transgenic plants with a low copy number of introduced genes (Hiei *et al.,* 1997).

5.2.3 Infection process

Agrobacterium tumefaciens infects the plants through wounds, either naturally occurring or caused by transplanting of seedlings and nursery stock. In natural conditions, the motile cells of *A. tumefaciens* are attracted to wound sites by chemotaxis. This is partly a response to the release of sugars and other common root components. Strains that contain the Ti plasmid respond more strongly, because they recognize wound phenolic compounds like acetosyringone even at very low concentrations (10^{-7} M). Acetosyringone plays a further role in the infection process by activating the virulence genes (*Vir* genes) on the Ti plasmid at higher concentrations (10^{-5} to 10^{-4} M). These genes coordinate the infection process. It is important to note that only a small part of the plasmid (T-DNA) enters the plant and the rest of the plasmid remains in the bacterium to serve further roles. When integrated into the plant genome, the genes on the T-DNA code for auxins, cytokinins and other novel plant metabolites (opines and agrocinopines). These plant hormones upset the normal balance of cell division leading to the production of galls. Opines are unique amino-acid derivatives and the agrocinopines are unique phosphorylated sugar derivatives. All these compounds can be used by the bacterium as the sole carbon and energy source.

5.2.4 Markers for Plant Transformation

5.2.4.1 Selectable markers

Genes conferring resistance to antibiotics like *neomycin phosphotransferase* II (nptII) (Baribault *et al.,* 1989), *hygromycin phosphotransferase (hpt)* (Le Gall *et al.,* 1994),

phosphinothricin acetyl transferase / bialaphos resistance (pat/bar) (Perl *et al.,* 1996) are being used to select transgenic cells. Another selectable marker gene, phosphomano-isomerase (pmi), which catalyzes conversion of mannose-6-phosphate to fructose-6-phosphate, an intermediate of glycolysis that positively supports growth of transformed cells, is also recently being used. Mannose absorbed by the plant cells gets converted into mannose-6- phosphate, an inhibitor of glycolysis, inhibits growth and development of nontransformed cells. Transformed cells having *PMI* gene can utilize mannose as a carbon source.

5.2.4.2 Screenable markers

The oncogenes of *Agrobacterium* are replaced by reporter / screenable marker genes like *β-glucuronidase* gene (*gus*) (Baribault *et al.,* 1990), *luciferase (luc)* gene for analyzing gene expression. Since the first demonstration of the *green fluorescent protein (gfp)* gene from jellyfish *Aequorea victiria* as a marker gene (Chalfie *et al.,* 1994), it has attracted increasing interest and is considered advantageous over other visual marker genes. Unlike other reporter proteins, GFP expression can be monitored in living cells and tissues in a non-destructive manner. This gene has been used as a visible reporter gene in genetic transformation of both monocots and dicots (Haseloff *et al*., 1997; Reichel *et al*., 1996; Kaeppler and Carlson, 2000). The fluorescence emission of GFP only requires the excitation of living cells by UV or blue light (390 nm strong absorption and 470 nm week absorption), which results from an internal 4-(*p*-hydroxybenzylidene)-5-imidazolinone chromophore generated by an autocatalytic cyclization and oxidation of a ser65-gly67 residues of the proein. The other advantage of *gfp* as a reporter gene is that no exogenously supplied substrate/ cofactors are needed for its fluorescence emission at 508 nm. Red Fluorescent Protein marker (DsRed2, a mutant form of DsRed from *Discosoma* sp.) was first used as a visual reporter gene for transient expression and stable transformation of soybean (Nishizawa *et al*., 2006). DsRed2 fluorescence can be monitored with any fluorescence stereomicroscope equipped with a filter set for excitation at 530–560 nm and emission at 590–650 nm.

5.3 Materials and methods

Leaf discs of *in-vitro* grown tobacco (*N. tabaccum* var. Anand 119) were used as the explants for the *Agrobacterium*-mediated transformation. Seeds of *Leucaena,* imbibed in distilled water after the treatment with conc. sulphuric acid (7-10 min) and mercuric chloride (0.1 % for 10 min) were used as source of embryo axes. Embryo axes excised from the seeds and inoculated on regeneration medium ([1/2 MS + TDZ (0.5 mg/L)) were used as the target material for the *Agrobacterium*-mediated and particle bombardment transformation

5.3.1 Bacterial strain and plasmid

Agrobacterium tumefaciens GV2260, *E. coli* (XL-1 blue). Plasmid vectors like pCAMBIA 1300 and pCAMBIA1301 binary plant transformation vector (CAMBIA, USA) and pGEMT-Easy (Promega USA), were used.

5.3.2 Plant Material

Tobacco explants: Fresh leaves were taken from axenic cultures of *Nicotiana tabacum* var. Anand 119 and ~ 6 mm diameter disc were punched out.

Leucaena leucocephala explants: Axenic cultured embryo axes of *L. leucocephala* were taken as explants for transformation experiments.

5.3.3 A. *tumefaciens* culture conditions, transformation and selection

The bacterial culture condition, transformation and selection were done as described in chapter 2: Section 2.8.2.3 and *A. tumefaciens* mediated *tobacco* transformation was discussed in chapter 2; section 2.11.

5.3.4 Transformation of *Leucaena leucocephala*

One day old embryo axes without cotyledons were used as explants for transformation. Seeds of *Leucaena* imbibed in distilled water after the treatment with concentrated sulphuric acid (7-10 min) and mercuric chloride (0.1 % for 10 min), were used as source of embryo axes. Embryo axes were excised from the seeds and inoculated on regeneration medium [1/2 MS + TDZ (0.5 mg/ L)]. The embryos were then used for transformation. LD_{50} for hygromycin in *L. leucocephala* was estimated as 15 mg/L in the previous experiment. The transformation was carried out by three methods namely, 1) Particle bombardment, 2) Particle bombardment followed by co-cultivation and 3) Agro-infusion method.

5.3.4.1 Particle bombardment

Calculated amount (approximately 500 µg) of gold particles of 1.0 µm size was used for bombardment. The gold particles taken in microcentrifuge tubes were washed with sterile water and supernatant was decanted after centrifuging at 13,000 g for 10 min. The washing of particles was repeated three times with 70% (freshly prepared) ethanol and supernatant was decanted after centrifuging at 13,000 g for 10 min. Again particles were suspended in 500 µL sterile water and were pelleted down after centrifuging at 13,000 g for 10 min. Plasmid DNA of 1.0 µg (plasmid/cassette) was dissolved in 100 µL of TE buffer, added to the prepared gold particles and mixed by vortexing. Spermidine (100 µL of 0.1 M) was added to it and mixed by vortexing. PEG 3000 (100 µg) was added and mixed by vortexing followed by addition of 100 µl of 2.5 M $CaCl_2$ and incubated in ice bath for 10 min. It was then given a brief spin to settle down the sample at the bottom. The sample was resuspended in absolute ethanol and sonicated briefly to resuspend the particle and was used for bombardment. Particle bombardment was done using BioRad PDS 1000/ He system. Helium gas pressure (1100 psi) was used to bombard the particle. Micro-carrier travel distance was adjusted to 6 cm. The embryo axes were arranged in the centre of the petridishes and bombarded with gold particles coated with DNA (plasmid/cassette). The explants were incubated in dark for 72 h after the bombardment.

5.3.4.2 Particle bombardment followed by co-cultivation

The embryo axes were first bombarded with gold particles coated with DNA (plasmid/cassette) as mentioned above followed by co-cultivation with *Agrobacterium* (GV2260) harboring pCAMBIA1301 containing LlCAld5H gene in sense orientation. Cultured *Agrobacterium* cells were pelleted by centrifugation at 5,000 g for 5 min and resuspended in the YEM medium at a density of $3\text{-}5\times10^9$ cells/mL, considering the optical density of bacterial culture at 600 nm. The selected embryo axes after bombardment were co-cultured with *Agrobacterium* for one hour by immerging them into the bacterial suspension. The agro infected embryo axes were then transferred onto the regeneration medium [1/2 MS + TDZ (0.5 mg/ L)] with or without 0.1 mM acetosyringone (3', 5'-dimethoxy-4'-hydroxy-acetophenone; Sigma-Aldrich Chemical Co.) and co-cultivated in the dark at 28 ± 2 °C for 3 days. After co-cultivation, the embryo axes were washed thoroughly with cefotaxime 250 mg/ L in sterile distilled water and transferred onto the regeneration medium [1/2 MS + TDZ (0.5 mg/ L)].

5.3.4.3 Agro-infusion method

In this method, the *Agrobacterium* was introduced forcibly with a sterile needle to the meristematic region of the explants. After transformation, embryo axes were cultured on regeneration medium [1/2 - MS + TDZ (0.5 mg/l)] without selection for one week. Then the axes were shifted to selection medium containing hygromycin (10 mg/ L) for 3 weeks followed by selection on hygromycin 15 mg/ L for another 3 weeks. The survived explants on hygromycin (15 mg/ L) were shifted to 1/2 MS without hygromycin selection. Cytokinin, 2ip (2-isopentenyl adenine; 0.5 mg/ L) was used in the medium to have better elongation of transformed shoots.

5.3.5 DNA extraction from transgenic lines

DNA was isolated using a modified protocol of Lodhi, *et al*., 1994. In short, one or two leaflets (*Leucaena*) and a piece of leaf (tobacco) were ground in 20 – 30 µL CTAB

extraction buffer (Chapter 2, section 2.8.4.2) using glass pestle in a microfuge tube. Extraction buffer (500 µL) and PVPP (polyvinyl polypyrolidone) to a final concentration of 0.2% were used and mixed gently. The tube was kept at 65 □C for 20 min with intermittent shaking. RNAse was added and kept for 10 more min at 65 □C. Tube was cooled and 400 µL of Chloroform: Isoamyl- alcohol (24:1) was added and mixed gently. Tube was centrifuged at 6000 g for 10 min at 4 □C, and supernatant was transferred to a fresh microfuge tube. Isopropanol (0.6 volume) or pre-chilled absolute ethanol (2 volumes) was added and kept at room temperature for precipitation. Tube was again centrifuged at 6000 g for 10 min at 4 □C to pellet DNA. DNA was washed with 70% cold ethanol, air dried and dissolved in 40 µL SMQ water.

5.3.6 ELISA (Enzyme-Linked Immunosorbent Assay)

Fresh tissues were collected, frozen in liquid nitrogen and crushed to a fine powder. Crude protein was extracted with 2 mL of protein extraction buffer (100 mM Tris HCl pH.7.5, 2% PVPP, 2% PEG 4000, DTT 5 mM and PMSF 1 mM). Total Protein was quantified using Bradford reagent. Equal amount of protein was coated on 96 well micro titre plates. ELISA protocol was performed as discussed in section 2.13.

5.3.7 PCR amplification, cloning and bacterial transformation

As discussed in chapter 2; section 2.8.4.9.

5.4 Results and Discussion

5.4.1 Cloning and construction of antisense vector

To introduce the partial-length antisense LlCAld5H in tobacco, a plasmid vector pCAMBIA1300-LlCAld5H500 was constructed as follows. DNA fragment of 540 bp, corresponding to the coding sequences of LlCAld5H was PCR amplified from plasmid containing cDNA of the *Leucaena leucocephala* coniferaldehyde-5-hydroxylase gene (GenBank accession no. EU041752). Standard cycling condition was followed in the PCR and amplification was obtained with a forward primer (5'-GAGCTCGGTCCAACGGGAACTGGCC-3') containing *Sac*I restriction site and reverse primer (5'- GGTACCAAAGAGAGGGCAGACCAAGC -3') containing *Kpn*I site. PCR-amplified fragment was cloned in pGEMT-Easy (Promega, Madison, WI, USA) and sequenced. The recombinant plasmid (referred as pGEMT-LC500) was digested with *Sac*I–*Kpn*I restriction endonuclease that has released a ~500 bp fragment (Figure 5.2).

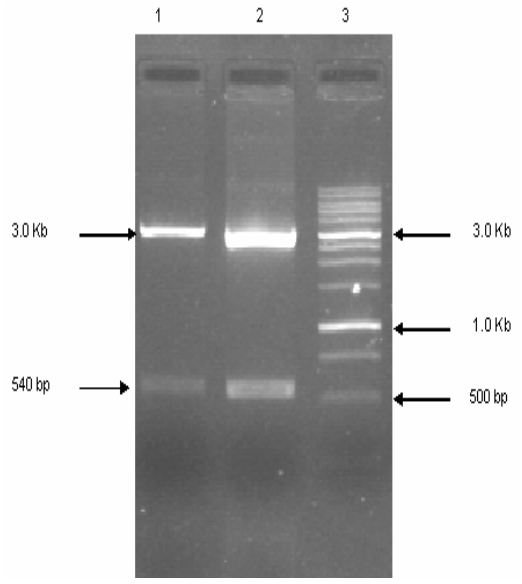


Figure 5.2. 1.5% agarose gel showing 540 bp partial lengths LlCAld5H gene as a digested fragment released from the linearized vector (lanes 1 and 2). The medium range DNA size marker has resolved in lane 3.

Prior to this, pCAMBIA 1301 was modified in our laboratory by cloning a cassette of CaMV 35S promoter and NOS terminator within *Eco*RI and *Hin*dIII site of pCAMBIA 1301. This cassette also contained a *Kpn*I (near CAMV 35S), *Bam*HI and a *Sac*I (near NOS terminator) restriction sites. The pCAMBIA 1300 binary vector was digested with *Kpn*I and *Sac*I restriction enzymes. To prevent the recircularization of the vector (which we had encountered many times during the antisense cloning), the digested vector was treated with calf intestinal alkaline phosphatase (0.5 unit/µg of plasmid DNA) with NEB buffer 3. The linearized pCAMBIA1300 vector was extracted with Phenol:chloroform (24:1) and used for ligation. *Sac*I–*Kpn*I digested pGEMT-L500 fragment was ligated into corresponding site of the linearized vector pCAMBIA1300. The ligation product was transformed into *E.coli* (XL -1 blue) cell line. The recombinant plasmid was confirmed by *Sac*I- *Kpn*I digestions which yielded a digestion fragment of 540 bp as depicted in Figure.5.3 and also sequenced. The recombinant clone (referred as pCAMBIA1300-*AsCAld5H*), was later mobilized into *A. tumefaciens* host cell.

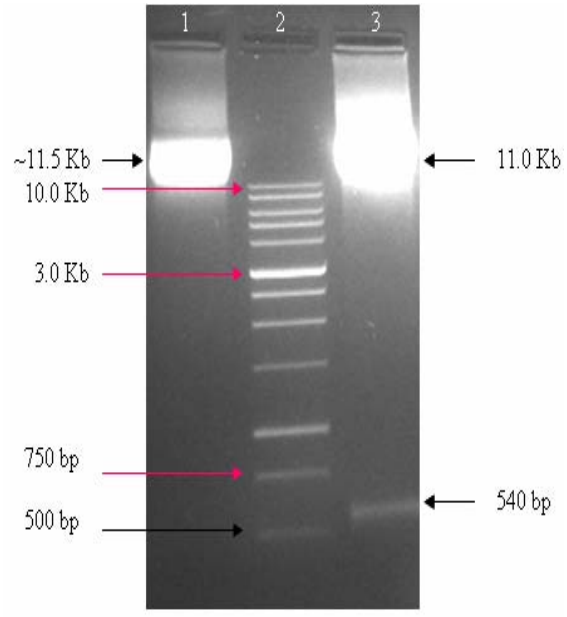


Figure 5.3. Partial length gene fragment released out as digestion product resolved on 1% agarose gel (lane 3). Undigested vector (lane 1), DNA marker (lane 2).

The right and left hand T-border of pCAMBIA1300 vector harbors the hygromycin resistant gene and multiple cloning sites (Figure 5.4). This vector does not have any reporter gene thus the transformants using this vector can not be analyzed by reporter gene assay.

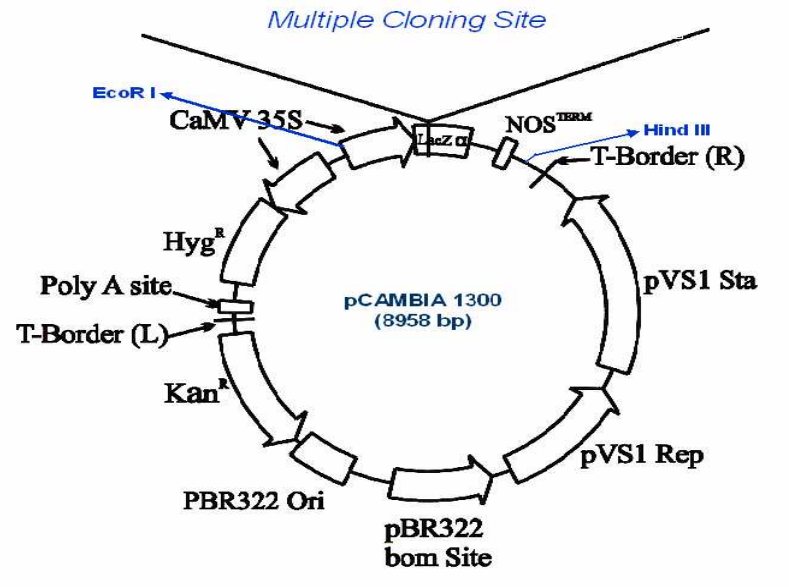


Figure 5.4. Map of pCAMBIA1300 shuttle vector.

Moreover, selection of the transformants cannot be done at the beginning of the transformation and failure of the experiment would be noticed at later stage. To avoid these short comings, pCAMBIA1301 vector was used for transformation. The right and left hand T-border of pCAMBIA1301 vector harbors the hygromycin resistance gene (selectable marker) for plant selection, multiple cloning sites and GUS (reporter gene) with axon and introns but the MCS of pCAMBIA1301 is devoid of any promoter and terminator to drive the gene expression and termination successfully. Presence of *Eco*RI and *Hin*dIII restriction sites at either side of multiple cloning sites of pCAMBIA1301 has enlightened the path of successful antisense cloning as the entire antisense construct including promoter and terminator sequences could be taken out as *Eco*RI – *Hin*dIII fragment from the pCAMBIA1300-*AsCAld5H* and cloned to the new vector. The shortcomings of both the vectors could be complemented with this cloning strategy.

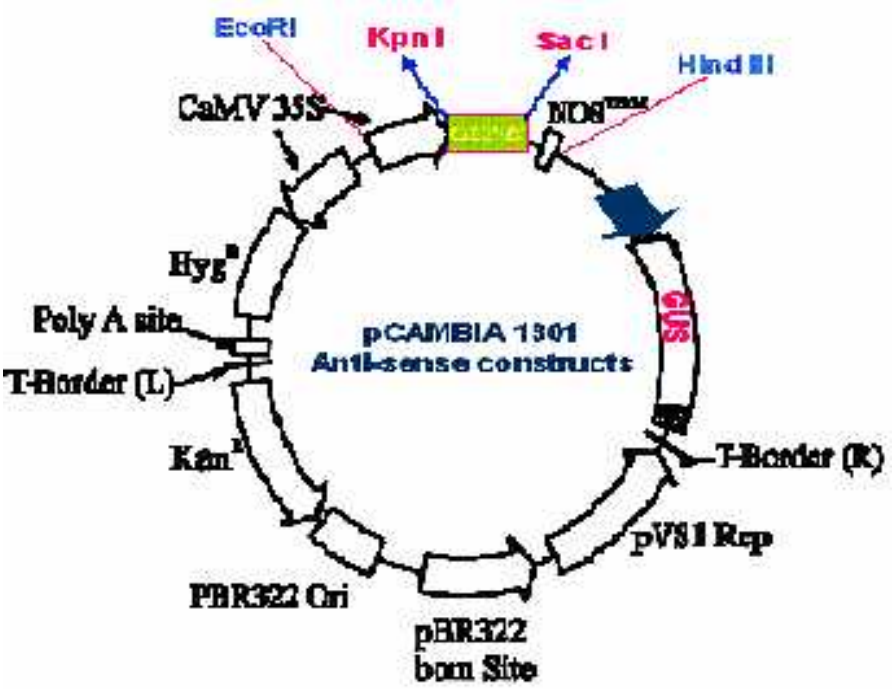


Figure 5.4. Map of pCAMBIA1301 shuttle vector depicting multiple cloning sites, the GUS gene, left and right boarder of the T-DNA gene and plant and bacterial selectable marker genes.

The antisense construct that was finally used in plant transformation can be represented as in Figure 5.5. The region between the *Eco*RI and *Hin*dIII of antisense construct pCAMBIA1301-*AsCAld5H* shown in figure5.5, originally belongs to pCAMBIA1300-*AsCAld5H* clone which contain the 35S promoter, the partial length gene cloned within *Sac*I and *Kpn*I sites and the NOS terminator. The recombinant plasmid was transformed into *Agrobacterium tumefaciens* (GV2260 strain). This construct was introduced into tobacco by *Agrobacterium* mediated transformation method.

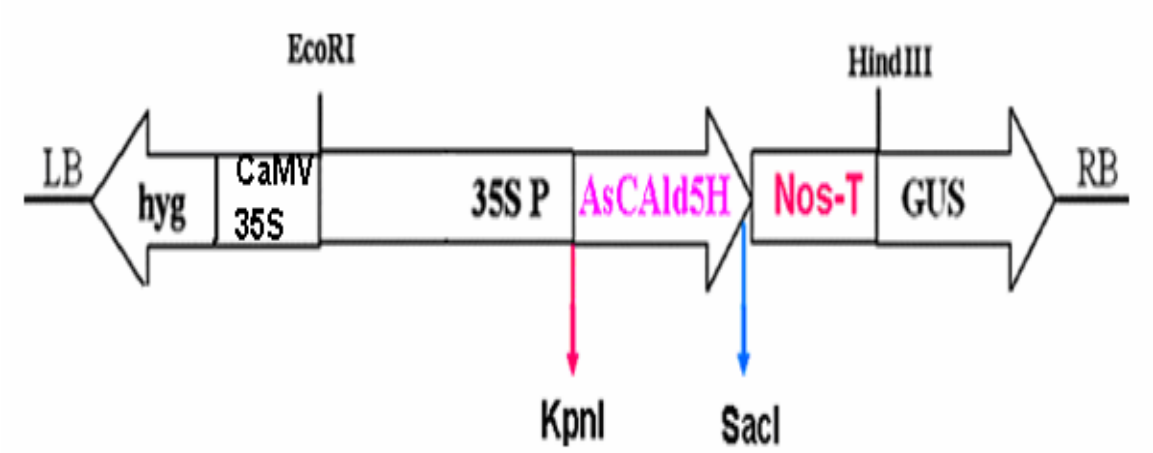


Figure 5.5. Diagrammatic representation of the antisense constructs pCAMBIA1301-*AsCAld5H*. Region between *Eco*RI and *Hin*dIII has taken out from pCAMBIA1300-*AsCAld5H* construct.

5.4.2 Tobacco transformation

Tobacco (*N. tabaccum* var. Anand 119) leaf discs were transformed separately with *A. tumefaciens* cultures harboring the antisense construct shown in figure 5.5. Shoots induction started after 2 weeks under selection pressure (hygromycin 3 mg/L) from the cut surface of the leaf disc (Figure 5.6 a). Proliferation of induced shoot started and noticed after 4 weeks (Figure. 5.6 b, c and d). The regenerants were allowed to grow for 12 weeks and then shifted to root induction medium. Roots were initiated within 2 weeks of shifting (Figure 5.6 e).

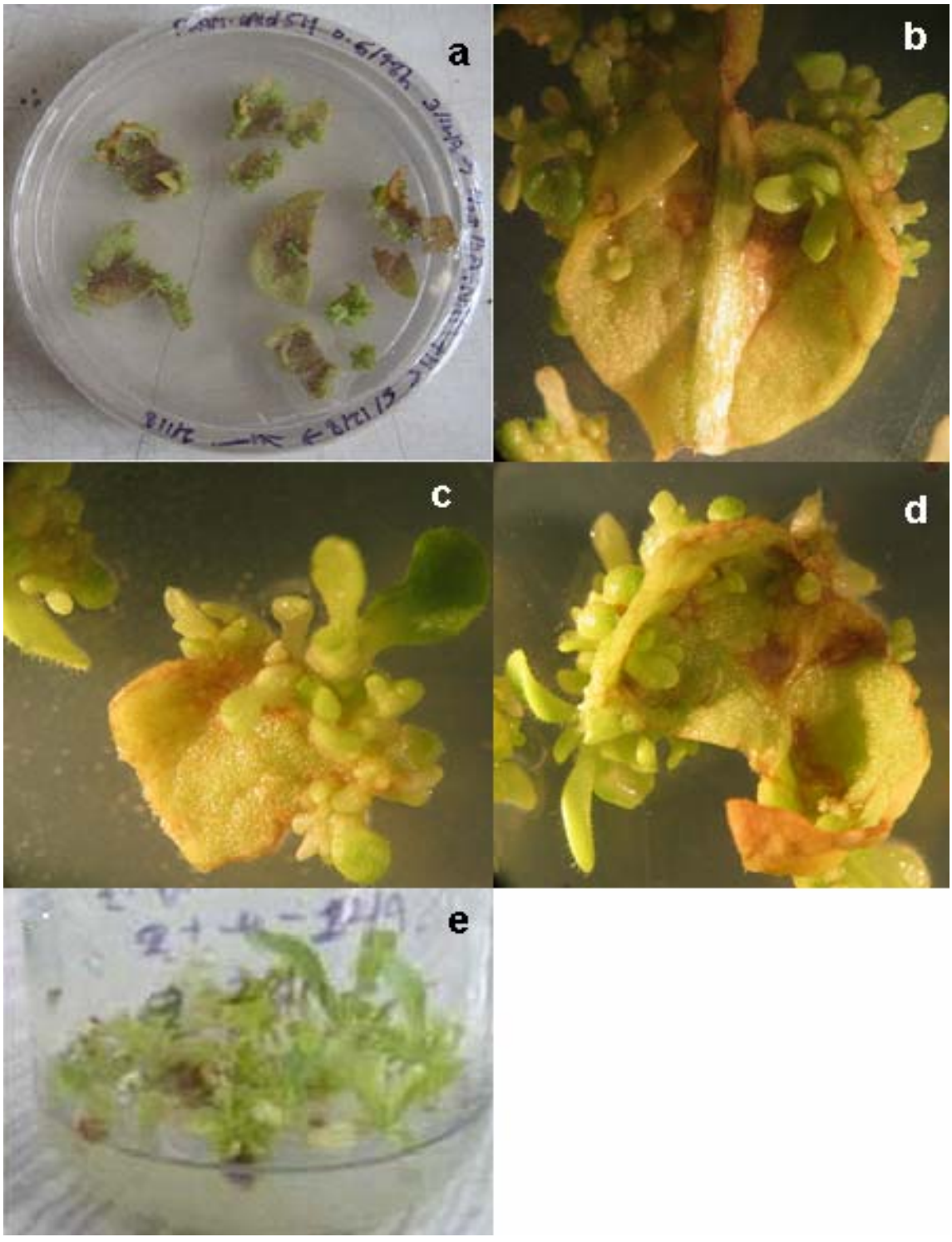


Figure 5.6. Putative transformed tobacco shoots regenerated *in vitro* on selection medium. Shoot bud induction after 2 weeks (a), proliferation of shoot bud after 4, 8 and 12 weeks (b, c and d) and shoots transferred to root inducing media (e).

5.4.3 Analysis of tobacco transformants

Tobacco was used as a standard transformation system. Transformation of tobacco was done according to the protocol given by Horsch, *et al*., 1985. Very few plants were survived after transferring to rooting medium. PCR based screening for putative transformants were done for integration of hygromycin phosphotransferase in the transformants.

5.4.3.1 PCR based screening using hygromycin gene specific primers

DNA was isolated from plants, which survived on hygromycin supplemented media. PCR was performed to amplify hygromycin phosphotransferase gene from putative transgenic lines using hygromycin phosphotransferase gene specific primers. The forward primer named as HygBF (5'-GTCGACCTATTTCTTTGCCCTCGGAC-3') and reverse primer HygBR (5'-GGATCCCCTGACCTATTGCATCTCCC-3') was used. An amplicon of ~900 bp was amplified from plausible transformed plants (Figure 5.7).

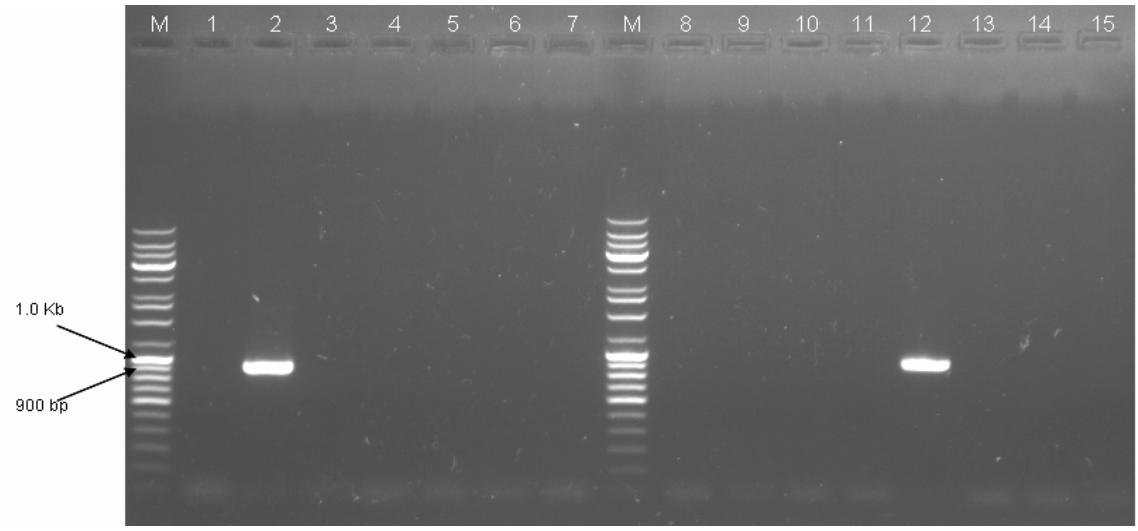


Figure 5.7 1% agarose gel shows PCR amplification of hygromycin resistance gene from plausible transformed tobacco plants (lane 2 and 3). No amplification was observed in other lanes. Medium range DNA size marker (Bangalore Genei, Bangalore, India) was separated in the lanes marked as M.

Two putative transformants of tobacco were characterized to be hygromycin gene positive. The less number of transformants and the less biomass, limits the further confirmation of the positive transformants either by ELISA or by slot blots.

5.4.4 Construction of plant transformation vector containing LlCAld5H gene in sense orientation

To express the complete gene and monitor its effect on the lignin biosynthesis and/or monomer composition with reference to *Leucaena,* the complete open reading frame of the LlCAld5H was cloned in sense orientation. As we have mentioned the cloning strategy of antisense, that is prior to final construction of gene of intrest in pCAMBIA 1301, the candidate gene was first cloned into the modified version of pCAMBIA1300, which contain the 35S promotor and terminator sequences to drive the gene expression. The similar strategy has been followed in the construction of gene in sense orientation as well. The sense construct was first cloned into the modified version of pCAMBIA-1300 (as described in antisense gene cloning) with *Kpn*I and *Bam*HI sites, as the presence of the *Sac*I site within the LlCAld5H gene limits, the cloning of the gene in *Kpn*I and *Sac*I. The construction of plant transformation vector in sense orientation could be stated as follow. The complete ORF of the gene from pGEMT-LlCAld5H clone was amplified with the introduction of *Kpn*I *and Bam*HI sites at 5' and 3' end of the gene respectively. The amplicon was cloned in pGEMT- vector. After digestion with *Kpn*I *and Bam*HI enzymes, it was subcloned in corresponding site of modified binary vector pCAMBIA1300 down stream to CaMV 35S promoter. The recombinant clone was refrred as *pCAMBIA1300-sCAld5H* and was confirmed by restriction digestion which released a ~1.5 Kb insert as depicted in figure 5.8. The sense construct with promoter and terminator sequences was taken out as an *EcoR*I – *Hin*dIII fragment and cloned into PCAMBIA1301. The recombinant plasmid thus constructed was termed as pCAMBIA-1301-*sCAld5H.*

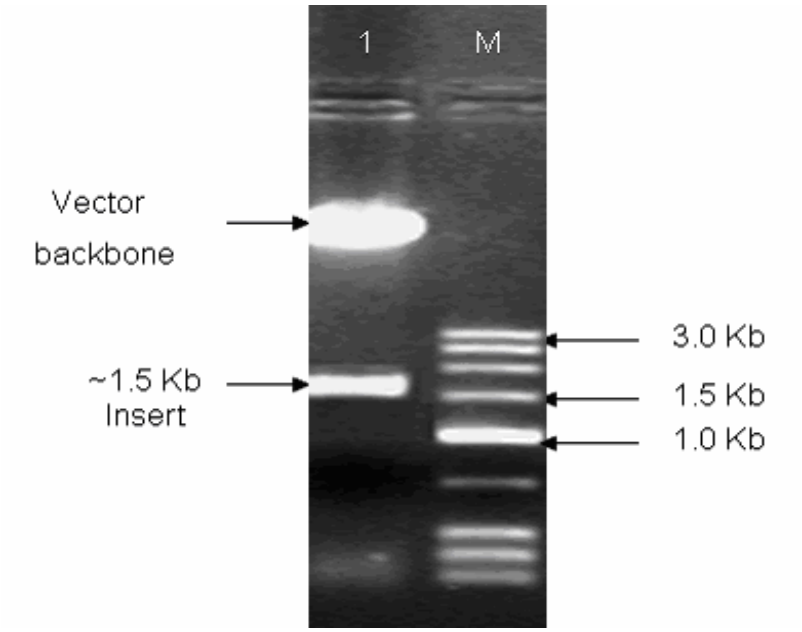


Figure 5.8. Full length gene fragment released as digestion product resolved on 1% agarose gel in lane 1, where as lane M shows DNA size marker from Biozymes.

In pCAMBIA1301-*sCAld5H* construct the insert was cloned in the sense orientation. The map of this construct between the left and right border is shown in figure 5.9. The LlCAld5H cDNA cloned in pCAMBIA 1301 plant transformation vector is under the transcriptional control of CaMV 35S promoter and poly-A sequences from nopaline synthase gene. The recombinant plant transformation vector further comprises a selection cassette containing hygromycin marker gene for selection of transformed plants and plant cells, also under the control of CaMV 35S promoter and poly-A sequences derived from nopaline synthase. The vector also includes a reporter gene construct comprising of GUS gene operably linked to CaMV 35S promoter and poly-A sequences from nopaline synthase gene. This construct was used for transformation of *Leucaena* plants. The construct pCAMBIA 1301+ *sCAld5H* gene was mobilized in *A. tumefaciens* strain (GV2260) for plant transformation. The constructs were mobilized into *Agrobacterium* by the freeze-thaw method.

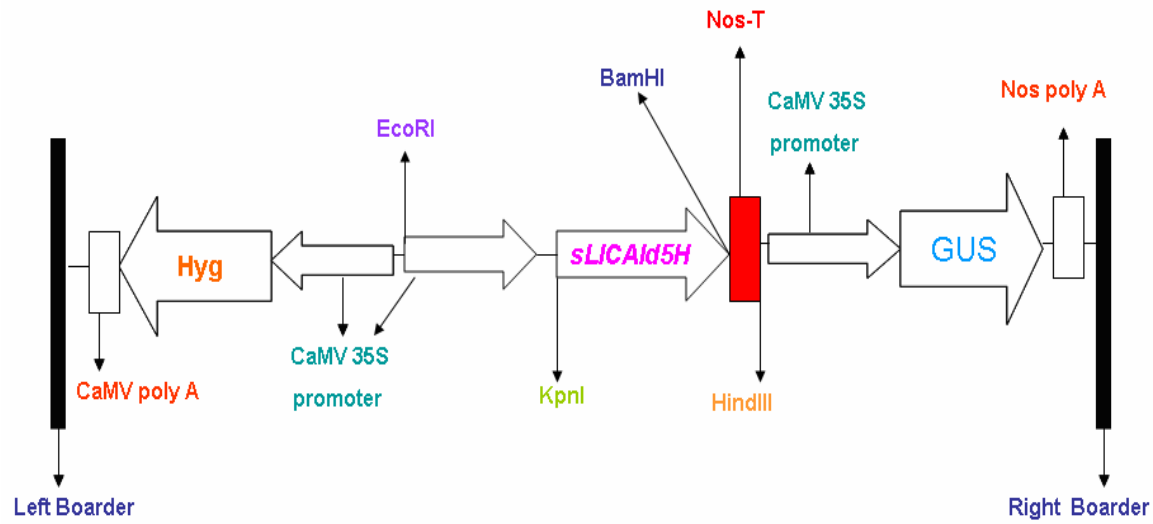


Figure 5.9. Diagrammatic representation of the sense constructs of pCAMBIA1301-*s*LlCAld5H between left and right boarder of the T-DNA.

5.4.5 Transformation of *Leucaena* and selection

Embryo axes of *Leucaena* were used as explants for transformation experiment as described previously. The non-transformed and transformed embryo axes (Particle bombardment / particle bombardment followed by co-cultivation /Agro-infusion) were kept on 1/2 MS + TDZ (0.5 mg/L) regeneration medium for one week. Then the embryo axes were shifted to selection medium containing hygromycin (10 mg/L) for 3 weeks. Transformed embryo axes were survived on selection medium (10 mg/L, hygromycin) while non transformed embryo axes turn black on selection medium (Figure 5.10 a). The survived embryo axes on hygromycin (10 mg/L) were further selected on hygromycin 15 mg/L for another 3 weeks. The survived explants on hygromycin (15 mg/L) were shifted to 1/2 MS without hygromycin selection. Cytokinins, 2ip (2-isopentenyl adenine, 0.5 mg/L) was used in the medium to have better elongation of transformed shoots. They were transferred to plastic cups with autoclaved soil and kept in culture room for few days before transferring to green house.

Out of 153 plants which survived on hygromycin supplemented media, 60 plants showed elongation on rooting media (Table 5.1). The transformed plants kept on root inducing medium are shown in figure 5. 11a. Majority of the transformants were survived but the growth was absolutely insignificant (Figure 5.10 d). Leaves of the plants that showed significant elongation on rooting media were normal in size when compared to nontransformed plant. Putative transformed plants when attained stem height of approximately 5-6 cm was transferred to greenhouse. Growth of putative transformed plants was observed to be less when compared to non-transformed plant.

5.4.6 Analysis of *Leucaena* transformants

The first step to analyze the transformed plant was the GUS assay. It was performed using transformed embryo axes and non-transformed embryo axes. Blue colored spots (Figure 5.12) were observed in most of the transformed embryo axes of *Leucaena* and whereas the non- transformed embryo axes were failed to display the particular blue spots. Primary screening was performed to check the integration of hygromycin phosphotransferase gene and *sCAld5H* gene. DNA of 21 plants, approximately of same height (5 - 6 cm) was isolated. Other plants were very weak and they needed few months to attain the biomass for these analysis. Due to the less biomass and small size attained by these plants, ELISA was performed from few selected transformed plants. Screening of putative transformants was done for integration of hygromycin phosphotransferase gene and *sCAld5H* gene by following methods;

1. PCR based screening using hygromycin phosphotransferase gene specific primers and *sCAld5H* specific primers.

2. ELISA with *CAld5H* protein extracted from transgenic lines of *Leucaena.*

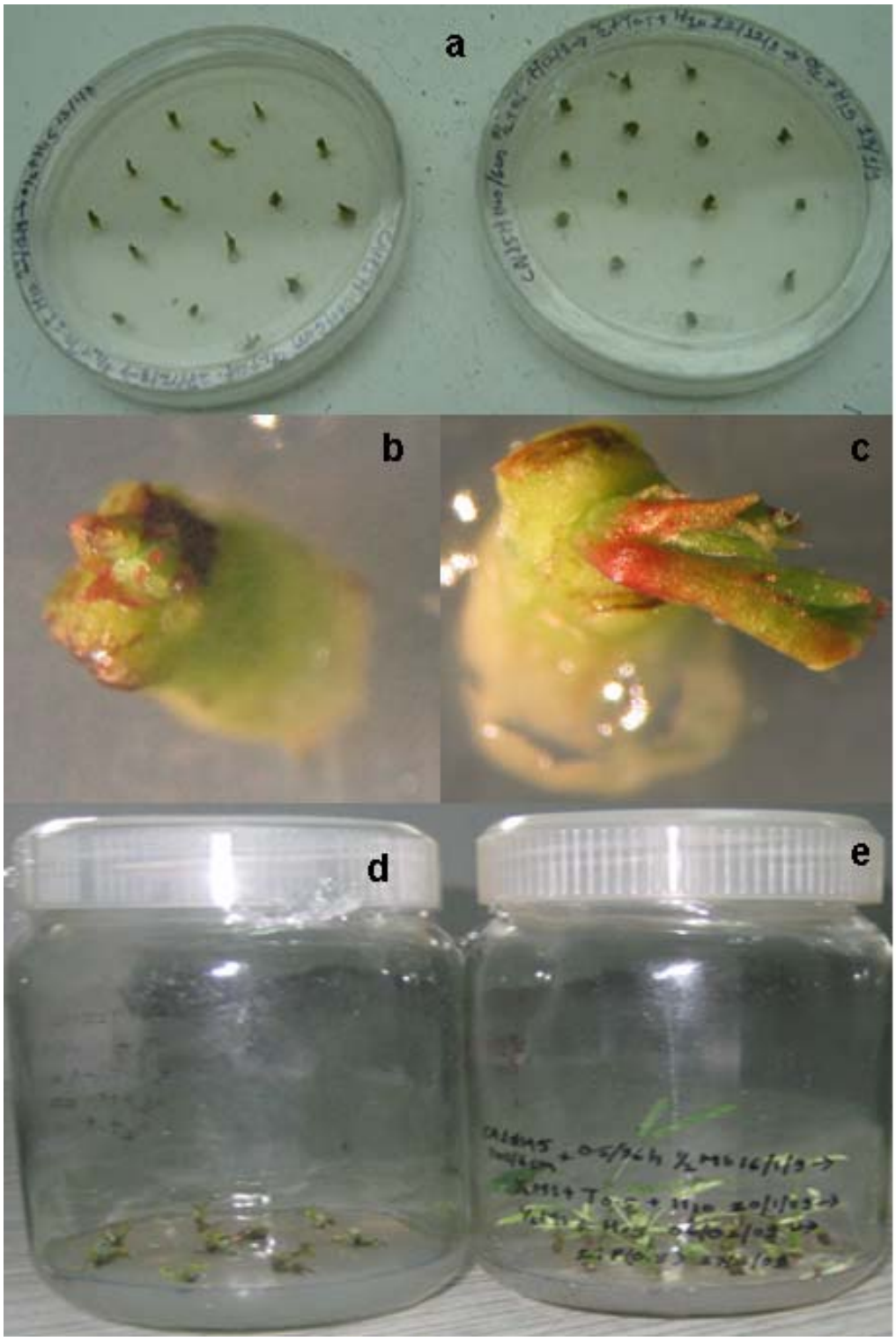


Figure 5.10. Putative transformed *Leucaena* shoots regenerated *in vitro* on selection medium. Transformed embryo axes on selection medium after two weeks (a), five weeks (b), 8 weeks (c) 12 weeks (d). 6 months old regenerated plants kept on nutrient media.

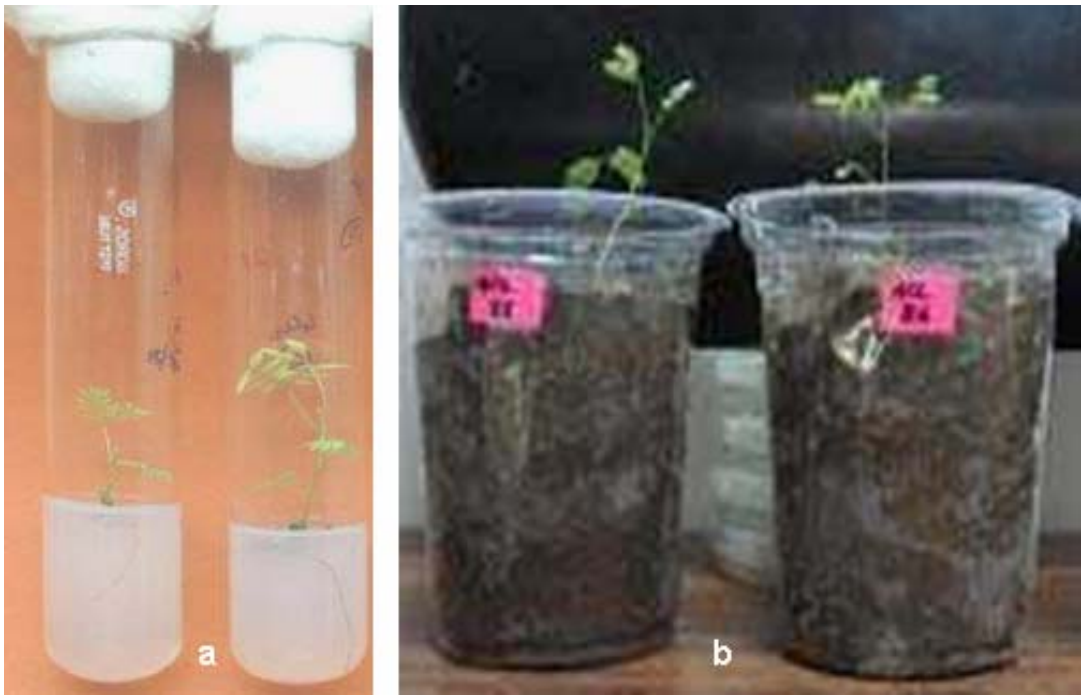


Figure 5.11. Regenerated plant on root inducing media **(a),** hardened transformed plants has shown in **b.**

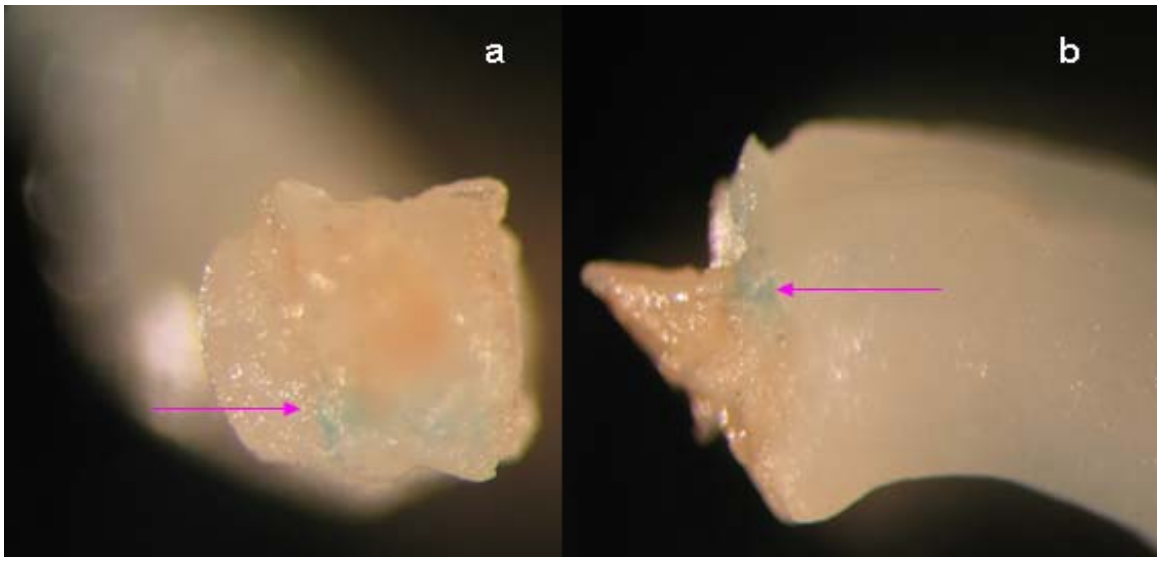


Figure 5.12. Embryo axes bombarded with *s*LlCAld5H gene construct coated particles, displaying the transient GUS expression as blue spots.

5.4.6.1. PCR based screening using hygromycin resistance gene specific primers

DNA was isolated from plants, which survived on hygromycin supplemented media. PCR amplification of hygromycin phosphotransferase gene from putative transgenic lines using hygromycin phosphotransferase gene specific primers was done. The forward primer HygBF: (5'-GTCGACCTATTTCTTTGCCCTCGGAC-3') and the reverse primer HygBR :(5'-GGATCCCCTGACCTATTGCATCTCCC-3') was used. An amplification product of 900 bp was obtained in putative transformants (Figure 5.13). Three putative transformants of *Leucaena* were characterized to be hygromycin gene positive. These putative transgenic plants were referred as LeuCAl-7, LeuCAl-32 and LeuCAl-36.

5.4.6.2 PCR based screening using LlCAld5H gene specific primers

To check the integration of LlCAld5H gene in the putative transformed plants, the amplification of the full-length candidate gene was done with the same g-DNA template that was used to amplify the hygromycin phosphotransferase gene in the previous experiment. The full-length gene was amplified from the two out of the 21 g-DNA used as template extracted from the putative transformed *Leucaena* (Figure 5.14). The size of the amplicon is exactly the same size of the open reading frame. The primers have failed to amplify the genomic clone of the gene as it did in the earlier trial with the genomic DNA template from non-transformed plant. The plausible reason of not getting the genomic clone may be that, the first axon is too small and the 3' reason of the primer might be falling into the intron reason where it could not anneal due to mismatch. The result obtained from PCR screening experiment was contradictory with reference to the actual numbers of transformed *Leucaena* plant, as three plants (LeuCAl-7, LeuCAl-32 and LeuCAl-36) were found to be positive for hygromycin phosphotransferase genes whereas only two of them (LeuCAl-7 and LeuCAl-36) showed amplification of the full length LlCAld5H gene. In an attempt to further confirmation of these transgenic lines the ELISA was done as mentioned in the next experiment.

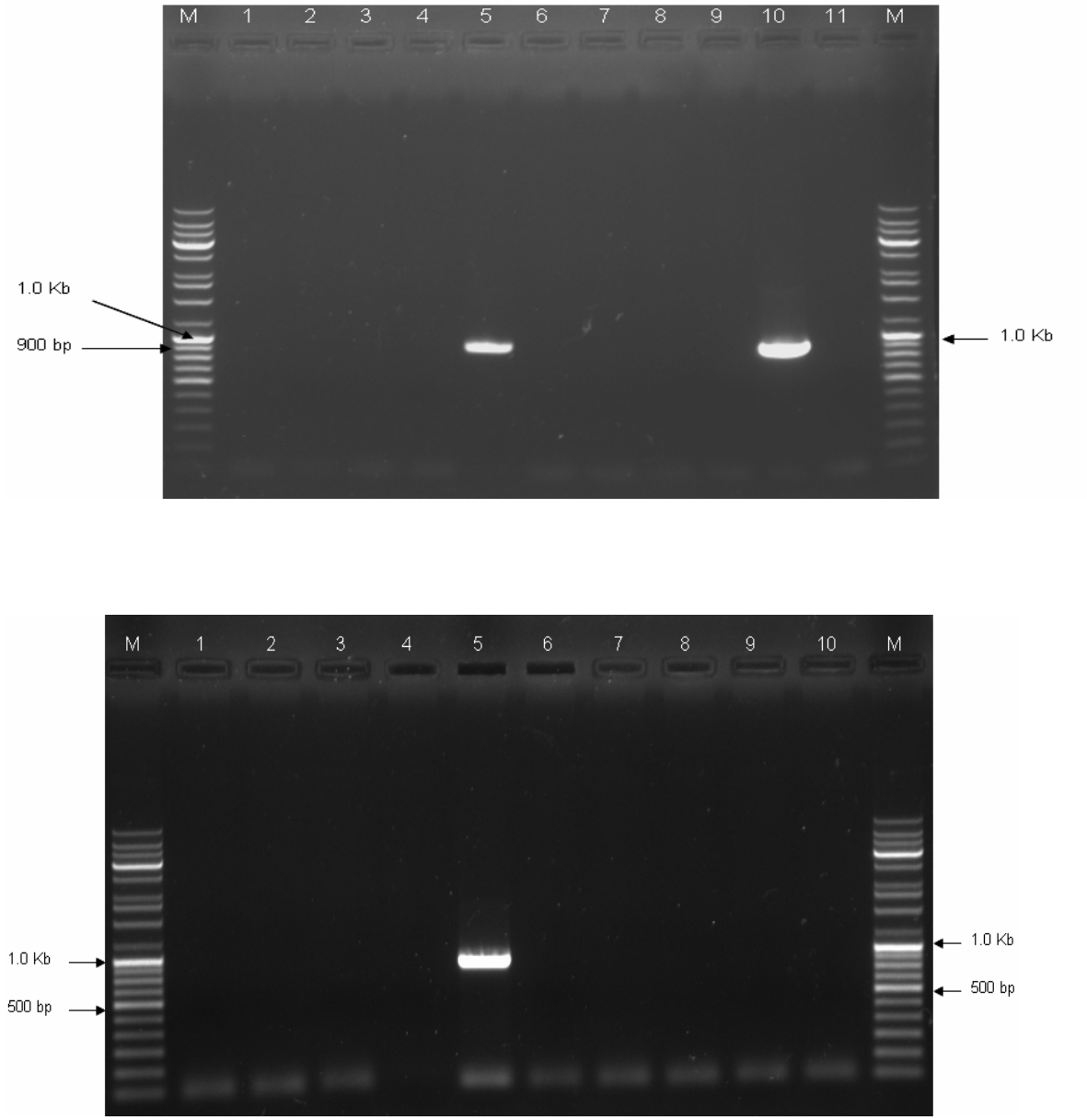


Figure 5.13. PCR amplification of hygromycin resistance gene from putative transformed *Leucaena* plants separated on 1% agarose gel (both panel). Lane 5 and 10 of upper panel and lane 5 of lower panel shows the amplification of the marker gene. No amplification was observed in other lanes. Medium range DNA size marker from Bangalore Genei, (India) was separated in the lanes marked as M in both the panel.

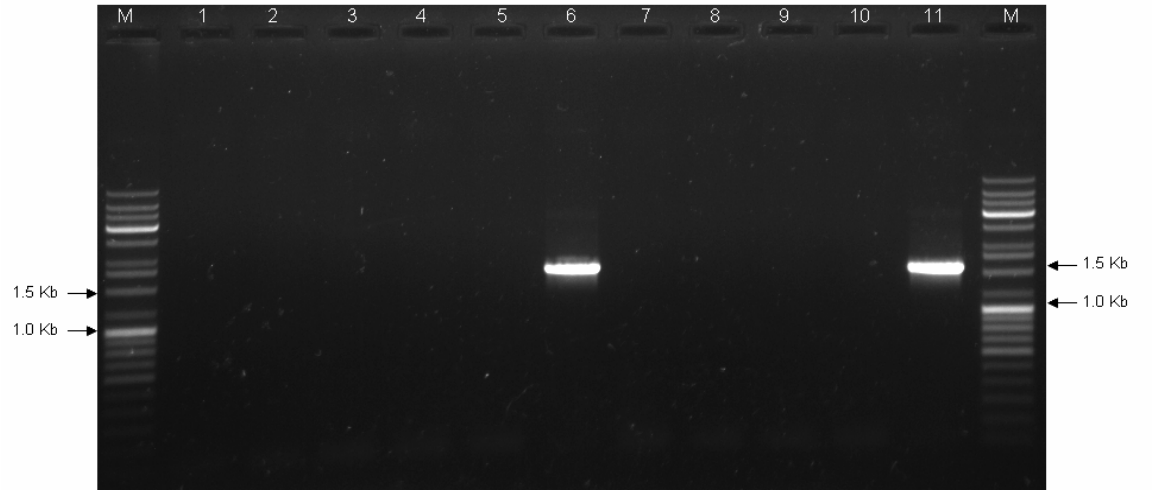


Figure 5.14. PCR amplification of LlCAld5H gene from putative transformed *Leucaena* plants separated on 1% agarose gel. Lane 6 and 11 of shows the amplification of the candidate gene. No amplification was observed in other lanes. Medium range DNA size marker from Bangalore Genei (Bangalore, India) was separated in the lanes marked as M in the first and last wells.

5.4.6.3 ELISA

ELISA was done using anti-LlCAld5H protein as a primary antibody. The total protein was extracted from the 2-4 leaves of the all the three putative transgenic plants mentioned in the previous section. Protein quantification was done according to Bradford methods. Crude protein (10 μg) was coated in each well of microtitre plate in triplicates. Equal quantity of crude protein, isolated from the non-transformed *Leucaena* was used as a control. The recombinant LlCAld5H protein (described in chapter 4) was introduced as positive control for ELISA. Primary antibodies and secondary antibodies were used at a dilution of 1:10,000 and 1:20,000 respectively. ELISA analyses of putative transformants have shown a significant increase in CAld5H protein level in two of the putative transgenic lines namely LeuCAl-7 and LeuCAl-36 as compared to non-transformed *Leucaena*.

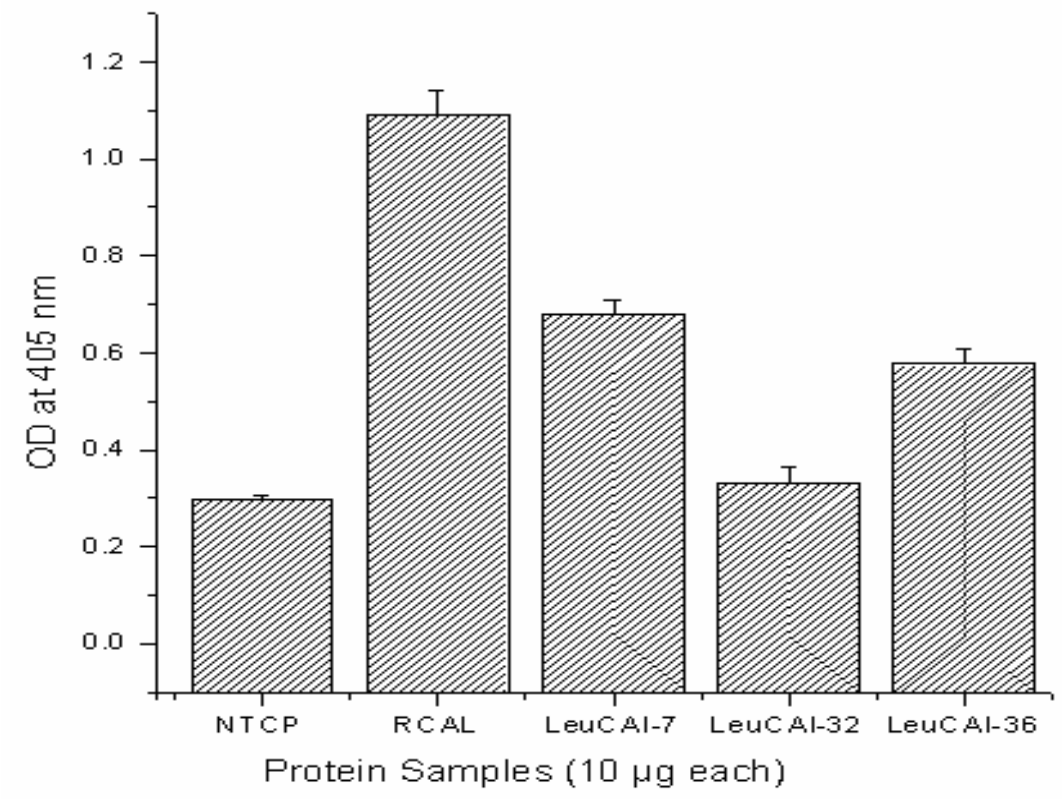


Figure 5.15. Graphical representation of ELISA profile of extracted LlCAld5H protein from transformants (LeuCAl-7, 32, 36). RCAL is the positive control sample and the NTCP is protein extracted from non transformants (X-axis). Values plotted on Y-axis are O.D at 405 nm. All values are plotted with standard deviation taken into account.

Table 5.1. Details of the number of *Leucaena* embryo axes used for transformation and PCR positives.

Gene	**Method used**	**Number of embryo axes used**	**Number of explants survived on selection (Hyg 15 mg/L medium)**	**No. of shoots elongated**	**Avg. shoot length (in cm)**	**Number of shoots used for DNA extraction and PCR**	**No. of samples that are PCR positive**	**Transform-ation efficiency confirmed through PCR (%)**
sCAld5H	Particle bombardment	74	54	25	3.47	7	2	2.7
sCAld5H	Particle bombardment + co-cultivation	99	74	31	2.79	14	1	> 1
sCAld5H	Agro infusion	34	5	4	2.00	2	0	0
	Total	**207**	**133**	**60**	**8.26**	**23**	**4**	**~4**

5.4.7. Discussion

To meet the increasing demand of high quality wood for paper industry, it is essential to provide genetically modified plant species and the present research work emphasizes on the same issue. Lignin biosynthesis genes are also being targeted to produce such plants using antisense and siRNA technology with low lignin content. Down-regulation of various key genes has been achieved to obtain transgenic plant with low or modified lignin content. In contrast to the other lignin biosynthetic genes, role of *CAld5H* gene is such that it's over expression will be good to achieve highly reactive lignin, which in turn will be good for pulp industries (Li *et al.,* 2003). Hence, work was initiated with the objective of cloning and characterizing *CAld5H* cDNA clones from *Leucaena* and also for its up regulation to raise transgenic lines for proper utilization in paper industry. Transformation protocol for *Leucaena* and tobacco was standardized in our laboratory. Transformed plants were analyzed using GUS assay, PCR and ELISA based experiments. GUS enzyme assay was used for primary screening of transformants followed by PCR screening. Two transgenic lines in tobacco and three transgenic lines in *Leucaena* were screened in primary experiments. In the plant transformation process, every transformation event behaves as an individual event and has to be different from other events *i.e.* two transformation events for the same gene need not behave identically. It is not necessary to get a plant positive for both the genes. Plant may be positive for only hygromycin phosphotransferase gene integration, only *sCAld5H* gene integration or both. As evident in case of *Leucaena*, three putative transformants showed integration of hygromycin phosphotransferase gene in the PCR screening, but only two of them showed integration of *sCAld5H* gene. ELISA studies reveal the increased level of CAld5H protein in two of the putative transgenic lines. Further analysis such as RT-PCR and Southern blot analysis is required indeed for confirming the integration of *CAld5H* gene in sense orientation to achieve transgenic lines up- regulated for LlCAld5H gene. It will take at least few more months to nurture the putative transgenic *Leucaena* in greenhouse, so that further analysis, such as Southern blot and lignin estimation (qualitative and quantitative) could be done to establish *Leucaena* transgenic lines up-regulated for *CAld5H*.

5.5. Conclusion

Primary screening of transformants were done by GUS enzyme assay that was followed by PCR and ELISA based screening methods. Three transgenic lines of *Leucaena* and two transgenic lines of tobacco have been established which will be further analyzed for abundance of LlCAld5H transcript using real-time PCR and lignin estimation (quantity and quality) to establish the transgenic lines of *Leucaena* and tobacco for up and down regulated of *CAld5H* gene respectively. Transgenic lines of *Leucaena* producing altered lignin monomer ratio will be utilized for paper and pulp industry in future.

Summary

A central and challenging question in paper and pulp industries is how to get rid from lignin from cellulose fibers. These two are the major cell wall constituents in plants. Pulp making process involves removal of lignin which also costs a huge loss of cellulose fiber. Pulping efficiency depends on lignin reactivity (which is associated with syringyl/guaiacyl S/G monolignol ratio). Wood- pulping kinetics revealed that every unit increase in the S/G ratio would roughly doubles the rate of lignin removal (Chapple *et al*,.1992). Considerable evidence is now available that shows that in angiosperm trees, the syringyl monolignol pathway branches out from guaiacyl pathway through coniferaldehyde and is regulated in sequence by three genes encoding coniferaldehyde 5-hydroxylase (CAld5H), 5-hydroxyconiferaldehyde *O*-methyltransferase (COMT) and sinapyl alcohol dehydrogenase (SAD). It has also been revealed by enzyme kinetics studies that CAld5H has a 6-to 50- times slower turn over rate than the other two syringyl pathway enzymes (Li *et al*,.2000), pointing to a key role for CAld5H in limiting syringyl monolignol biosynthesis and , therefore, lignin S/G ratio.

Paper industry in India mainly uses bamboos, *Eucalyptus* sp., *Casuarina* sp. and *Leucaena* sp. as a source for paper pulp. *Leucaena* is exclusively used in India and about 25% of raw material for pulp and paper industry comes from this hard wood tree.

This study is the first instance towards isolation and characterization of lignin biosynthetic pathway gene(s) for the development of transgenic *Leucaena* plants. This study is intended to lay down criteria for the development of the transgenic plants which would ultimately strive for: (a) altered ratio of S/G lignin (b) altered lignin content.

To achieve the objective, a partial cDNA clone, LlCAld5H was isolated and sequence of the clone was submitted to NCBI GenBank with accession no. DQ986906. Southern hybridization and slot blot analysis with homologous probes reveals that the CAld5H gene in *L. leucocephala* exists as a gene family of possibly 2-3 members. Rapid amplification of both 5’ and 3’ ends of the cDNA ends was done and the sequences corresponding to 5’ and 3’ regions of the gene were isolated. A complete open reading frame of 1536 bp of the gene was isolated, cloned and characterized. The GenBank accession no. of full length gene is EU041752. Approximately 300 bp of 3’ untranslated

region (3'UTR) and a small 5' UTR were isolated and characterized. The clone containing complete ORF of the CAld5H from *Leucaena leucocephala* was referred as LlCAld5H. The LlCAld5H nucleotide shows 77% sequence homology withCAld5H gene of sweetgum. The LlCAld5H encodes a protein of 511 amino acids with molecular weight of 57.24 KDa and the theoretical pI (isoelectric point) of the protein is 6.52. The amino acids sequence shows 83% sequence identity with CYP84A16 of *Glycine max*. The CAld5H protein sequence displays all the characteristic features of a plant P450 protein including Heme-binding ligands, Stop transfer sequence etc. LlCAld5H protein and the LlCAld5H gene are phylogenetically closer to CYP84A16 of *Glycine max* (both the plants belongs to *Fabaceae* family).

The signal P prediction revealed the presence of a signal peptide of 29 amino-acids. LlCAld5H sequence encoding the mature coniferaldehyde-5-hydroxylase protein was cloned in pET 28a (+) expression vector. The recombinant protein was purified in denatured condition and used to raise polyclonal antibody against it. The candidate protein, coniferaldehyde-5-hydroxylase was detected in *Leucaena leucocephala* by Western blotting. Optimization of expression of native protein was done and the best expression was achieved when bacterial cultures were grown in Terrific broth for 24-30 h after induction with 0.03 mM IPTG at 25°C. Specific activity with crude extract was estimated. The homology modeling of LlCAld5H protein sequence was done based on the template Human Microsomal Cytochrome P450 1A2 (2HI4). The RMSD value less than 1.00 showed that although the identity of the template was less than 30%, the model built was good in all the domains modeled in the predicted structure. The PROCHECK analysis showed 88.7%, 9.1%, 2.0% residues in most favorable, additionally allowed and generously allowed regions, respectively, of the Ramachandran map. Only one residue was found to lie in the disallowed region. The analysis of the interactions at the ligand binding site revealed that the Hydrogen bond between Phe441 and Cys 448 is crucial in maintaining the structure of the domain.

Spatio-temporal expression profile was investigated using real-time PCR analysis after optimizing a plethora of conditions such as primer concentration, RNA extraction, cDNA preparations and PCR cycling conditions. It was found that LlCAld5H gene transcript was expressed in radical, root, stem and to a slightly lowere extent in leaves. However, a unique pattern of expression is exhibited by stem and leaves as in both the tissue the expression is increasing from 5-10 days but the 20 days old plants showed the decline in expression followed by further increase in expression of 30 days samples. LlCAld5H protein was immuno-cytolocalized in tissues undergoing lignification.

Genetic transformation was done in Leucaena and tobacco with sense and antisense construct respectively to study the effect on lignin quality and quantity. Primary screening of transformants were done by GUS enzyme assay that was followed by PCR and ELISA based screening methods. Three transgenic lines of *Leucaena* and two transgenic lines of tobacco have been established which will be further analyzed for abundance of LlCAld5H transcript using real-time PCR and lignin estimation (quantity and quality) to establish the transgenic lines of *Leucaena* and tobacco for up and down regulated of *CAld5H* gene respectively. Transgenic lines of *Leucaena* producing altered lignin monomer ratio will be utilized for paper and pulp industry in future.

Future Prospects

1. Over expressed LlCAld5H protein could be purified in soluble form and substrate specificities may be determined for different substrates of LlCAld5H.
2. Putative transformed *Leucaena* plants could be analyzed for quality and quantity of lignin being produced.
3. Transgenic lines of *Leucaena* most suited for paper industry could be utilized for paper production.

References

Abbott JC, B. A., Pincon G, Legrand M, Lapierre C, Mila I, Schuch W, Halpin C. (2000). "Simultaneous suppression of multiple genes by single Transgenes: Downregulation of three unrelated lignin biosynthetic genes in tobacco." Plant Physiology 128: 844-853.

Abbott JC, B. A., Pincon G, Legrand M , Lapierre C, Mila I, Schuch W, Halpin C (2002). "Simultaneous suppression of multiple genes by single transgenes. Downregulation of three unrelated lignin biosynthetic genes in tobacco." Plant Physiology 128: 844-853.

Abravaya K, H. J., Marshall R (2003). "Molecular beacons as diagnostic tools: technology and applications." Clin Chem Lab Med. 41:468-474.

Adler E. (1977). "Lignin chemistry-past present and future." Wood Sci.Technol 11: 169-218.

Akhtar M, S. G., Lentz MJ, Horn E, Scaney RE, Kirk, TK (1998). "Commercialization of biopulping for mechanical pulping." Biotechnology in the Pulp and Paper Industry I.

Alpeter, E., Mackben, M., & Trauzettel-Klosinski, S. (2000). "The importance of sustained attention for patients with maculopthies." Vision Research 40: 1539-1547.

An (1987). "Binary Ti vectors for plant transformation and promoter analysis." Meth Enzymol. 153: 292-305.

Anterola AM, L. N. (2002). "Trends in lignin modification: a comprehensive analysis of the effects of genetic manipulations/mutations on lignification and vascular integrity." Phytochemistry 61: 221-294.

Arnheim N, E. H. (1992). "Polymerase Chain Reaction Strategy." Annu Rev Biochem. 61:131-156.

Bao W, O. M. D., Whetten R, Sederoff RR (1993). "A laccase associated with lignification in loblolly pine xylem." Science 260: 672-74.

Baribault, T. J. S., K. G. M. Cain, P. A. and Scott, N.S.() "Transgenic Grapevines: Regeneration of shoots expressing -Glucuronidase." Journal of experimental botany 8: 1045-1049.

Baucher M, H. C., petit-Conil M, Boerjan W (2003). "Lignin: Genetic Engineering and Impact on Pulping." Critical Review of Biochem. Mol. Biol 38(305-350).

Baucher M, H. C., Petit-Conil M, Boerjan W. (2003,). "Lignin: Genetic Engineering and Impact on Pulping." Critical Reviews in Biochemistry and Molecular Biology 38: 305-350.

Baucher M, M. B., Van Montagu M, Boerjan W (1998). "Biosynthesis and genetic engineering of lignin." Critical Review of Plant Science 17(125-97).

Boerjan W, R. J., Baucher M (2003). "Lignin Biosyn." Annu. Rev. Plant Biol 54(519-546).

Borson, N., Sato WL, Drewes LR (1992). "A lock-docking oligo (dT) primer for 5' and 3' RACE PCR." PCR Methods Appl. 2: 144-148.

Boudet AM, G. D., Grima-Pettenati J (1997). "In Lignin and Lignification: Recent biochemical and biotechnological advances." C R Acad. Sci. Paris 319(317-331).

Brazeau, D. A. (2004). "Combining genome wide and targeted gene expression profiling in drug discovery: microarrays and real-time PCR." Drug Discov. Today, 9: 838-845.

Breen A, S. F. (1999). "Fungi in lignocellulose breakdown and biopulping." Current Opinion in Biotechnology, 10: 252-258.

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

Bubner, B., Baldwin, I.T. (2004). "Use of real-time PCR for determining copy number and zygocity in transgenic plants." Plant Cell Rep. 23: 263-271.

Busam G, G. B., Kneusel RE, Matern U: (1997). Plant Physiol, 113:: 1003-1005.

Bustin, S. A. (2002). "Quantification of mRNA using real-time reverse transcription PCR (RT-PCR): trends and problems." J. Mol. Endocrinol. 29: 23-39.

Bustin, S. A., Nolan, T., (2004). "Pitfalls of quantitative real time reverse transcription polymerase chain reaction." J. Biomol. Technol. 15: 155-166.

Campbell MM, S. R. (1996). "Variation in lignin content and composition. Mechanisms of control and implications for the genetic improvement of plants." Plant Physiol. 110: 3-13.

Carsten Kegler, K. G. A. R. M. (2006). "Establishment of a real-time PCR protocol for expression studies of secondary metabolite biosynthetic gene clusters in the G/C-rich myxobacterium Sorangium cellulosum So ce56." Journal of Biotechnology 121:: 201-212.

Chabannes M, B. A., Lapierre C, Marita JM, Ralph J, Pean M, Danoun S, Halpin C, Grima-Pettenati J, Boudet AM: (2001). "Strong decrease in lignin content without significant alteration of plant development is induced by simultaneous down-regulation of cinnamoyl CoA reductase (CCR) and cinnamyl alcohol dehydrogenase (CAD) in tobacco plants." Plant Journal. 28: 257-270.

Chalfie M, Y. T., Euskirchen G, Ward WW, and Prasher DC. "Green fluorescent protein as a marker for gene expression." Science, 263, Issue ,(5148): 802-805.

Chang H M, S. K. V. (1973). Techn Assoc Pulp Pap Ind 56: 132-143.

Chapple CCS, V. T., Ellis BE, Somerville CR (1992). "An Arabidopsis mutant defective in the general phenylpropanoid pathway." Plant Cell 4: 1413-1424.

Chen C, M. H., Burggraeve B, De Rycke RM, Inoue K, Vleesschauwer VD, Steenackers M, Van Montagu MC, Engler GJ, Boerjan, WA (2000). "Cell-specific and conditional expression of caffeoyl-coenzyme A-3-O-methyltransferase in poplar." Plant Physiol 123: 853-867.

D., H. K. A. S. (1989). "Physiology and molecular biology of phenyl propanoid metabolism." Annual review of plant physiology and plant molecular biology 40: 347-369.

Daguma B, K. B., Okali DUU (1988). "Factors affecting germination of *Leucaena leucocephala* (Lam.) de Wit. seed." Seed Science and Technology 16: 489-500.

Dauwe R, M. K., Goeminne G, Gielen B, Rohde A, Beeumen JV, Ralph J, Boudet A-M, Kopka J, Rochange SF, Halpin C, Messens E, Boerjan.W. (2007). "Molecular phenotyping of lignin-modified tobacco reveals associated changes in cellwall metabolism, primary metabolism, stress metabolism and photorespiration." The Plant Journal 52: 263-285.

DD, B. (1987). "Creating a ribonuclease-free environment." Meth. Enzymol. 152: 20-24.

Delmer DP, A. Y. (1995). "Cellulose biosynthesis." Plant Cell. 7: 987-1000.

Dhar, A. K, K.S. Veazey, G. and Read, B. (2009). "Validation of reference gene for quantification of immune gene expression in shrimp." Molecular Immunology 46:(8-9:): 1688-95.

Dharmawardhana DP, E. B., Carlson JE: (1995). "A beta-glucosidase from lodgepole pine xylem specific for the lignin precursor coniferin." Plant Physiology 107: 331-39.

Dixon RA, C. F., Guo D, Parvathi K (2001). "The biosynthesis of monolignols: a "metabolic grid," or independent pathways to guaiacyl and syringyl units?" Phytochemistry 57: 1069-1084.

D.R N. (1999). "Cytochrome P450 and the individuality of species." Arch Biochem Biophys 369: 1-10.

Dwivedi UN, C. W., Yu J, Datla RSS, Bugos RC, Chiang VL, Podila GK (1994). "Modification of lignin biosynthesis in transgenic Nicotiana through expression of an antisense O-methyltransferase gene from Populus." Plant Mol. Biol. 26: 61-71.

E., Z. (2008). "Molecular cloning, characterization and expression analysis of three aldehyde oxidase genes from Pisum sativum L." Plant Physiology and Biochemistry 46: 19-28.

Eckert KT, K. T. (1990). "High fidelity DNA synthesis by the Thermus aquaticus DNA polymerase." Nucleic Acids Res. 18:: 3739- 3744.

Elkind Y, E. R., Mavandad M, Hedrick SA, Ribak O, Dixon RA, Lamb CJ: (1990,). "Abnormal plant development and down-regulation of phenylpropanoid biosynthesis in transgenic tobacco containing a heterologous phenylalanine ammonia-lyase gene." Proc Natl. Acad. Sci.(USA). 87: 9057-9061.

Elkind Y, E. R., Mavandad M, Hedrick SA, Ribak O, Dixon RA, Lamb CJ. (1990,). "Abnormal plant development and down-regulation of phenylpropanoid biosynthesis in transgenic tobacco containing a heterologous phenylalanine ammonia-lyase gene." Proceedings of National Academy Science U S A. 87(22): 9057- 9061.

EO, S. (1987). "A method of retaining phloroglucinol proof of lignin." Stain Technol. 62: 279-280.

Feinbeng AP, V. B. (1983). "A technique for radiolabelling DNA restriction endonucleases fragments to high specific activity." Anal Biochem. 132: 6-13.

Feinbeng AP, V. B. (1984). "A technique for radiolabelling DNA restriction endonucleases fragments to high specific activity." Addendum.Anal BioChem. 137: 266-267.

Fengel D, W., G: (1984). "Lignin in Wood, Chemistry, Ultrastructure, Reactions." pp 132-181.

Fergus BJ, G. D. (1970). "The location of guaiacyl and syringyl lignins in birch xylem tissue." Holzforschung. 24: 113-17.

Feyereisen, D. W.-R. a. R. (2000). "Cytochrome P450: a success story." Genome Biology 1(6): 30031-3003.9.

Feyereisen D. W.-R. a. R. (2000). "Cytochromes P450: a success story." Genome Biology. 1(6): 3003.1-3003.9.

Franke R, H. M., Denault JW, Ruegger MO, Humphreys JM, Chapple C: b, (2002). "Changes in secondary metabolism and deposition of an unusual lignin in the ref8 mutant of Arabidopsis." Plant Journal. 30: 47-59.

Franke R, M. C., Meyer K, Shirley AM, Cusumano JC, Chapple C: (2000). "Modified lignin in tobacco and poplar plants over-expressing the Arabidopsis gene encoding ferulate 5-hydroxylase." Plant Journal. 22: 223-234.

Freeman WM, W. S., Vrana KE (1999). "Quantitative RT-PCR: pitfalls and potential." Biotechniques. 26: 112-122, 124-125.

Freudenberg K, N. A., ed. (1968). "Constitution and Biosynthesis of Lignin." Berlin: Springer-Verlag.: 129.

Freudenberg K, N. A., ed. Springer-Verlag. (1968). Constitution and Biosynthesis of Lignin. Berlin: pp 129.

Fucharoen S, F. G., Fucharoen P, Fukumaki Y (1989). "A novel ochre mutation in the beta-thalassemia gene of a Thai. Identification by direct cloning of the entire beta-globin gene amplified using polymerase chain reactions." J Biol Chem. 264: 7780-7783.

Ginzinger, D. G. (2002). "Gene quantification using real-time quantitative PCR: An emerging technology hits the mainstream." Experimental hematology 30: 503-512.

Goffner D, C. M., Campargue C, Clastre M, Borderies G, Boudet A, Boudet AM: (1994). "Purification and characterization of cinnamoyl-CoA:NADP oxidoreductase in Eucalyptus gunnii." Plant Physiol. 106: 625-632.

Goodenow M, H. 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.

Goujon T, F. V., Mila I, Pollet B, Ruel K, Burlat V, Joseleau J-P, Barri`ere Y, Lapierre C, Jouanin L: (2003). "Down-regulation of the AtCCR1 gene in Arabidopsis thaliana: effects on phenotype, lignins and cell wall degradability." Planta. 217: 218-228.

Graham SE, P. J. (1999). "How similar are P450s and what can their differences teach us." Arch Biophys Biochem 369: 24-29.

Grand C, B. A., Boudet AM: (1983). "Isoenzymes of hydroxycinnamate: CoA ligase from poplar stems: properties and tissue distribution." Planta. 158: 225-229.

Grima-Pettenati J, G. D. (1999). "Lignin genetic engin. revisited." Plant Science 145: 51-65.

Guo D, C. F., Inoue K, Blout JW, Dixon RA: (2001). "Downregulation of caffeic acid 3-Omethyltransferase and caffeoyl CoA 3-O-methyltransferase in transgenic alfalfa: impacts on lignin structure and implications for the biosynthesis of G and S lignin." Plant Cell. 13:(73-88.).

Gupta H.K, A. 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." Animal Feed Science and Technology. 78: 159.

Halpin C, H. K., Chojecki J, Oliver D, Chabbert B, Monties B, Edwards K, Barakate A, Foxon GA: (1998). "Brown-midrib maize (bm1)-a mutation affecting the cinnamyl alcohol dehydrogenase gene." Plant Journal. 14: 545-553.

Halpin C, K. M., Foxon GA, Campbell M M, Boudet AM, Boon JJ, Chabbert B, Tollier MT and Schuch W. (1994). "Manipulation of lignin quality by down-regulation of cinnamyl alcohol dehydrogenase." Plant Journal 6: 339-350.

Harris PJ, K. M., Kendon MF and McKenzie RJ (1997). Biochem Syst Ecol. 25: 167-179.

Haseloff, J, Siemering, K. R., Prasher, D. C., Hodge, S. (1997). "Removal of a cryptic intron and subcellular localisation of green fluorescent protein are required to mark transgenic Arabidopsis plants brightly." Proc. Natl. Acad. Sci. U.S.A.

He L, T. N. (1991). "Formation and structure of lignin in monocotyledons. IV. Deposition process and structural diversity f the lignin in the cell wall of

sugarcane and rice plant studied by ultraviolet icroscopic spectroscopy." Holzforschung. 45: 191- 98.

Heid CA, S. J., Livak KJ, Williams PM (1996). "Real time quantitative PCR." Genome Res. 6: 986-994.

Heller HC, Orians GH, P. W., Sadava D."Life:The science of Biology,." 6th edition Sunderland (MA): Sinauer Associates, Inc; 251: 322-323.

Henrik Nielsen, J. E., Søren Brunak and Gunnar von Heijne. (1997). Protein Engineering, 10: 1-6.

Hiei Y, K. T., and Kubo T, (1997). "Transformation of rice mediated by Agrobacterium tumefaciens." Plant molecular biology 35(1-2): 205-18.

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

Highley TL, D. W. (1998). "Biotechnology in the study of brown-and white-rot decay." In Forest Products Biotechnology.: 15-36.

Higuchi R, D. G., Walsh PS, and Griffith R (1992). "Simultaneous amplification and detection of specific DNA sequences." Biotechnology. 10: 413-417.

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

Higuchi, T. (1990). "Lignin biochemistry: biosynthesis and biodegradation." Wood Sci Technol , 24: 23-63.

Horch R, F. J., Hoffmann N, Eichholtz P, Rogers R, Fratey T (1985). "A simple method for transferring genes into plants." Science. 227: 1229-1231.

Horsch RB, F. J., Hoffmann NL, Eichholtz D, Rogers SG and Fraley RT. (1985). "A simple and general method for transferring genes into plants." Science 227: 1229-1231.

Hu W-J, H. S., Lung J, Popko JL, Ralph J, Stokke DD, Tsai C-J, Chiang, VL (1999). "Repression of lignin biosynthesis promotes cellulose accumulation and growth in transgenic trees." Nat. Biotechnology 17: 808-812.

Hu W-J, H. S., Lung J, Popko JL., Ralph J, Stokke DD, Tsai C-J, Chiang, VL: (1999). "Repression of lignin biosynthesis promotes cellulose accumulation and growth in transgenic trees." Nature Biotechnology. 17: 808-812.

Humphreys JM, C. C. (2002). "Rewriting the lignin roadmap." Current Opinion in Plant Biology. 5: 224-229.

Humphreys JM, H. M., Chapple, C: (1999). "New routes for lignin biosynthesis defined by biochemical characterization of recombinant ferulate 5-hydroxylase, a multifunctional cytochrome P450-dependent monooxygenase." PNAS (USA). 96: 10045-10050.

Imai, T., H.Globerman, J.M. Gertner, N.Kagawa, and M.R.Waterman. (1993). " Expression and purification of functional human 17 alpha-hydroxylase/17, 20-lyase (P450c17) in Escherichia coli. Use of this system for study of a novel form of combined 17-alpha -hydroxylase/17, 20-lyase deficiency." J. Biol.Chem 268: 19681-19689.

J, G. (1985). "Chemistry of delignification. Part 1: General concept and reaction during pulping." Wood Sci. Technol. 19: 289-312.

J, S. (2003). "Evolution of water transport and xylem structure." International Journal of Plant Science. 164: 115-S127.

JA, R. (1977). "The evolution of vascular land plants in relation to supracellular transport processes." Adv Bot. Res. 5: 153-219.

Jannick Dyrlov Bendtsen, H. N., Gunnar von Heijne and Soren Brunak. (2004). J. Mol. Biol, 340: 783-795.

Jansson, I., I. Stoilov, M.Sarfarazi, and J.B. Schenkman, (2000). "Enhanced expression of CYP1B1 in Escherichia coli." Toxicology 144: 211-219.

Jefferson RA, K. T., Bevan MW (1987). "GUS fusions: β-glucuronidase as a sensitive and versatile gene fusion marker in higher plants." EMBO J 6: 3901-3907.

Jones L, E. A., Turner SR. (2001). "Cloning and characterization of irregular xylem4 (irx4): a severely lignin deficient mutant of Arabidopsis." Plant Journal 26: 205-216.

Joseleau J-P, R. K. (1997). "Study of lignification by noninvasive techniques in growing maize internodes. An investigation by Fourier transform infrared cross-polarizationmagic angle spinning 13C-nuclear magnetic resonance spectroscopy and immunocytochemical transmission electron microscopy." Plant Physiol. 114: 1123- 1133.

Jouanin L, G. T., de Nada V, Martin M-T, Mila I, Vallet C, Pollet B, Yoshinaga A, Chabbert B, Petit-Conil M, Lapierre C: (2000). "Lignification in transgenic poplars with extremely reduced COMT activity." Plant Physiol. 123: 1363-1373.

K, E. (1977). "Anatomy of Seed Plants,." Edn 2nd. San Francisco: John Wiley & Sons;

Kaeppler H.F., M. G. K., Skadsen R.W., Nuutila A.M., Carlson A.R. (2000). "Genetic engineering of oat via visual selection of cells expressing green fluorescent protein." Plant Cell Rep.

Kahn R, D. F. (2000). "Function and evolution of plant cytochrome P450. Recent Adv Phytochem." Recent Adv. Phytochem. 34: 151-189.

Kahn R, D. F. R. A. (2000). "Function and evolution of plant cytochrome P450." Phytochem. 34: 151-189.

Kajita S, K. Y., Omori S: (1996). "Alterations in the biosynthesis of lignin in transgenic plants with chimeric genes for 4-coumarate: coenzyme A ligase." Plant Cell Physiol. 37: 957-965.

KB, M. (1990). "The unusual origin of the polymerase chain reaction." Sci Am. 262: 56-61.

K.B. M. (1990). Sci Am, 262: 56-61.

Kellogg DE, R. I., Chen S, Mukhamedova N, Vlasik T, Siebert P, Chenchik A (1994). "Hot start PCR facilitated by a neutralizing monoclonal antibody directed against Taq DNA polymerase." BioTechniques. 16: 1134-1137.

Kimmel AR, B. S. (1987). "Preparation of cDNA and the generation of cDNA libraries." Overview. Meth Enzymol. 152: 307-316.

KM, C. (1994). "Reverse transcriptase can inhibit PCR and stimulate primer dimer formation." PCR Methods Appl. 4: 62-64.

Knut Meyer, J. C. C., Chris Somerville and Clint C.S.Chapple: (1996). "Ferulate-5-hydroxylase from Arabidopsis thaliana defines a new family of cytochrome P450-dependent monoxygenases." PNAS. (USA), 93: 6869-6874.

Korth KL, B. J., Chen F, Rasmussen S, Lamb C, Dixon RA: (2001). "Changes in phenylpropanoid metabolites associated with homology dependent silencing of phenylalanine ammonialyase and its somatic reversion in tobacco." Physiol Plant. 111: 137-143.

Kwok S, K. D., 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.

Kyte, J. A. D., R (1982). "A simple method for displaying the hydropathic character of a protein." J. Mol. Biol, 157: 105-132.

L, R. (2000). "Basic principles of quantitative PCR." Mol Biotechnol. 15: 115-122.

LA, D. (2001). "Lignification and lignin topochemistry-an ultra-structural view." Phytochemistry 57: 859-873.

Lacombe E, H. S., Doorsselaere JV, Piquemal J, Goffner D, Poeydomenge O, Boudet, AM: (1997). "Cinnamoyl CoA reductase, the first committed enzyme of the

lignin branch biosynthetic pathway: Cloning, expression and phylogenetic relationships." Plant Journal. 11: 429-441.

Le Gall O, T. L., Danglot Y, Candresse T, Bouquet A (1994). "Agrobacterium- mediated genetic transformation of grapevine somatic embryos and regeneration of transgenic plants expressing the coat protein of grapevine chrome mosaic nepovirus (GCMV)." Plant Science 102: 161-170.

Lee D, M. K., Chapple C, Douglas CJ: (1997). "Antisense suppression of 4-coumarate: coenzyme A ligase activity in Arabidopsis leads to altered lignin subunit composition." Plant Cell. 9: 985- 998.

Lee LG, C. C., Bloch W (1993). "Allelic discrimination by nick-translation PCR with fluorogenic probes." Nucleic Acids Res. 21: 3761-3766.

Lewis NG, Y. E. (1990). "Lignin: Occurrence, biogenesis, and biodegradation." Annu Rev Plant Physiol Plant Mol Biol, 41: 455-496.

Lewis NG, Y. E. (1990). "Lignin: occurrence, biogenesis and biodegradation." Annual Review of Plant Physiology and Plant Molecular Biology 41: 455-496.

Li L, C. X., Leshkevich J, Umezawa T, Harding SA, Chiang VL (2001). "The last step of syringyl monolignol biosynthesis in angiosperms is regulated by a novel gene encoding sinapyl alcohol dehydrogenase." Plant Cell 13: 1567-1585.

Li L, C. X., Lu S, Nakatsubo T, Umezawa T, Chiang VL: (2005). "Clarification of cinnamoyl co-enzyme A reductase catalysis in monolignol biosynthesis of aspen." Plant Cell Physiol. 46: 1073-1082.

Li L, O. Y., Joshi CP, Chiang VL (1999). "Secondary xylem specific expression of caffeoyl-coenzyme A 3-O-methyltransferase plays an important role in the methylation pathway associated with lignin biosynthesis in loblolly pine." Plant Mol Biol. 40: 555-565.

Li L, P. J., Zhang X.-H, Osakabe K, Tsai C-J, Joshi C P, Chiang VL (1997). "A novel multifunctional O-methyltransferase implicated in a dual methylation pathway associated with lignin biosynthesis in loblolly pine." PNAS. 94: 5461- 5466.

Li L, P. J., Umezawa T, Chiang VL (2000). "5-Hydroxyconiferylaldehyde modulates enzymatic methylation for syringyl monolignol formation, a new view of monolignol biosynthesis in angiosperms." J Biol Chem 275: 6537-6545.

Li L, Z. Y., Cheng XF, Sun J, Marita J, Ralph J, Chiang VL (2003). "Combinatorial genetic reduction of lignin quantity and augmentation of lignin syringyl constituents in trees through multigene co-transformation." Proc Natl Acad Sci USA 100: 4939-4944.

Livak KJ, F. S., 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 MA, G.-N. Y., Norman FW, Bruce IR (1994). "A simple and efficient method for DNA extraction from grapevine cultivars, Vitis species and Ampelopsis." Plant Mol Biol Reporter. 12: 6-13.

MacKay JJ, O. M. D., Presnell T, Booker FL, Campbell MM, Whetten RW, Sederoff, RR: (1997). "Inheritance, gene expression, and lignin characterization in a mutant pine deficient in cinnamyl alcohol dehydrogenase." Proc. Natl. Acad. Sci. 94: 8255-8260.

Maher EA, B. N., Ni W, Elkind Y, Dixon RA, Lamb CJ (USA). (1994). "Increased disease susceptibility of transgenic tobacco plants with suppressed levels of preformed phenylpropanoid products." Proc. Natl. Acad. Sci. 91: 7802-7806.

Marchler-Bauer A, P. A., Shoemaker BA, Thiessen PA, Geer LY, Bryant SH (2002). Nucleic Acids Res, 30: 281-283.

Marita JM, R. J., Hatfield RD, Chapple C: (1999,). "NMR characterization of lignins in Arabidopsis altered in the activity of ferulate 5-hydroxylase." Proc. Natl. Acad. Sci. (USA). 96: 12328-12332.

Martz F, M. S., Pincon G, Legrand M: (1998). "cDNA cloning, substrate specificity and expression study of tobacco caffeoyl-CoA 3-O-methyltransferase, a lignin biosynthetic enzyme." Plant Mol. Biol. 36: 427-437.

Mellerowicz EJ, B. M., Sundberg B, Boerjan W: (2008). "Unraveling cell wall formation in the woody dicot stem." Plant Mol. Biol. 47: 239- 247.

Mendelman LV, B. M., Petruska J,Goodman MF (1989). "Nearest neighbor influences on DNA polymerase insertion fidelity." J Biol Chem. 264: 14415-14423.

Mhlanga MM, M. L. (2001). "Using molecular beacons to detect single-nucleotide polymorphisms with real-time PCR." Methods. 25: 463-471.

Mierendorf, R. C., Percy, C. and Young, R.A. (1987). Meth.Enzymology.: 152, 458.

Moffat B.A, S. F. W. (1986). "Use of bacteriophage T-7 RNA polymerase to direct selective high level expression of cloned genes [abstract]." In Molecular Biology 189((1)): 113-130.

M.Q. Z. (1997). Proc Natl Acad Sci, 94: 565-568.

MR, G. O., Coon MJ, Estabrook RW, et al (1996). " P450 superfamily: update on new sequences, gene mapping, accession numbers and nomenclature." Pharmacogenetics 6: 1-42.

Mullis KB, F. F. (1987). "Specific synthesis of DNA in- vitro via a polymerase catalysed reaction." Methods Enzymol. 255: 335-350.

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

MW. P. (2001). “A new mathematical model for relative quantification real-time RT-PCR." Nucleic acid research 29(9): 2002-2007.

Nelson DR, K. L., Kamataki T, Stegeman JJ, Feyereisen R, Waxman DJ, Waterman MR, Gotoh O, Coon MJ, Estabrook RW: (1996). "P450 superfamily: update on new sequences, gene mapping, accession numbers and nomenclature." Pharmacogenetics 6: 1-42.

Ni W, P. N., Dixon RA (1994) Transgenic Res.1994, 3: 120-126.

Nimz HH, Tutschek R: (1977). "Reduced lignin in transgenic plants containing a caffeic acid O -methyltransferase antisense gene." Holzforschung. 31: 101-106.

O, G. (1992). "Substrate recognition sites in cytochrome P450 family 2 (CYP2) protein inferred from comparative analyses of amino acid and coding nucleotide sequences." J Biol Chem 267: 83-90.

O, G. (1998). "Divergent structures of Caenorhabditis elegans cytochrome P450 genes suggest the frequent loss and gain of introns during the evolution of nematodes." Mol Biol Evol 15: 1447-1459.

O'Connell A, H. K., Piquemal J, Grima-Pettenati J, Boudet A, Pollet B, Lapierre C, Petit-Conil M, Schuch W, Halpin C: (2002). “Improved paper pulp from plants with suppressed cinnamoyl-CoA reductase or cinnamyl alcohol dehydrogenase." Transgenic Res. 11: 495-503.

Okayama H, B. P. (1982). "High-efficiency cloning of full-length cDNA." Mol Cell Biol. 2:: 161- 170.

Olof Emanuelsson, S. B., Gunnar von Heijne, Henrik Nielsen (2007). Nature Protocols, 2: 953-971.

Omura T, S. R. (1964). "The carbon monoxide-binding pigment of liver microsomes. I. Evidence for its hemoprotein nature." J Biol Chem , 239: 2370-2378.

Osakabe K, T. C., Li L, Popko JL, Umezawa T, Carraway DT, Smeltzer RH, Joshi CP, Chiang VL: (1999). "Coniferyl aldehyde 5-hydroxylation and methylation direct syringyl lignin biosynthesis in angiosperms." Proc. Natl. Acad. Sci. (USA). 96: 8955-8960.

P, H. (1995). "Purification of His-Tag fusion proteins from Escherichia coli." Trends Biochem Sci. 20: 285-286.

P, H. (1995). "Purification of His-Tag fusion proteins from Escherichia coli." Trends Biochem Sci. 20: 285-286.

P, M. (2001). "Microfluidics- downsizing large-scale biology." Nat. Biotechnol, 19: 717-721.

Petsko GA, S. S. (2000). "The catalytic pathway of cytochrome P450cam at atomic resolution." Science 287: 1615-1622.

Pincon G, C. M., Lapierre C, Pollet B, Ruel K, Joseleau J-P, Boudet AM, Legrand M: (2001). "Simultaneous down-regulation of caffeic/5-hydroxy ferulic acid-Omethyltransferase I and cinnamoyl-coenzyme A reductase in the progeny from a cross between tobacco lines homozygous for each transgene. Consequences for plant development and lignin synthesis." Plant Physiol. 126: 145-155.

Piquemal J, L. C., Myton K, O'Connell A, Schuch W, Grima-Pettenati J, Boudet A-M: (1998). "Down-regulation in cinnamoyl-CoA reductase induces significant changes of lignin profiles in transgenic tobacco plants." Plant Journal. 13: 71-80.

Ralph J, H. R., Piquemal J, Yahiaoui N, Pean M, Lapierre C, Boudet AM: (1998). "NMR characterization of altered lignins extracted from tobacco plants down-regulated for lignification enzymes cinnamyl alcohol dehydrogenase and cinnamoyl CoA reductase." Proc. Natl. Acad. Sci. (USA). 95: 12803-12808.

Ralph J, L. K., Brunow G, Lu F, Kim H, Schatz PF, Marita JM, Hatfield RD, Christensen JH, Boerjan W: (2004). "Natural polymers from oxidative coupling of 4-hydroxyphenylpropanoids." Lignins: Phytochemistry Review. 3: 29-60.

Reid ID, P. M. (1994). "Biological bleaching of kraft pulps by white-rot fungi and their enzymes." FEMS Microbiol. Review. 13: 369-376.

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

Richard Sibout, M. B., Melinda Gatineau, Jan Van Doorsselaere, Isabelle Mila, Brigitte Pollet, Bruno Maba, Gilles Pilate, Catherine Lapierre, Wout Boerjan, Lise Jouanin. (2002). "Expression of a poplar cDNA encoding a ferulate 5-hydroxylase/coniferaldehyde 5-hydroxylase increases S lignin deposition in Arabidopsis thaliana." Plant Physiol. Biochem. 40: 1087-1096.

Richardson, T. H., F. Jung, K. J. Griffin, M.Wester, J. L. Raucy, B. Kemper, L. M. Bornheim, C. Hassett, C. J. Omiecinski, and E. F. Johnson. (1995). "A universal approach to the expression of human and rabbit cytochrome P450s of the 2C subfamily in Escherichia coli." Arch. Biochem.Biophys. 323: 87-96.

Rochus Franke, C. M. M., Knut Meyer, Amber M. Shirley, Joanne C. Cusumano and Clint Chapple (2001). "Modified lignin in tobacco and poplar plants over-expressing the Arabidopsis gene encoding ferulate 5-hydroxylase." The Plant Journal 22(3): 223 - 234.

Ruegger M., K. M., J.Cussumano, C. Chapple. (1999). "Regulation of ferulate 5-hydroxylase expression in Arabidopsis in the context of sinapate ester biosynthesis." Plant Physiol. 119: 101-110.

Ruiqin Zhong, W. H. M. I., David S. Himmelsbach, Farris L. Poole II, and Zheng-Hua Ye (2000). "Essential Role of Caffeoyl Coenzyme A O-Methyltransferase in Lignin Biosynthesis in Woody Poplar Plants." Plant Physiol, 124: 563-578.

S Koutaniemi, T. W., A arkonen, E Alatalo, CG Fossdal, P aa, T Laakso, KV Fagerstedt, LK Simola, L Paulin, S Rudd, TH Teeri (2007). "Expression profiling of the lignin biosynthetic pathway in Norway spruce using EST sequence." Plant Mol Biol Reporter. 65: 311-28.

SA, B. (1969). "Biochemistry of lignin formation." Bioscience 19: 115-121.

Saha BK, T. B., Bucy RP (2001). "Quantitation of HIV-1 by real-time PCR with a unique fluorogenic probe." J Virol Methods. 93: 33-42.

Saiki R, S. S., Faloona F, Mullis KB, Horn GT, Erlich HA, Arnheim N (1985). "Enzymatic amplification of beta-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia." Science. 230: 1350-1354.

Saiki RK, G. D., Stoffel S, Scharf S, Higuchi R, Horn GT, Mullis KB, Erlich HA (1988). "Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase." Science. 239: 487-491.

Saka S, G. D. e. H. T. (1985). "Localization of lignins in wood cell walls." In Biosynthesis and Biodegradation of Wood Components: 51-62.

Sambrook, E. F. F.A. T. M. (1989). molecular cloning: A Laboratory manual (second Edition), cold spring Harbour Laboratory, Cold Spring harbor 8.3-11.58.

Sambrook J, F. E., Maniatis T 1989 Molecular cloning; A Laboratory Manual, 2nd ed., New York: Cold Spring Harbor Laboratory Press, Cold Spring Harbor. Scharf SJ, Horn GT, Erlich HA (1986). "Direct cloning and sequence analysis of enzymatically amplified genomic sequences." Science. 233: 1076-1078.

Sarkanen KV, L. C. (1971). "Lignins: Occurrence, Formation, Structure, and Reactions." New York: Wiley-Intersci : 916.

Sarni F, G. G., Boudet AM: (1984). "Purification and properties of cinnamoyl-CoA reductase and cinnamyl alcohol dehydrogenase from poplar." Eur.J. Biochem. 139: 259-265.

Sellner LN, C. R., MacKenzie JS (1992). "Reverse transcriptase inhibits Taq polymerase activity." Nucleic Acids Res. 20: 1487-1490.

Sewalt V.J.H, N. W., Blount JW, Jung HG, Masoud SA, Howles PA, Lamb C, Dixon RA: (1997). "Reduced lignin content and altered lignin composition in transgenic tobacco down-regulated in expression of L-phenylalanine ammonia-lyase or cinnamate 4-hydroxylase." Plant Physiology 115: 41- 50.

Sibout R, E. A., Pollet B, Goujon T, Mila I, Granier F, Seguin A, Lapierre C, Jouanin L: (2003). "Expression pattern of two paralogs encoding cinnamyl alcohol dehydrogenase in Arabidopsis. Isolation and characterization of the corresponding mutants." Plant Physiol. 132: 848-860.

Siebert PD, C. A., Kellogg DE, Lukyanov KA, Lukyanov SA (1995). "An improved method for walking in uncloned genomic DNA." Nucleic Acids Res. 23: 1087-1088.

Solinas A, B. L., McKeen C, Mellor JM, Nicol J, Thelwell N, Brown T (2001). "Duplex Scorpion primers in SNP analysis and FRET applications." Nucleic Acids Res. 29: E: 96.

Stafford H.A, D. S. (1972). "4-Hydroxycinnamic acid hydroxylase and polyphenolase activities in Sorghum vulgare." Plant Physiol. 49: 590-595.

T.F., U. (1977). "Show me the Money: Prokaryotic expression vectors and Purification system." The Scientist. 11: (17).

T, H. (1997). "Biochemistry and Molecular Biology of Wood." (Springer Series in Wood Science,) Springer, Berlin.

Tan W, W. K., Drake TJ (2004). "Molecular beacons." Curr Opin Chem Biol. 8: 547-553.

TE, T. (1986). "Determination, Structure, Chemistry, Topochemistry, Physical Properties, Origin, and Formation of Compression Wood." Compression Wood in Gymnosperms, I.(Berlin: Springer-Verlag.): 706.

Terashima N, F. K., He LF, Takabe K (1993). "Comprehensive model of the lignified plant cell wall." In Forage Cell Wall Structure and Digestibility, 1993. ed. Jung HG, Buxton DR, Hatfield RD, Ralph D: 247-70.

Terry CF, S. D., Ballam LD, Harris N, McDowell DG, Parkes HC (2002). "Real-time detection of genetically modified soya using Lightcycler and ABI 7700 platforms with TaqMan, Scorpion, and SYBR Green I chemistries." JAOAC Int. 85: 938-944.

Timothy J. Tranbarger, B. S. f. a. S. M. (2000). "Regulation of NADPH-cytochrome P450 reductase expressed during Douglas -fir germination and seedling development." Plant Molecular Biology 44: 141-153.

Todd JA, B. J., McDevitt HO., (1987) " HLA-DQ β-gene contributes to susceptibility and resistance to insulin-dependent diabetes mellitus." Nature. 329: 599-604.

UK, L. (1970). "Cleavage of structural proteins during the assembly of the head of bacteriophage T4." Nature. 227: 680-685.

Van der Velden VH, H. A., Cazzaniga G, Szczepanski T, Gabert J, van Dongen JJ (2003). "Detection of minimal residual disease in hematologic malignancies by real-time quantitative PCR." Leukemia. 17: 1013-1034.

Van Doorsselaere J, B. M., Chognot E, Chabbert B, Tollier M-T, Petit-Conil M, Leple J-C, Pilate G, Cornu D, Monties B, (1996). Plant J 8: 855-864.

Vet JA, M. S. (2005). "Design and optimization of molecular beacon real-time polymerase chain reaction assays." Methods Mol Biol. 288: 273-290.

Vet JA, V. d. R. B., Blom HJ (2002). "Molecular beacons: colorful analysis of nucleic acids." Expert Rev Mol Diagn. 2: 77-86.

Ward., H. A. H. A. J. M. (2003). "Enhanced heterologous expression of two Streptomyces griseolus Cytochrome P450s and Streptomyces coelicolor Ferredoxin Reductase as potentially efficient hydroxylation catalysts." Applied And Environmental Microbiology 69,(1): .73-382.

Whetten RW, M. J., Sederoff RR: (1998). "Recent advances in understanding lignin biosynthesis." Annu Review Plant Physiology Plant Mol. Biol. 49: 585-609.

Whetten RW, M. J. a. S. R. (1998). "Recent advances in understanding lignin biosynthesis." Annual Review Plant Physiology and Plant Molecular Biology 49: 585-609.

Williams PA, C. J., Sridhar V, Johnson E, McRee DE: (2000). "Mammalian microsomal cytochrome P450 monooxygenase: structural adaptations for membrane binding and functional diversity." Molecular Cell. 5: 121-131.

Wittwer CT, H. M., Moss AA, Rasmussen RP (1997). "Continuous fluorescence monitoring of rapid cycle DNA amplification." Biotechniques. 22: 130-131, 134-138.

Wittwer CT, R. K., Andrew RV, David DA, Gundry RA, Balis UJ (1997). "The LightCycler: a microvolume multisample fluorimeter with rapid temperature control." Biotechniques. 22: 176-81.

Yanofsky M, L. B., Montoya A, Rubin R, Krul W, Gordon M, Nester E (1985). "Molecular and genetic analysis of factors controlling host range in Agrobacterium tumefaciens." Mol Gen Genet 201: 237-246.

Ye Z-H, K. R., Matern U, Varner JE. (1994). "An alternative methylation pathway in lignin biosynthesis in Zinnia." Plant Cell. 6: 1427-1439.

Zhong R, M. W. I., Negrel J, Ye Z-H: (1998). " Dual methylation pathways in lignin biosynthesis." Plant Cell. 10: 2033-2046.

Zhong R, M. W. I., Negrel J, Ye Z-H: (1998). "Dual methylation pathways in lignin biosynthesis." Plant Cell. 10: 2033-2046.

Zhu YY, M. E., Chenchik A, Li R & Siebert PM (2001). "Reverse transcriptase template switching: A SMART™ approach for full-length cDNA library construction." Biotechniques. 30: 892-897.

Zhu YY, M. E., Chenchik A, Li R & Siebert PM: (2001). Biotechniques, 30: 892-897.

Zimmermann HH, N., E. Seemuller: (1985). Holzforschung. 39: 45-49.

Publications

Research papers published/under review/under preparation:

1. Noor M. Shaik, Manish Arha, Nookaraju A, Sushim. K. Gupta, Sameer Srivastava, **Arun K. Yadav**, Pallavi S. Kulkarni, Abhilash O. U., Rishi K. Vishwakarma, Somesh Singh, Rajeshri Tatkare, Kannan Chinnathambi, Shuban K. Rawal, Bashir M. Khan (October, 2009). *Improved method of in vitro regeneration in Leucaena leucocephala - a leguminous pulpwood tree species*. Physiology and molecular biology of plants, 15(4).

2. 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 388-3936.

3. **A. K. Yadav**, M. Arha, S. K. Gupta, N. M. Shaik, Sameer Srivastava P. Kulkarni, O. U. Abhilash and B. M. Khan (2008). *Molecular cloning, characterization and expression analysis of Coniferaldehyde 5-hydroxylase gene in developing seedlings of Leucaena leucocephala, a paper and pulp yielding tree species* (Manuscript under preparation).

4. **A. K. Yadav**, M. Arha, S. K. Gupta, N. M. Shaik, Sameer Srivastava P. Kulkarni, O. U. Abhilash and B. M. Khan (2009).*Homology modelling and docking studies of Coniferaldehyde 5-hydroxylase from Leucaena leucocephala, a paper and pulp yielding tree species.* (Manuscript under preparation).

5. 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).

Abstract Published

1. 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.

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 (2005). *High frequency regeneration and transformation of Leucaena leucocephala*. In Proceedings of National Symposium on Plant Biotechnology: New Frontiers, CIMAP, Lucknow, Uttar Pradesh, India

3. **A. K. Yadav**, M. Arha, S. K. Gupta, N. M. Shaik, S. Srivastava, 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.

4. **A. K. Yadav**, M. Arha, S. K. Gupta, N. M. Shaik, S. Srivastava, 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.

5. **A. K. Yadav**, M. Arha, S. K. Gupta, N. M. Shaik, S. Srivastava, 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.

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.

Appendix

***leucoplast Leucaena leucocephala* [gbpln]: 1 CDS's (327 codons)**

Fields: [triplet] [frequency: **per thousand**] ([number])

UUU	0.0 (0)	UCU	6.1 (2)	UAU	12.2 (4)	UGU	15.3 (5)
UUC	42.8 (14)	UCC	30.6 (10)	UAC	36.7 (12)	UGC	39.8 (13)
UUA	3.1 (1)	UCA	3.1 (1)	UAA	3.1 (1)	UGA	0.0 (0)
UUG	9.2 (3)	UCG	9.2 (3)	UAG	0.0 (0)	UGG	21.4 (7)
CUU	3.1 (1)	CCU	6.1 (2)	CAU	0.0 (0)	CGU	3.1 (1)
CUC	27.5 (9)	CCC	30.6 (10)	CAC	9.2 (3)	CGC	15.3 (5)
CUA	3.1 (1)	CCA	27.5 (9)	CAA	21.4 (7)	CGA	0.0 (0)
CUG	21.4 (7)	CCG	6.1 (2)	CAG	27.5 (9)	CGG	6.1 (2)
AUU	9.2 (3)	ACU	6.1 (2)	AAU	3.1 (1)	AGU	3.1 (1)
AUC	18.3 (6)	ACC	21.4 (7)	AAC	36.7 (12)	AGC	27.5 (9)
AUA	6.1 (2)	ACA	6.1 (2)	AAA	3.1 (1)	AGA	18.3 (6)
AUG	21.4 (7)	ACG	18.3 (6)	AAG	24.5 (8)	AGG	12.2 (4)
GUU	3.1 (1)	GCU	21.4 (7)	GAU	18.3 (6)	GGU	27.5 (9)
GUC	12.2 (4)	GCC	36.7 (12)	GAC	36.7 (12)	GGC	58.1 (19)
GUA	0.0 (0)	GCA	9.2 (3)	GAA	3.1 (1)	GGA	27.5 (9)
GUG	15.3 (5)	GCG	21.4 (7)	GAG	12.2 (4)	GGG	21.4 (7)

Coding GC 60.55% 1st letter GC 53.21% 2nd letter GC 55.66% 3rd letter GC 72.78%

Table of Standard Genetic Code

	T	C	A	G
T	TTT Phe (F) TTC " TTA Leu (L) TTG "	TCT Ser (S) TCC " TCA " TCG "	TAT Tyr (Y) TAC TAA **Ter** TAG **Ter**	TGT Cys (C) TGC TGA **Ter** TGG Trp (W)
C	CTT Leu (L) CTC " CTA " CTG "	CCT Pro (P) CCC " CCA " CCG "	CAT His (H) CAC " CAA Gln (Q) CAG "	CGT Arg (R) CGC " CGA " CGG "
A	ATT Ile (I) ATC " ATA " **ATG** Met (M)	ACT Thr (T) ACC " ACA " ACG "	AAT Asn (N) AAC " AAA Lys (K) AAG "	AGT Ser (S) AGC " AGA Arg (R) AGG "
G	GTT Val (V) GTC " GTA " GTG "	GCT Ala (A) GCC " GCA " GCG "	GAT Asp (D) GAC " GAA Glu (E) GAG "	GGT Gly (G) GGC " GGA " GGG "