"Molecular studies on Glycosyltransferase gene(s) from *Bacopa monniera*"

A THESIS
SUBMITTED TO THE
UNIVERSITY OF PUNE

FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY
IN
BIOTECHNOLOGY

BY
RUBY

UNDER THE GUIDANCE OF
DR. B. M. KHAN

PLANT TISSUE CULTURE DIVISION
NATIONAL CHEMICAL LABORATORY
PUNE- 411008
INDIA
FEBRUARY, 2011

Dedicated to my beloved
Parents and Grand mother...

CERTIFICATE

This is to certify that the work incorporated in the thesis entitled **"Molecular studies on Glycosyltransferase gene(s) from *Bacopa monniera*"** submitted by Ruby for the degree of Doctor of Philosophy, 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 thesis entitled **"Molecular studies on Glycosyltransferase gene(s) from *Bacopa monniera*"** has been carried out at Plant Tissue culture Division, National Chemical Laboratory, Pune, under the guidance of **Dr. Bashir M. Khan**. The work is original and has not been submitted in part or full by me for any other degree or diploma to any other university. I further declare that the materials obtained from other sources have been duly acknowledged in the thesis.

(Ruby)

Date: February, 2011

Place: Plant Tissue Culture Division,
National Chemical Laboratory (NCL),
Pune – 411008, India.

ACKNOWLEDGEMENTS

First of all I would like to express my deep sense of gratitude towards my research guide, ***Dr. B. M. Khan*** *for his subtle guidance, unmatched humanity, constant support and encouragement, constructive criticism, valuable suggestions and for giving me freedom in finalizing the research and preparation of this manuscript without which the task would have been difficult. His enthusiastic attitude, innovative ideas and scientific knowledge have inspired me profoundly. I truly feel privileged to have joined his research group. I sincerely thank for the care and affection that I received from him and his family in the entire period.*

I would be equally indebted to ***Dr. S. K. Rawal****, Former Head, PTC Division, NCL, for his moral support, valuable comments and impeccable suggestions during the progress and completion of this work.*

Special and sincere thanks to ***Dr. B. D. Kulkarni.***

I am really greatful to ***Mrs. M. V. Shirgurkar*** *for her help in tissue culture work,* ***Dr. C. G. Suresh*** *for his help and analysis in homology modeling and docking studies,* ***Dr. Diman Sarkar*** *for his guidance in cell culture work and* ***Dr. Srinivas Hotha*** *for providing me the LC-MS facility.*

My sincere thanks to Ranu Sharma for her help in homology modeling and docking studies, Shampa Sarkar for her help in cell culture studies, Asif shiekh for his help in conducting LC-MS experiments and Yashwant for his help in MALDI analysis.

I would like to extend my thanks to Dr. Sucheta, Dr. D.C. Agarwal, Dr. Ashok Giri, Mrs. S. Kendurkar and all other members of the scientific and supporting staff of the Plant Tissue Culture Division, Biochemical Sciences Division and NCL Library.

Many thank to my seniors Dr. Rohini, Dr. Neelima, Dr. Nookaraju, Dr. Manish, Dr. Sushim, Dr. Noor, Dr. Sameer, Dr. Arun, Dr. Pallavi and Dr. Abhilash, for their encouragement and help during the course of this work.

I would like to thank my labmates and colleagues R.J.Santosh Kumar, Rishi (rishi baba), Somesh (somu), Prashant (sonya), Poonam (pooo), Neha (Sweetoo), Parth (motu), Krunal (kurkur), Sumita (sumi), Santosh Gupta, Kannan (kanzz), Shakeel (shaks), Uma (umi), Trupti, Sonali, Nazneen, Zohaib, Sneha, Gayatri, Pratik, Amar, Anuja, Sobia, Shruti, Disha, Richa, Srikanth, Sudarshan, Sameena, Yasir, Nishi, Rajshree, Gayatri I.T, Lumbini, Jay, Azfar, Malini, Ulfat for maintaining lively and cordial atmosphere in the lab.

I thank my trainees Sulakhana, Asma S, Aditi, preeti and rashmi for their help during my work period.
It gives me great pleasure to thank my friends Fazal (Fazlu), Sameera, Poonam L, Neha K, Meena, Swapna, Trupti J, Varsha, Raju D, Bhuban, Sweta S, Nitasha, Rita, Asad, Shadab, Srikanth, Ravi, Viku, Ashish, Chetan, Pushkar, Ejaz, Pooja, Jyoti, Tonima, Saurabh, Ashwini. R,

*Rasika, Varsha, Madhurima, Ansari, Avinash, Geeta, Sridevi, Mayank, Amrish, Nishant, Tulika, Urvashi, Manas, Priyabrath, Ketaki, **Kashmir University group** (Qulsum, Nissar W, Tousief, Nissar S, Javaid, Eijaz, Asiya, Rukhsana, Riyaz, Muneera, Gowher W and all other batchmates), **School group** (Sheema, Iqra, Shaheen, Saima, Asma and Ifat) and **TLEP 2008 group** for their great support.*

*Special mention is a must for **GJ group** Manisha (nishu), Shampa (shampu), Manasi (mantu), Manaswini (manu), Payal (PG), Mun Mun (mooni), Rupa (Roops), Shreeja (Chechi), Savita (roomy), Priyanka, Gawri, Mandakini, Lalitha, Sangeeta, Tanaya, Jaya, Anjali, Shivranjini, Jhumur and all other hostelites. I would always cherish the unlimited fun I shared with them.*

*I take this opportunity to express my earnest respect to my school teachers, graduation tutors especially **Dr. Khursheed I. Andrabi** and **Dr. Bashir A. Mir** for their way of teaching that build up my research career in science.*

*I am highly indebted to my friend **Sweta Shipra** for constant support and making my stay more comfortable by her presence. I also cherish all those moments I shared with her.*

*Words would not suffice to thank my best buddy **Rizwan Habib Wani** (Rizu) for his never ending support and being on my side in any instance. I will treasure the laughs, fights and tears I have shared with him till my last breath.*

Special thanks to Prof. Habibullah wani, Mrs. Zarifa Banu, Samina H, Shamim H, Imran H and Muneer R for their unwavering support and dearest little Ayaan for his lisping which would relieve my tension at times.

*I owe my deepest love and gratitude to my parents (**Mr. Mohd Maqbool Zargar** and **Mrs. Akther**), Grandmother, my uncles (Bashir A. Zargar and Dr. Ashraf), my sisters (Wajida, Wahida, Nasima and Zuryat), my brothers (Mohsin and Sakib) and my aunt Halima for their unconditional love, support and encouragement that has helped me reach this far.*

I am grateful to Dr. Saurav Pal, Director, NCL for providing necessary facilities and permitting me to submit my findings in the form a thesis.

The fellowship awarded by the CSIR, India is duly acknowledged.

*Last but not the least I thank **almighty God** for giving me the strength and courage at every step of life.*

Ruby

CONTENTS

Abbreviations

AA	Amino acid
ADHD	Attention deficit hyperactivity disorder
AraT	Arabinosyltransferase
ABA	Abscisic acid
bp	Base pairs
BM	*Bacopa monniera*
BMGT	*Bacopa monniera* glycosyltransferase
BR	Brassinosteroids
BSA	Bovine serum albumin
cDNA	Complementary DNA
Ci/ mmol	Curie per milli mole
CIAP	Calf Intestinal Alkaline Phosphatase
Cps	Counts per second
CAT	Catalase
Da	Dalton
DEPC	Diethylpyrocarbonate
DNA	Deoxyribose nucleic acid
DTT	Dithiothritol
EDTA	Ethylene Diamine Tetra Acetic acid
EDTA	Ethylene diamine tetra acetic acid disodium salt
ELISA	Enzyme linked immuno sorbent assay
EtBr	Ethidium bromide
GT	Glycosyltransferase
g /L	Grams per litre
g	Gram
G	Guaiacyl
gDNA	GenomicDNA
GSP	Gene Specific Primers

GSH-Px	Glutathione peroxide
GABA	Gamma amino butyric acid
GAT	Glucuronic acid transferases
Glc T	Glycosyltransferases
Gat T	Galactosyltransferases
h	Hour(s)
HSP 70	Heat-shock protein (Hsp70)
IPTG	Isopropyl ß-D-thiogalactoside
Kb	Kilobase pairs
KDa	Kilo Daltons
Kg	Kilogram
L	Litre
LB	Luria-Bertani
LD_{50}	Lethal dose 50%
MS	Murashige and Skoog
MCS	Multiple cloning sites
mg	Milligram
min	Minute(s)
mL	Millilitre
µL	Micro liter
µg	Microgram
mM	Millimolar
mRNA	Messenger RNA
nM	Nano molar
nm	Nanometer
NUP	Nested Universal Primers
O/N	Overnight
OD	Optical density
pI	Isoelectric point
PEG	Polyethylene glycol
pg	Picogram

pmol	Picomole
PMSF	Phenyl methyl sulphonyl fluoride
ppm	Parts per million
psi	pounds per square inch
PVPP	Poly vinyl pyro phosphate
pmol	Pico mole
PSPG	Plant secondary product glycosyltransferase
Q-PCR	Quantitative polymerase chain reaction
RACE	Rapid amplification of cDNA ends
RNase A/ (H)	Ribonuclease A/(H)
RNA	Ribose nucleic acid
rpm	Rotations per minute
RT	Room temperature
RhaT	Rhamnosyltransferases
RT-PCR	Reverse transcriptase polymerase chain reaction
RT-PCR	Real-time polymerase chain reaction
s	Second(s)
SDS	Sodium dodecyl sulphate (sodium lauryl sulphate)
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
SMQ	Sterile Milli Q
Soln	Solution
sp.	Species
SSC	Saline sodium citrate
SOD	Superoxide dismutase
SA	Salicylic acid
TAE	Tris acetic EDTA buffer
TE	Tris EDTA buffer
TEMED	Tetramethylethylenediamine
TLC	Thin layer chromatography
U	Units
UGT	UDP-glucose glycosyltransferases

UDP-GT	UDP-glycosyltransferase
UPM	Universal Primer Mix
UTR	Untranslated Region
UV	Ultraviolet
V	Volt
v/v	Volume / Volume
WHO	World Health Organization
w/v	Weight / Volume
X-gal	5-bromo-4-chloro-3-indolylß-D-galactoside
α	Alpha
β	Beta
λ	Lamda
%	Percentage
°C	Degree Celsius
µg	Microgram
µg/L	Micrograms per liter
µL	Microlitre
µm	Micrometer
µM	Micromolar

Abstract

Bacopa monniera (Family: Scrophulariaceae) is a medicinal herb, found throughout the Indian subcontinent in wet, damp and marshy areas. It is often called as “Brahmi” or “The thinking person’s herb”. It is used in traditional Indian medicine, the Ayurveda, for the treatment of anxiety, and in improving intellect and memory for several centuries. In addition to memory boosting activity, it is also claimed to be useful in the treatment of cardiac, respiratory and neuropharmacological disorders like insomnia, insanity, depression, psychosis, epilepsy and stress. It was reported to possess anti-inflammatory, analgesic, antipyretic, sedative, free radical scavenging and anti-lipid peroxidative activities also. The active constituents of *Bacopa monniera* are saponins, alkaloids, and phytosterols. The pharmacological properties of *Bacopa monniera* were studied extensively and the activities were attributed mainly due to the presence of characteristic saponins called as “bacosides.” Bacosides are complex mixture of structurally closely related compounds, glycosides of either jujubogenin or pseudojujubogenin.

Plant-derived glycosides have attracted much attention due to their widespread applications. The isolation and purification of such glycosides from plant sources is, however, tedious and usually results in low yields. Thus, simple approaches for the generation of such glycosides would be highly beneficial. Since glycosylation confers amphipathic properties on the molecule and is normally critical for biological activity, this is clearly an important area in which an effort should be invested.

Glycosyltransferases (GTs):

Glycosylation reactions are of great biological importance to both prokaryotes and eukaryotes, and require the coordinated action of a large number of enzymes, the glycosyltransferases (GTs). These enzymes transfer the sugar moiety from an activated nucleotide–sugar to an acceptor, which may be a growing oligosaccharide, a lipid or a protein. The enzymes can be classified into families on the basis of sequence similarity, catalytic specificity and the existence of consensus sequences. There are over 91 glycosyltransferase families. Despite the fact that many GTs recognize identical donor or acceptor substrates, few regions of sequence homology have been

found among the different classes of eukaryotic glycosyltransferases and enzymes that are structurally related most often catalyze the same or a similar reaction. It seems these glycosyltransferases are "managers of metabolism", playing a role in cellular homeostasis and through their activity, regulating metabolic flux, levels of active hormones and the detoxification of xenobiotics.

Plants are exposed to a wide range of toxic and bioactive low-molecular-weight molecules from both exogenous and endogenous sources. Glycosylation is one of the primary sedative mechanisms that plants utilize in order to maintain metabolic homeostasis. Plants are capable of synthesizing several thousand different low molecular weight compounds, defined as secondary plant metabolites. Part of this diversity arises from decoration with glycosyl-, carboxyl-, methyl- and hydroxyl- groups by glycosyltransferases, acyltransferases, methyltransferases and cytochrome P450s, respectively. The conjugation of endogenous and exogenous organic molecules with sugar is one such important tool employed by all organisms. Secondary plant metabolites are glycosylated to O (OH- and COOH-), N, S and C atoms by glycosyltransferases (GTs) using nucleotide-activated sugars as donor substrates. Hydroxylated molecules are the most common acceptors, whilst UDP-glucose is the most common donor.

Glycosylation converts reactive and toxic aglycones into stable and non-reactive storage forms, thereby limiting their interaction with other cellular components. Generally speaking, the attachment of carbohydrate moieties to nucleophillic molecules will reduce the possibility of electron transfer from the unglycosylated molecule to other cellular components, thereby lowering the reactivity, and consequentially improving the stability of that molecule. Since the nucleophillic sites are, in many cases, the particular part of the molecules that interact damagingly with other cellular components, the addition of sugars will block the interactive site and consequently reduce toxicity. Sugars are highly polar, and the addition of carbohydrate moieties to hydrophobic substrates will serve to increase the water solubility of the resultant glycoside.

The present study was taken up with the objective to isolate, clone and characterize Glycosyltransferase gene (GT) from *Bacopa monniera*, which plays crucial role in solubilisation of medicinally important secondary metabolites. The other objective was to study the expression of protein in heterologous system and see the expression pattern in different parts of the plant. The thesis embodies five chapters.

Chapter 1. General Introduction

This chapter gives general information on medicinal plants in India. Background of research done on *Bacopa monniera*, a medicinal plant has been dealt in detail. A thorough literature survey of work done in the area of glycosyltransferase gene with regards to the current status of research in this area has been presented. A special emphasis on the techniques and strategies used in study of the glycosyltransferase gene has been dealt with. Role of the key enzyme in the formation of different secondary metabolites with special reference to health has been provided. Finally, the scope of the present study and objectives of the thesis work have been discussed.

Chapter 2. Materials and Methods

Overall materials and the protocols followed during the course of this work in detail have been included in this chapter.

Chapter 3. Isolation, Cloning and Characterization of *B. monniera* Glycosyltransferase gene(s)

This chapter deals with the PCR based approach for fishing out the full-length c-DNA clones of glycosyltransferase gene(s) from *Bacopa monniera*. Primers were designed on the basis of consensus regions of various reported nucleotide sequences of the GT genes from the NCBI GenBank database. A partial GT gene sequence was amplified and the sequence showed maximum homology with GT members of lamiaceae family. The partial GT gene sequence was used to design gene specific primers in order to get the full length genes. Rapid Amplification of cDNA ends (RACE) was performed to fish out the full length cDNA clones. Characterization of the gene(s) encoding GT is embodied in this chapter.

Chapter 4.

(4.A): Heterologous expression of *B. monniera* GT gene(s), its purification and characterization

This chapter deals with the cloning of cDNA of GT gene(s) in the expression vector pET 30b+ and its expression in *E. coli* BL21 (DE3). The protein was purified from inclusion bodies using Ni-chelated affinity column. Polyclonal antibodies have been raised against purified GT protein in rabbit which has been used for western analysis. This section also covers characterization of the protein.

With reference to the anti-cancerous properties of the aglycone moiety (Genistein), comparative study of aglycone part and glycosylated product has been done by cell proliferation MTT assay using MCF-7 and THP-1 cell lines.

(4.B): Tissue specific expression and immuno-cytolocalization of *B. monniera* GT gene(s):

This chapter will describe the expression pattern of GT gene(s) in different plant part. The spatial and temporal expression of the GT genes was studied using quantitative Real Time PCR. Primers were designed from the non conserved region of the gene so as to distinguish it from the other GT genes. The 18S RNA was used as the internal standard. Different plant parts *i.e*, leaf, root and stem were used for expression studies. This section also deals with the expression studies of samples under different stress conditions.

Chapter 5. Homology modeling and docking studies of the *B. monniera* GT gene(s)

This chapter deals with the molecular modeling studies of the GT gene(s) using MODELLER 9v8 software and for the docking studies AUTODOCK 4.0 and FLEX software was used. Relative affinity of various substrates towards the GT has been discussed in this chapter.

Chapter: 1

Introduction

1.1 Medicinal plants

The medicinal properties of plants have made an outstanding contribution in the origin and evolution of many traditional herbal therapies. These traditional knowledge systems have started to disappear with the passage of time due to scarcity of written documents and relatively low income in these traditions. Over the past few years, however, the medicinal plants have regained a wide recognition due to an escalating faith in herbal medicine in view of its lesser side effects compared to allopathic medicine and the necessity of meeting the requirements of medicine for an increasing human population.

India is endowed with a rich wealth of medicinal plants. In India, of the 17,000 species of higher plants, 7500 are known for medicinal uses (Shiva MP, 1996). This proportion of medicinal plants is the highest proportion of plants known for their medical purposes in any country of the world for the existing flora of that respective country (Table 1). Ayurveda, the oldest medical system in Indian sub-continent, has alone reported approximately 2000 medicinal plant species, followed by Siddha and Unani (Table 2). The Charak Samhita, an age-old written document on herbal therapy, reports on the production of 340 herbal drugs and their indigenous uses (Kumar *et al*., 2003). Most of these continue to be gathered from wild plants to meet the demand of the medical profession. Thus, despite the rich heritage of knowledge on the use of plant drugs, little attention had been paid to grow them as field crop in the country till the latter part of the nineteenth century.

Table 1: Distribution of medicinal plants.

Country/ region	**Total no. of Native species in Flora**	**No. of medicin-al plant species reported**	**% of medicinal plants**	**Source**
World	297000	52885	10	Schippmann *et al.,* 2002
India	17000	7500	44	Shiva 1996
Indian Himalayas	8000	1747	12	Samant *et al.,* 1998

Table 2: The status of various medicinal systems in India.

Characteristics	**Medical systems**				
	Ayurveda	**Siddha**	**Unani**	**Tibetan**	**Homeopathy**
Medicinal plants known	2000	1121	751	337	482
Licenced pharmacies	8533	384	462	-	613
Hospitals	753	276	74	-	223
Dispensaries	15193	444	1193	-	5634
Registered practitions	438721	17560	43578	-	217460
Undergraduate college	219	6	37	-	178
Postgraduate college	57	3	8	-	31

During the past seven or eight decades, there has been a rapid extension of the allopathic system of medical treatment in India. It generated a commercial demand for pharmacopoeial drugs and products in the country. Thus efforts were made to introduce many of these drug plants into Indian agriculture and studies on the cultivation practices were undertaken for those plants which were found suitable and remunerative for commerical cultivation. The average annual foreign trade in crude drugs and their phytochemicals is between 60 and 80 million rupees and this accounts for a little over 0.5 per cent of the world trade in these commodities. Currently, approximately 25% of drugs are derived from plants, and many others are synthetic analogues built on prototype compounds isolated from plant species in modern pharmacopoeia (Rao *et al*., 2004).

The northern part of India harbours a great diversity of medicinal plants because of the majestic Himalayan range. So far about 8000 species of angiosperms, 44 species of gymnosperms and 600 species of pteridophytes have been reported in the Indian Himalaya (Singh and Hajra, 1996), of these 1748 species are known as medicinal plants (Samant *et al*., 1998). The maximum medicinal plants (1717 species) have been reported around the 1800 m

elevation range. On the regional scale, the maximum species of medicinal plants have been reported from Uttaranchal (Kala, 2002), followed by Sikkim and North Bengal (Samant *et al*., 1998). The trans-Himalaya sustains about 337 species of medicinal plants (Kala, 2002), which is low compared to other areas of the Himalaya due to the distinct geography and ecological marginal conditions (Kala, 2002).

The curative properties of drugs are due to the presence of complex chemical substances of varied composition (present as secondary plant metabolites) in one or more parts of these plants. These plant metabolites according to their composition are grouped as alkaloids, glycosides, corticosteroids, essential oils, etc. The alkaloids form the largest group, which includes morphine and codein (Poppy), strychnine and brucine (nux vomica), quinine (cinchona), ergotamine (ergot), hypocampus (beeladona), scolapomine (datura), emetine (ipecac), cocaine (coco), ephedrine (ephedra), reserphine (*Rauwolfia*), caffeine (tea dust), aconitine (aconite), vascine (vasaca), santonin (*Aremisia*), lobelin (*Lobelia*) and a large number of others. Glycosides form another important group represented by digoxin (foxglove), stropanthin (*Strophanthus*), glycyrrhizin (liquorice), barbolin (aloe), sennocide (senna), etc. Corticosteroids have come into prominence recently and diosgenin (*Dioscorea*), solasodin (*Solanum* sp.), etc. now command a large world demand. Some essential oils such as those of valerian kutch and peppermint also possess medicating properties and are used in pharmaceutical industry. However, it should be stated in all fairness that our knowledge of the genetic and physiological make-up of most of the medicinal plants is poor and we know still less about the biosynthetic pathways leading to the formation of active constituents for which these crops are valued.

1.2. *Bacopa monniera*

1.2.1 History

Bacopa monniera L. (BM) Pennell, commonly known as Water Hyssop, brahmi, jal brahmi and nir-brahmi, has been used for centuries in the Ayurveda, a holistic system of medicine originating from India. The name brahmi is derived from the word “Brahma” the mythical “creator” in the Hindu pantheon. Because the Brain is the centre for creative activity, any compound that improves the brain health is called brahmi which also means ‘bringing

knowledge of the supreme reality' (Russo & Borrelli, 2005). According to Hackman (1998), early Hindu religious practices, prior to written history, require the memorization and repetition of lengthy, orally transmitted Vedic scriptures by scholars. *Bacopa* was reputedly used in these early times to enhance the scholar's capacity to memorise these epic hymns and scriptures (ibid.). Other Sanskrit (ancient Indian language) names for this plant are "Bahuphene", "Ahiphena", "Phenavati". The word "Phena" means "foaming property". When mixed with water *Bacopa* plant parts produce a stable froth that is attributed to the saponins present in the plant.

The herb has been described in Ayurvedic texts since around 800 BC and recorded as a treatment for the range of mental disorders in the 'Charak Samhita' (Singh & Dhawan, 1997), which according to the literature, was in the 16th century AD (Chowdhuri *et al*., 2002; Russo & Borrielli, 2005). In the later treatise of the 16th century, the Bhavprakasa Varg-Prakarana, *Bacopa's* actions are set down as follows: bitter, laxative, astringent, brain tonic, memory enhancing and longevity promoting. As well as brain conditions such as epilepsy, insanity and neuroses, other indications described in this treatise include aneamia, leprosy, renal disease, blood disease, poisoning and cough (Singh & Dhawan, 1982).

Ayurvedic medicine classifies *Bacopa* as belonging to a group of plant medicines- known as medhya rasayana- that improve mental health, intellect and memory (medhya) and promote longevity and rejuvenation (rasayana) (Singh & Singh, 1980). Hence *Bacopa* shares its Sanskrit name, brahmi, with another herbal nervous system restorative, *Centella asiatica.*

In a priority list of the most important medicinal plants, evaluated on the basis of their medicinal importance, commercial value and potential for further research and development, *B. monniera* was placed second according to a sector study by the Export- Import Bank of India. According to an estimate, the annual requirement of the plant was projected to be about 12,700 tonnes of dry material, valued at approximately Rs 15 billion.

1.2.2 Description

Bacopa monniera, a member of the *Scrophulariaceae* family (figwort or snapdragon family), is an annual or perennial, decumbent or erect stemmed plant. It is a creeping, diffusely branched, somewhat succulent herb, 10-30 cm long with numerous ascending branches.

Leaves are sessile or short-petioled, ovate-oblong or decussate, 0.6-2.5 cm long and 0.2-1 cm wide, fleshy and dotted with black specks. Flowers are white to pale blue or violet, solitary, borne in leaf axils on slender pedicels 0.6-32 cm long; the glabrous calyx is divided to the base, sepals about 6 mm long, ovate or lanceolate, acute; the corolla 8 mm long has a broadly campanulate tube and nearly equal lobes, rounded with shining dots when fresh. Fruits (capsules) about 5 mm long, ovoid, acute, included in the persistent calyx; seeds 0.85 mm long, oblong, striate, pale. Dispersal and propagation is by seeds and stem fragments. Crushed leaves have a distinctive 'lemon' scent. The herb is found in wetlands throughout the Indian subcontinent in damp and marshy or in sandy area near streams in tropical regions. The genus *Bacopa* includes over 100 species of aquatic herbs distributed throughout the warmer regions of the world, apart from India, Nepal, Srilanka, Pakistan, Afganistan, China, Taiwan and Vietnam and is also found in Africa, Australia, Florida and other Southern states of the USA. The herb can be found at elevations from sea level to altitudes of 4,400 feet, and is easily cultivated if adequate water is available. Flowers and fruit appear in summer and the entire plant is used medicinally (Bone, 1996; Chopra, 1958; Anon, 2004).

Fig-1.1: Bacopa plant (A) whole plant, (B) flower and (C) leaves.

1.2.3 Classification

- Kingdom: Plantae
- Phylum: Magnoliophyta
- Class: Magnoliopsida
- Order: Scrophulariales
- Family: Scrophulariaceae
- Genus: *Bacopa*
- Species: *monniera*

1.2.4 Active Constituents

In view of the importance of this plant in the indigenous system of medicine, systematic chemical examinations of the plant have been carried out by several groups of researchers. Detailed investigations were first documented in 1931, when Bose and Bose reported the isolation of the alkaloid ‘‘brahmine’’ from *Bacopa monniera*. Later, other alkaloids like nicotine and herpestine have also been reported (Chopra *et al*., 1956). Several glycosides such as asiaticoside and thanakunicide, flavonoids such as apigenin and luteonin, phytochemicals such as betulinic acid, wogonin, oroxindin, betulic acid, stigmastarol, beta-sitosterol, as well as brahamoside, brahminoside, brahmic acid, isobrahmic acid, vallerine, pectic acid, fatty acids, tannin, volatile oil, ascorbic acid, thanakunic acid and asiatic acid (Rastogi *et al*., 1965; Rastogi *et al*., 1967; Rastogi *et al*., 1973; Rastogi *et al*., 1974; Rastogi *et al*., 1977) have also been reported. The isolation of D-mannitol, and a saponin, hersaponin and potassium salts by Sastri *et al*., (1959) provided further details of the chemical components of *Bacopa monniera.*

The major chemical entity shown to be responsible for the memory-facilitating action of *Bacopa monniera*, bacoside A, was assigned as 3-(α-L-arabinopyranosyl)-O-β-D-glucopyranoside-10,20-dihydroxy-16-keto-dammar-24-ene (Chatterji *et al*., 1965). Bacoside A usually co-occurs with bacoside B, the latter differing only in optical rotation and probably an artifact produced during the process of isolating bacoside A (Rastogi, 1990) as shown in Fig.1.2. The chemical composition of bacosides, contained in the polar fraction, has been established on the basis of chemical and physical degradation studies. On acid hydrolysis,

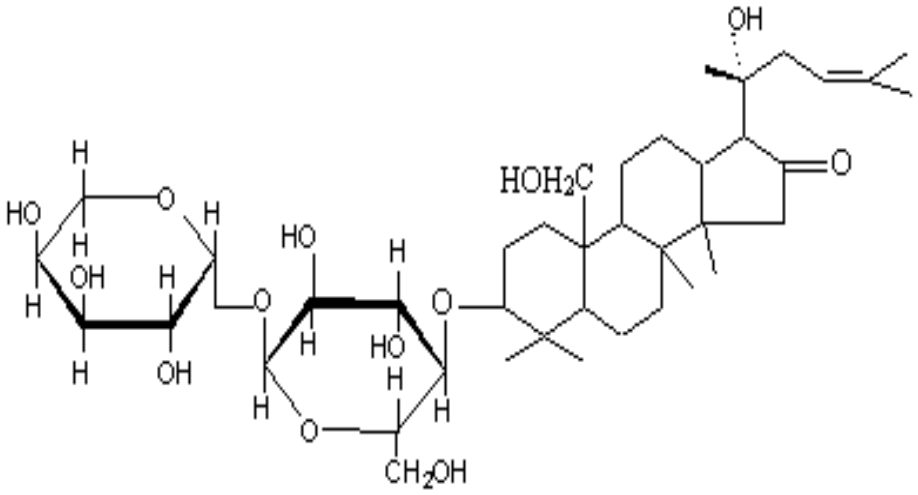
bacosides yield a mixture of aglycones, bacogenin A1, A2, A3 (Kulshreshtha and Rastogi, 1973, 1974; Chandel *et al*., 1977), which are artifacts, and two genuine sapogenins, jujubogenin and pseudojujubogenin (Rastogi *et al*., 1994). Another bacogenin, A4 was identified as ebelin lactone pseudojujubogenin (Rastogi *et al*., 1994). Successively, a minor saponin bacoside A1 was isolated and characterized as 3-O-[α-L-arabinofuranoyl(1-3)-β-L-arabinopyranosyl] jujubogenin (Rastogi *et al*., 1994). Rastogi *et al.* (1994) evidenced a new triperpenoid saponin, bacoside A3. Its structure was established as 3-β-[O-β-D-glucopyranosyl(1-3)-O-[α-L-arabinofuranosy(1-2)]-O-β-D-glucopyranosyl)oxy] jujubogenin by chemical and spectral analyses. Garay *et al.* (1996a) isolated from *Bacopa monniera* three new dammarane-type triterpenoid saponins of biological interest, bacopasaponins A, B and C, identified as 3-O-α-L-arabinopyranosyl-20-O-α-L-arabinopyranosyl-jujubogenin, 3-O-[α-L-arabinofuranosyl(1-2)α-L-arabinopyranosyl] pseudojujubogenin and 3-O-[β-D-glucopyranosyl(1-3){α-L-arabinofuranosyl(1-2)}α-Larabinopyranosyl] pseudojujubogenin by spectroscopic and chemical transformation methods. The hitherto undetermined configurations at C-20 and C-22 of pseudojujubogenin were elucidated by phase-sensitive ROESY, and 1H and 13C signals of the saponins were assigned by DEPT, 1H–1H COSY, HSQC and HMBC techniques. Successively, the same authors (Garay *et al*., 1996b) isolated a new dammarane-type pseudojujubogenin glycoside, bacopasaponin D, defined as 3-O-[α-L-arabinofuranosyl(1-2)β-D glucopyranosyl] pseudojujubogenin by spectroscopic and chemical transformation methods. It is noteworthy that dammarane-type triterpenoid saponins are major constituents of a number of reputed herbal drugs, including *Ginseng*. Although jujubogenin glycosides have been isolated from several reputed medicinal plants (i.e. *Rhamnaceae* and *Scrophulariaceae*), pseudojubogenin glycosides have only been reported in this Indian herbal drug.

In view of the increasing interest on this herbal drug, Chakravarty *et al.* (2001) undertook a thorough chemical reinvestigation of the glycosidic fraction of the methanol extract of the plant and were able to isolate two new pseudojujubogenin glycosides designated as bacopaside I and II. Their structures have been elucidated as 3-O-α-L-arabinofuranosyl-(1-2)-[6-O-sulphonyl-β-D-glucopyranosil-(1-3)]-α-L-arabinopyranosylpseudojujubogenin and 3-O-α-L-arabinofuranosyl-(1-2)-[β-D-glucopyranosil-(1-3)]β-D-glucopyranosylpseudojujubog-

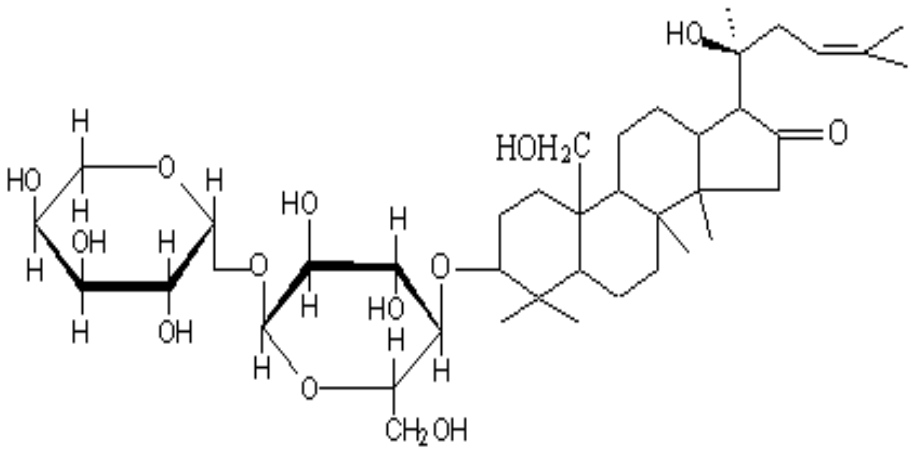
enin, mainly on the basis of dimensional (2D) NMR and other spectral analyses. Subsequently, the same authors (Chakravarty *et al*., 2003) isolated three new saponins from *Bacopa monniera*, designated as bacopasides III, IV and V, with structures 3-O-α-L-arabinofuranosyl-(1-2)-β-D-glucopyranosyl jujubogenin, 3-O-β-D-glucopyranosyl-(1-3)-α-L-arabinopyranosyl jujubogenin, 3-O-β-D-glucopyranosyl-(1-3)-α-L-arabinofuranosyl pseudoju jubogenin, on the basis of dimensional (3D) NMR and other spectral analyses.

In addition, Hou *et al.* (2002) have isolated a new saponin, 3-O-[α-1-arabinofuranosyl-(1-2)]-α-L-arabinopyranosyl jujubogenin, named bacopasaponin G, a new matsutaka alcohol derivative, (3R)-1-octan-3yl-(6-O-sulfonyl)-β-D-glucopyranoside, a new phenylethanoid glycoside,3,4-dihydroxyphenylethyl alcohol (2-O-feruloyl)-β-D-glucopyranoside, and a new glycoside, phenylethyl alcohol [5-O-p-hydroxybenzoyl-β-D-apiofuranosyl-(1-2)]-β-D-glucopyranoside.

Moreover, three new phenylethnoid glycosides, viz. monnierasides I–III along with the known analogue plantainoside B have been isolated from the glycosidic fraction of BM (Chakravarty *et al*., 2002).



a. Bacosides A



b. Bacoside B

Fig. 1.2: Chemical structure of (a) Bacosides A (levorotatory) and (b) Bacoside B (dextrorotatory).

1.2.5 The Mechanism for the sustainable memory

Russo *et al*. (2003a) suggested that because of its ability to reduce NO-induced cellular alterations, brahmi has a therapeutic potential in treatment or prevention of neurological diseases. It was also reported that extract of brahmi rich in saponins, is able to induce a dose-related increase in superoxide dismutase (SOD), catalase (CAT) and glutathione peroxide (GSH-Px) activities in rat frontal cortex, striatum and hippocampus (Bhattacharya *et al*., 2000).

The hippocampus, the brain's seat of memory, is located in the temporal (left and right) sides of the brain. It processes signals sent to the brain by the senses into the templates of memory, which are then stored in other parts of the brain, creating a long-term memory. Signals are converted into electrical impulses in the nerve cells due to a rapid change in protein composition. These impulses are then conducted across neurons and through synapses, which connect nerve cells. This process continues until the bond between the nerve cells strengthen and memory is created.

Normal synaptic activity is a process mediated by neurotransmitters. Each neuron is a single nerve cell. It has one or more arms called axons that send signals (impulses) and one or

more arms called dendrites that receive signals. When a signal is transmitted through an axon terminal, spherical bodies called vesicles fuse with its membrane. Neurotransmitters are released when the vesicles burst open into the synaptic space, the minute space between the sending and the receiving cells used to discharge neurotransmitters ("Passengers"). To end the signal, the axons reabsorb some neurotransmitters; and the enzymes in the synapse neutralize the other neurotransmitters. It is evident that a disruption in any part of normal synaptic activity would affect memory. This normally occurs with advanced age and continuous electrical activity, which wears out the synapses. As a result, new memory creation is impaired and memory loss occurs.

Chemical substances and plants extracts that are known to restore the memory work in different ways. The bacosides are the memory chemicals in Brahmi (Rastogi *et al*., 1994; Sivaramakrishna *et al*., 2005). The active ingredients vary from 0.9 to 1.8% (Mathur *et al*., 2002). According to scientists at the Central Drug Research Institute (CDRI) located in Lucknow, India, the bacosides help to repair damaged neurons by adding muscle to kinase, the protein involved in the synthesis of new neurons to replace the old ones. Depleted synaptic activity is thus restored, leading to augmented memory functions.

1.2.6 Mechanism of Action

Since *Bacopa's* primary therapeutic use is to enhance cognitive function, most of the research has been focused on the mechanism behind these properties. The triterpenoid saponins and their bacosides are responsible for *Bacopa's* ability to enhance nerve impulse transmission. The bacosides aid in repair of damaged neurons by enhancing kinase activity, neuronal synthesis, and restoration of synaptic activity, and ultimately nerve impulse transmission and boosting the synthesis of new protein in the brain (Singh & Dhawan, 1997).

Loss of cholinergic neuronal activity in the hippocampus is the primary feature of Alzheimer's disease. It is indicated by depleted concentrations of the neurotransmitters, acetylcholine, a reduction in the choline esterase activity and decreased muscarinic cholinergic receptor binding in the frontal cortex and hippocampus in rats (Enz *et al*., 1993). Animal research has shown that *Bacopa* extracts modulate the expression of certain enzymes involved in generation and scavenging of reactive oxygen species in the brain (Chowdhuri *et*

al., 2002). *In vitro* research has also shown that *Bacopa* exerts a protective effect against DNA damage in astrocytes and human fibroblasts (Russo *et al*., 2003a, b).

In animals, *Bacopa* has a relaxant effect on pulmonary arteries, aorta, trachea, and ileal and bronchial tissue, possibly mediated by inhibition of calcium- ion influx into cell membranes (Channa *et al*., 2003). *Bacopa* appears to stabilize mast cells *in vitro*, (Samiulla *et al*., 2001) and possesses anti-inflammatory activity via inhibition of prostaglandin synthesis and lysosomal membrane stabilization (Jain *et al*., 1994). *In vitro* research suggests an anticancer effect for *Bacopa* extracts, possibly due to inhibition of DNA replication in cancer cell lines (Elangovan *et al*., 1995).

1.2.7 Traditional uses of *Bacopa monniera*

Bacopa monniera often called "Brahmi"or "The thinking person's herb," is widely used in traditional Indian Ayurvedic System of medicine as a potent nervine tonic to enhance memory function, improve intellectual and cognitive functions (Rastogi *et al*., 1994). It is also used in treatment of asthama, leprosy, hoarseness, water retention and blood cleaning (Singh and Dhawan, 1982).

B. monniera has been used in Ayurvedic formulations for conditions ranging from catarrhal complaints, gastrointestinal disturbances due to excessive tobacco use, habitual abortions and high blood sugar due to anxiety disorders, hysteria, epilepsy etc. (Chopra *et al*., 1956; Nadkarni, 1976). In certain parts of India, Brahmi is believed to be an aphrodisiac. In Srilanka, under the name of Loonooweela, Brahmi is prescribed for fevers. In the Phillipines, it is used as a diuretic (Uphof, 1968).

The traditional use of the plant, of particular relevance to contemporary medicine, is its validation efficacy in promoting memory functions and providing relief to patients with anxiety neurosis. In Ayurveda, Brahmi is described as "medhya rasayana"or brain tonic with the ability to promote mental functioning along with providing general rejuvenative effects (Singh and Singh, 1980).

Brahmi has a bitter taste. Traditionally, the fleshy leaves and stems were made into paste or pressed for juice extraction. Sugar, jaggery or honey was added to make it more palatable. Some of the known preparations with brahmi are Brahmi Ghrita (in clarified

butter), sarasvatarisha (a decoction used as a brain tonic), Brahmi Rasayana (a rejuvenating formulation with other herbs), Brahmi Sarbat (a cooling drink).

In recent years, "Memory Plus", a product that contains the standardized extract of bacosides from brahmi, has been marketed in India. Under the definition of herbal drugs in the guidelines of herbal medicines issued by the World Health Organization (WHO) in 1991, an herbal product that has been used traditionally without demonstrable harm does not require specific regulatory action unless new evidence demands a revised risk/ benefit analysis. Subsequently, several formulations containing *B. monniera* extracts standardized for bacoside content have appeared in the global market place.

The evaluation of BacoMind^{TM} in healthy adult volunteers revealed that, at the given oral dose of 300 mg once a day for first 15 days and 450 mg once a day for next 15 days, the herbal supplement was found to be safe and tolerable. Though minor gastrointestinal side effects were reported in 3 out of 23 volunteers, the general physical, systemic, hematological, biochemical and electrocardiographic parameters were within the normal limits (Pravina *et al*., 2007). It was also reported that BacoMind^{TM} can be used as a natural dietary supplement for supporting optimal health and mental functioning and is safe for oral consumption. It also has the capacity to be used as an anti-oxidant and an anti-carcinogenic agent particularly against those carcinogens that mediate their effects via formation of oxygen free radicals (Deb *et al*., 2008).

Therefore, ayurvedic medicines have been used for traditional treatment of a number of disorders, particularly those involving anxiety, intellect and poor memory (Singh and Dhawan, 1997), epilepsy, insanity and retardation (Mathur *et al*., 2002), to counteract the effects of mental stress and neurosis and revitalization of sensory organs (Sivarajan and Balachandran, 1994). The well-known nootropic plant has been reported to possess sedative (Malhotra and Das, 1959), cardiotonic (Mathur *et al.*, 2002), cognitive enhancer (Nathan *et al*., 2001; Roodernrys *et al*., 2002), broncho-vasodilator (Channa *et al*., 2003), hepatoprotective (Sumathy *et al*., 2001; Sumathi and Nongbri, 2008), antidepressant (Sairam *et al*., 2002), calcium antagonistic (Dar and Channa, 1999), smooth muscle relaxant (Dar and Channa, 1997), neuropharmacological (Russo and Borrelli, 2005) cell stabilizing (Samiulla *et al*., 2001) and antiulcer (Sairam *et al*., 2001) properties. The said plant also showed antistress effect, via modulation of Hsp 70 expression, superoxide dismutase and cytochrome p450

inhibitory activity in rat brain (Chowdhuri *et al*., 2002) and antioxidant activity (Tripathy *et al*., 1996). *Bacopa monniera* is reported to play a protective role on morphine-induced brain mitochondrial enzyme status in rats (Sumathy *et al.*, 2002). It is also active against leishmaniasis (Sinha *et al.*, 2002).

It helps to regain general mental health through its rejuvenative effect. It is currently being marketed in Western countries as a memory enhancing agent. Studies have shown that the herb contains many active constituents, including a number of alkaloids and saponins. However, the major constituents are the steroidal saponins, bacosides A and its optical isomer bacoside B, several other chemicals like stigmastanol, betasitosterol and stigmasterol are found in all parts of the plant (Mathur *et al*., 2002). The recent discovery of the memory-enhancing property of the bacosides has enhanced the demand of this plant and resulted in its extensive use in several commercial preparations. Therefore, mass cultivation of this plant in consideration with its quantitative and qualitative improvement is the prime interest of herbal research.

A growing interest is centered on the protection of host organism by means of antioxidants because of their putative role in the modulation of smoking induced stress protein expression and apoptosis (Pinot *et al*., 1997; Aoshiba *et al*., 2001). Besides, the plant extract has been shown as potent free radical scavenger, antilipid peroxidative agent and antioxidant (Tripathy *et al*., 1996), antiaging potential (Jyoti and Sharma, 2006). Preliminary studies indicated that bacosides, the major saponins are responsible for the facilitatory and modulatory effects of *Bacopa monniera* (Singh *et al*., 1988).

Anbarasi *et al*. (2005a, b, 2006) have reported the protective effects of bacoside A on cigarette smoking-induced brain damage. The role of bacoside A against chronic cigarette smoking-induced hsp70 expression and apoptosis in rat brain cerebral cortex. Apoptosis was monitered by DNA fragmentation and terminal deoxynucleotidyl transferase-mediated deoxyuridine triphosphate nick end labeling (TUNEL) staining, and further confirmed by electron microscopy (Anbarasi *et al*., (2005c).

The studies have shown that it improves the rate of learning in a brightness discrimination task and a conditional avoidance task, that it improves retention, as demonstrated by savings in relearning, and that it attenuates amnesia induced by immobilization, electroconvulsive shock and scopolamine (Singh and Dhawan, 1997). This

latter finding involved administration of the extracted bacosides A and B, and suggests that they influence cholinergic systems. Bacosides, Brahmi's active principles responsible for improving memory related functions, are attributed with the capability to enhance the efficiency of transmission of nerve impulses, thereby strengthening memory and cognition.

1.2.8 Preclinical and Clinical studies

1.2.8a.1 Beneficial effects on learning and memory

The plant, plant extract and isolated bacosides have been extensively investigated in several laboratories for their neuropharmacological effects and a number of reports are available confirming their nootropic action.

Preliminary studies established that the treatment with the plant (Malhotra and Das, 1959) and with the alcoholic extract of BM plant (Singh and Dhawan, 1982) enhanced learning ability in rats. Subsequent studies indicated that the cognition-facilitating effect was due to two active saponins, bacoside A and B, present in the ethanolic extract (Singh and Dhawan, 1992). These active principles, apart from facilitating learning and memory in normal rats, inhibited the amnesic effects of scopolamine, electroshock and immobilization stress (Dhawan and Singh, 1996). The mechanism of these pharmacological actions remains conjectural. It has been suggested that the bacosides induce membrane dephosphorylation, with a concomitant increase in protein and RNA turnover in specific brain areas (Singh *et al*., 1990). Further *B. monniera* has been shown to enhance protein kinase activity in the hippocampus which could also contribute to its nootropic action (Singh and Dhawan, 1997). A study of Bhattacharya *et al.* (1999) reported that a standardized bacoside-rich extract of *B. monniera,* administered for 2 weeks in rats, reversed cognitive deficits induced by intracerebroventricularly administrated colchicines and by injection of ibotenic acid into the nucleus basalis magnocellularis.

The central cholinergic system is considered the most important neurotransmitter involved in the regulation of cognitive functions. Cholinergic neural loss in hippocampal area is the major feature of Alzheimer's disease and enhancement of central cholinergic activity by anticholinesterase is presently the mainstay of the pharmacotherapy of Alzheimer's disease

type senile dermantia. Administration of *B. monniera* for 2 weeks also reversed the depletion of acetylcholine, the reduction in choline acetylase activity and the decrease in muscarinic cholinergic receptor binding in the frontal cortex and hippocampus, induced by neurotoxins, colchicines (Bhattacharya *et al*., 1999).

It has been suggested that the behavioural effects of cholinergic degeneration can be alleviated by a reduction in noradrenergic function (Sara, 1989). *B. monniera* is known to lower norepinephrine and increase 5-hydroxytryptamine levels in the hippocampus, hypothalamus and cerebral cortex (Singh and Dhawan, 1997). *B. monniera* may thus, also indirectly, modify acetylcholine concentrations, through its influence on other neurotransmitter systems.

1.2.8a.2 Antidepressant and Antianxiety effects

Research using a rat model of clinical anxiety demonstrated that a *B. monniera* extract containing 25% bacoside A exerted anxiolytic activity comparable to lorazepam, a common benzodiazephine anxiolytic drug, and it was attentively noted that the BM extract did not induce amnesia, side effects associated with lorazepam, but instead had a memory-enhancing effects (Bhattacharya *et al*., 1998; Singh *et al*., 2000). The antidepressant potential of BM has been evaluated in an earlier study, wherein it showed a significant antidepressant activity in the most commonly used behaviour paradigms in animal models of depression, namely, forced swim test and learned helplessness tests (Bhatacharya *et al*., 2002). In the study, the BM extract in the dose range of 20-40 mg/kg was given once daily for 5 days and it was found comparable to standard antidepressant drug imipramine in antidepressant activity in rodent animals. The same study has postulated the role of serotonin and GABA (gamma amino butyric acid) in the mechanism of action attributed for its antidepressant action along with its anxiolytic potential, based on the compelling evidence that the symptoms of anxiety and depression overlap each other (Shader and Greenblatt, 1995).

1.2.8a.3 Anti-Epileptic effects

Although BM has been indicated as a remedy for epilepsy in Ayurvedic medicine (Shanmugasundaram *et al*., 1991), One Indian study examined the anticonvulsant properties of BM extracts in mice and rats. Researchers determined that intraperitonial injections of high doses of BM extract (close to 50% of LD_{50}) given for 15 days demonstrated anticonvulsant activity. When administered acutely at lower doses (approaching 25% of LD_{50}), anticonvulsant activity was not observed (Ganguly and Malhotra, 1967). It is postulated that the anti-convulsive effects could be mediated through GABA which is involved in neural impulse transmission, because substances which stimulate GABA are known to possess anticonvulsant, pain relieving and sedative activities.

1.2.8a.4 Antioxidant and Adaptogenic properties

BM extracts or bacosides have shown an antioxidant (Singh *et al*., 2006; Bafna *et al*., 2005; Sumathy *et al*., 2002; Pawar *et al*., 2001; Tripathy *et al*., 1996; Kapoor *et al*., 2008) and antistress activity (Dhawan *et al*., 1969). A previous study suggests an involvement of GABA-ergic system in the mediation of these central nervous system effects of BM (Singh, 1996). Based on animal study results, bacosides were shown to have antioxidant activity in the hippocampus, frontal cortex and striatum (Bhatacharya, 2000). Animal research has shown that the BM extracts modulate the expression of certain enzymes involved in generation and scavenging of reactive oxygen species in the brain (Govindarajan, 2005). It was suggested that the adaptogenic properties of the herb would be beneficial in the management of stress related conditions as BM showed the potential to be effective in stress in a study on rats (Chowdhuri, 2002). In the study BM extract was found not only to induce the constitutive expression of heat-shock protein (Hsp70) but also induce the CYP 450 enzymes in all regions of brain. The level of Hsp70 was found to be increased in brain as a response of stress. On the other hand, the group that was pre-treated for 1 week with 20-40 mg/kg/daily, before giving stress, the Hsp70 was found to be in lower concentration. An increase in the activity of CYP 450-dependent enzymes 7-pentoxyresorufin-odealkkylase (PROD) and 7-ethoxyresorufin-o-deethylase (EROD) was observed in all the brain regions

after exposure to stress alone and with both doses of BM extract, although the magnitude of induction observed was less with a higher dose of the same. Thus it was suggested that the BM primed the brain for stress by stockpiling these useful enzymes even before stressful conditions and that our susceptibility to stress could be lowered by using this medicinal herb. It was speculated that this induction may be an adaptive response to the stress which needs further investigation. The level of superoxide dismutase (SOD) was also increased in brain in the groups pre-treated with BM extracts. The data indicated that BM extracts has a potential to modulate the activities of Hsp70, CYP 450 and SOD and thereby possibly allowing the brain to be prepared to act under adverse conditions like stress.

Researchers concluded that BM helps in coping with combined hypoxia, hypothermic and immobilization stress that could lead to onslaught of 'free radicals' (Rohini, 2004). The results of the above-mentioned study have indicated that this extract exhibits interesting antioxidant properties, expressed by its capacity to scavenge superoxide anion hydroxyl radical, and to reduce H_2O_2-induced cytotoxicity and DNA damage in human fibroblast cells (Tripathy *et al*., 1996; Seiss H, 1993; Singh *et al*., 2003). BM extract has shown neuroprotective effect against aluminium-induced oxidative stress in the hippocampus of rat brain (Jyoti and Sharma, 2006). Aqueous extract of BM reduced nicotine-induced lipid peroxidation (LPO) and conferred geno protection in swiss mice in one study (Vijayan and Helen, 2007). Yet another study suggested that BM extract reduces amyloid levels in PSAPP mice and can be used in the therapy of Alzheimer's disease (Holcomb *et al*., 2006). One of the recent studies has shown the protective role of bacoside A against chronic cigarette smoking-induced damage in rat brain (Anbarasi *et al*., 2006). This antioxidant activity of BM is able to explain, atleast in part, the reported antistress, cognition-facilitating and antiaging effects produced by it in experimental animals and in clinical situations (Aloe *et al*., 2002) and may justify further investigation of its other beneficial biological properties.

1.2.8a.5 Gastrointestinal effects

Some *in vitro*, animal and human studies have investigated the effects of BM extract on the gastrointestinal tract. *In vitro* studies have demonstrated direct spasmolytic activity on intestinal smooth muscles, via inhibition of calcium influx across cell membrane channels.

This property suggests that BM extracts may be of benefit in conditions characterized by intestinal spasm such as irritable bowel syndrome (IBS) (Dar and Channa 1997, 1999). The results indicated the direct action of the extract on smooth muscles. Also calcium chloride-induced responses observed in the rabbit's blood vessels jejunum were reduced in the presence of the BM extract (10-700 mcg/ml), suggesting direct interference with the influx of calcium ions. However, since the extract did not affect contractions induced by noradrenalin caffeine, the author concluded that the extract had no appreciable effect on the mobilization of intracellular calcium. Based on the results of the experiment, it is postulated that the spasmolytic effect of BM extract on smooth muscles is predominantly due to the inhibition of calcium influx, appreciable to both electrical impulse-mediated and receptor-mediated calcium channels in the cell membrane. Animal and *in vitro* studies suggested that BM may have a protective and curative effect on gastric ulcers, and studies were reported for its antiulcerogenic activity (Dorababu *et al*., 2004; Jain *et al*., 1994; Rao *et al*., 2000; Dharmani and Palit, 2006; Goel and Sairam, 2002). In rats, a BM extract standardized for bacoside A was evaluated for its prophylactic healing effects in five models of gastric ulcers (Sairam *et al*., 2001). At a dose of 20 mg/kg for 10 days, BM extract significantly healed penetrating ulcers induced by acetic acid, significantly strengthened the mucosal barrier and decreased mucosal exfoliation. The extract also alleviated stress-induced ulcers as observed by significant reduction in LPO in rat gastric mucosa. BM's antioxidant properties and its ability to balance SOD catalase levels were postulated to account for this effect (Sairam *et al*., 2001). Recently *in vitro* study also demonstrated its specific anti-microbial activity against *Helicobacter pylori*, a bacterium associated with chronic gastric ulcers. When the extract was incubated with human colonic mucosal cells and *H. pylori*, it resulted in the accumulation of prostaglandin E and prostacycline, Prostaglandins known to be protective for gastric mucosa (Goel *et al*., 2003).

1.2.8a.6 Miscellaneous Studies

In vitro research has shown a protective effect of BM against DNA damage in astrocytes (Russo *et al*., 2003a) and human fibroblasts (Russo *et al*., 2003b). *In vitro* research has suggested that an anticancer effect of BM extracts is possible due to inhibition of DNA

replication in cancer cell lines (Elangovan *et al*., 1995). A study in mice demonstrated high doses (200 mg/kg) of BM extract increased the thyroid hormones, T4, by 41% when given orally. T3 was not stimulated, suggesting that the extract may directly stimulate synthesis and/or release of T4 at the glandular level, while not affecting conversion of T4 to T3. While this study indicated that BM extract did have a stimulatory effect on thyroid functions, the doses were very high and it was assumed that the typical 200-400 mg daily dose in humans may not have the same effect (Kar *et al*., 2002). BM was also reported to possess anti-inflammatory activity via inhibition of prostaglandin synthesis and lysosomal membrane stabilization (Jain *et al*., 1994; Channa *et al*., 2006). *In vitro* research using rabbit aorta and pulmonary artery has demonstrated that BM extract exerts a vasodilatory effect on calcium chloride-induced contraction in both tissues. It is believed to exert this effect via interference with calcium channel flux in tissue cells (Sairam *et al*., 2001). Animal studies have demonstrated that BM extract has a relaxant effect on chemically-induced bronchoconstriction, probably via inhibition of calcium influx into cell membranes. An earlier *in vitro* study demonstrated the broncho-vasodilatory activity of BM on rabbit and guinea pig trachea pulmonary artery and aorta (Dar and Channa, 1997). A subsequent rat study with BM extract confirmed the earlier results in which methanol sub-fractions of BM extract were given to anesthetized rats prior to induction of bronchoconstriction with carbachol, an acetylcholine analogue. Nearly all of the BM extract subfractions inhibited carbachol-induced bronchoconstriction, hypotension and bradycardia in this animal model (Channa *et al*., 2003). An *in vitro* study also demonstrated that a methanol extract of BM possessed potent mast cell stabilizing activity comparable to disodium cromoglycate, a commonly used allergy medication (Samiulla *et al*., 2001). These studies indicated the potential usefulness of BM extracts in bronchoconstrictive and allergic conditions and warrant human studies.

1.2.8b Clinical trials

In the clinical trials described below, *Bacopa* has been found to improve various aspects of cognitive function in children and adults. Sharma *et al*. (1987) found that learning, memory, perception and reaction times improved in 20 primary school children given *Bacopa* in syrup form at a dosage of 350 mg three times daily for three months. No side effects were reported.

This study was strengthened by the use of a matched placebo control group (n=20). However, the study was limited in that it was not double-blinded, and used a small sample size.

Negi *et al*. (2000) reported children with attention deficit hyperactivity disorder (ADHD) were found to benefit from *Bacopa* administration. A randomised, doubleblind, placebo-controlled trial of 36 children with ADHD was carried out by the Indian researchers. *Bacopa* was given at a dosage of 50 mg twice daily for 12-weeks, and a battery of cognitive function tests administered at baseline, 4, 8, 12 and 16-weeks (i.e. 4-weeks post trial). Improvements were reported in the active treatment group (n=19) at 12-weeks, as measured by tests of sentence repetition, logical memory, and paired associate learning tasks. Interestingly, improvement was still apparent at 16-weeks, four weeks after cessation of *Bacopa* administration.

In an open trial, 35 adults with anxiety neurosis were treated with *Bacopa* at the dose of 12 gm of dried herb daily in syrup form for 4-weeks. No significant side effects were observed and results were highly favourable as overall anxiety levels, concentration and memory span were all significantly improved along with other major anxiety-related physical symptoms and biochemical markers of anxiety (Singh & Singh, 1980). Whilst the findings of this study are encouraging, the lack of a control group is a significant design flaw, and further, the researchers do not delineate the selection criteria for inclusion in the study.

Bacopa's anxiolytic action was supported in a later animal study in which it was compared with a major pharmacological anxiolytic agent- the benzodiazepine lorazepam, in validated rat models of anxiety. It proved to be as effectual as the drug in every outcome measured without producing any motor deficits (a common side-effect of lorazepam) (Bhattacharya & Ghosal, 1998). Whilst promising, these findings indicate the need for further rigorous clinical trials to establish whether they are applicable to humans. It is noteworthy that support for an anxiolytic action in humans was provided by a well designed clinical trial in which state anxiety was significantly improved by Bacopa (Stough *et al*., 2001).

Some recent studies demonstrate that *Bacopa* may be effective for enhancing cognition in longer rather than shorter term administration. One study, using a randomised, double-blind, placebo-controlled design, tested the effects of *B. monniera* on cognitive function in forty six healthy adults between 18-60 years of age. Participants took *Bacopa* (300 mg daily) or placebo for 12-weeks and measurements, using a battery of well validated

neuropsychological tests, were recorded at baseline, 5-weeks and 12-weeks. Significant improvements were found in the active treatment group in speed of information processing (as measured by the Inspection Time task), learning rate and memory consolidation (as measured by the Rey Auditory Verbal Learning Test) and state anxiety levels (as measured by Spielberger's State Anxiety Inventory). Improvements were only found at 12-weeks and not earlier, suggesting that chronic administration is required to elicit *Bacopa's* effects (Stough *et al*., 2001). This study is well designed and rigorous; however, a limitation is that it does not take into account age, gender or education effects, all of which are known to effect performance on memory tests.

In contrast to the findings of Stough *et al*. (2001) members of the same research team reported that *Bacopa* had no acute effect on memory, when outcomes were measured at 2-hours post-administration (Nathan *et al*., 2001). In this study a randomised, double-blind, placebo-controlled trial was used to assess the acute effects of *Bacopa* (300 mg daily) on thirty eight healthy adults between 18-60 years of age. The same well validated neuropsychological test battery as employed in the Stough study (above) was used to assess the effects of *Bacopa* before and then 2-hours after administration. No effects were found.

Additionally, in a later randomised, double-blind, placebo-controlled study of 85 adults (aged 19-68 years) which tested cognitive effects of a combined tablet of 300 mg *Bacopa* with 120 mg *Gingko biloba*, no significant effects were found in outcomes measured at 2-weeks and 4-weeks of treatment duration (Nathan *et al*., 2004). Despite extensive searching of the literature it appears that the effects of *Bacopa* observed at 12-weeks have not been reported for lesser durations in well designed, larger human trials.

Another randomised, double-blind, placebo-controlled study confirmed the efficacy of *Bacopa* in improving memory in chronic administration. In this study, seventy six adults 40-65 years of age, were given *Bacopa* (dose 300 mg, or 450 mg for persons over 90 kg) or placebo and measured on tasks of attention, memory and psychological state at baseline, 12-weeks and 6-weeks post trial (Roodenrys *et al*., 2002). *Bacopa* significantly improved the ability to retain information over time as measured by a task requiring delayed recall of word pairs. The authors commented that this may be due to less information being lost from memory, that is, the results are due to decreased forgetting, as opposed to enhanced acquisition because learning trials did not show any effect of *Bacopa*. Outcome measures in

this study which failed to show a significant effect for *Bacopa* were tasks of short term memory, working memory, attention, retrieval of prior knowledge and psychological state (anxiety, stress and depression).

Roodenrys *et al*.'s (2002) finding of decreased forgetting rate as measured by a word pairs task supports the aforementioned findings of Stough *et al*. (2001) who also reported decreased forgetting rate, as measured by the AVLT. In contrast to Roodenrys *et al*. (2002) however, Stough's study, as well as the animal study of Singh and Dhawan (1997), did find an improvement in learning rate. A further contrast between the two studies is Stough *et al.*'s (2001) finding of *Bacopa's* efficacy in reducing anxiety. Additionally, Roodenrys *et al*.'s (2002) findings also contrasts with the anxiolytic effect observed in aforementioned studies by Singh & Singh (1980) and Bhattacharya & Ghosal (1998). The Roodenrys *et al*. (2002) study demonstrates a sturdy design, using well validated instruments to measure outcomes; however as in the Stough *et al*. (2001) study, age and education were not controlled for nor were the data adjusted for the effects of multiple measures. Thus there is an increased chance of a type one error.

A double-blind, placebo-controlled toxicological study in which bacosides were administered in various single doses (ranging from 20 mg to 300 mg) as well as multiple doses (100 mg and 200 mg) to healthy male volunteers for one month demonstrated an absence of any side-effects (Singh & Dhawan, 1997). The herb is TGA approved in Australia for over the counter sale as it is considered safe, and no adverse reactions are recorded in the literature. The traditional recommended dosage regimen is 5-10 gm of the powdered dried herb daily (Anonymous, Indian Herbal Pharmacopoeia, 2002:36).

1.2.9 Biotechnology and tissue culture studies on *Bacopa monniera*

In an effort to study the genetic diversity in the germplasm of *Bacopa monniera*, 24 accessions from different geographical regions of India and one introduction from Malaysia maintained at the Central Institute of Medicinal and Aromatic Plants, Lucknow, India, were analyzed for random amplified polymorphic DNA variation. The similarity between accessions was found to be in the range of 0.8-1.0, which is indicative of a narrow genetic base and a low to medium level of polymorphism. The individual accessions could be

differentiated, showing differences in morphological and growth properties at the DNA level. The low level of genetic variation was attributed to the interplay between sexual and vegetative reproduction and similarity of local environments in the habitats of the plant (Mulichak *et al*., 2001).

The plant has very high morphogenic potential, because of which explants from it respond very readily to treatment with plant growth regulators. Several protocols have been reported for the rapid multiplication of shoots and shoot cultures, and for the micropropagation of the plant (Mulichak *et al*., 2004; Freund *et al*., 2000). In most of the tissue culture studies, a very good response was obtained by 6-benzyl amino purine in producing adventitious shoots from different explants (Burkle *et al*., 2005; Llyod KO, 2000; Burcher *et al*., 2005; Rupprath *et al*., 2005) and somatic embryogenesis (Appel and Neu, 1997), which led to the development of successful protocols for the mass propagation of the plant (Appel and Neu, 1997; Schwalbe *et al*., 1987).

As *B. monniera* has a high morphogenic potential, the explants readily regenerated shoot buds in medium containing small amounts of cytokinins like 6-benzyl amino purine or kinetin (Freund *et al*., 2000). With a high regenerative potential, the stem and leaf explants showed a tendency to regenerate shoots and/ or roots even in media that are known to support callus initiation and growth (e.g., media supplemented with 2,4-dichlorophenoxyacetic acid), while the callusing response of the explants was very poor. The best response for shoot regeneration from both stem and leaf explants was obtained with 2 µM 6-benzyl amino purine in Murashige and Skoog (MS) media gelled with 0.2% Gelrite, where profuse introduction of multiple shoot buds was obtained in 9 days of incubation. Furthermore, after 3 weeks of shoot culture growth, when the shoots were harvested and the explant base was transferred into fresh medium, it continued to expand exponentially and regenerate new shoot buds. In addition, the leaf explants from shoot cultures proved to be superior explant material compared to the leaf explants from field-grown plants. The shoot cultures performed better when grown on medium gelled with Gelrite (0.2%) rather than with agar (0.7%). The phytochemical profile of the regenerated shoots was found to be similar to that of field-grown plants, as revealed by TLC analysis (Freund *et al*., 2000).

In an interesting study, *B. monniera* was shown to respond to the antibiotic trimethoprim and the fungicide bavistin, with bavistin inducing a good number of shoots from

internodal explants (Martone WJ, 1998). However, the growth of the regenerated shoots was stunted and for further optimum shoot growth it required the supplementation of plant growth regulators. This calls for further exploration of the potential of bavistin in shoot regeneration, especially since the induction response itself was not gradual and was a concentration-dependent response, causing a sudden surge of induction at 300 mg/L in internode and explants (Martone WJ, 1998).

Cell suspension cultures were established from leaf explants of *in vitro* plants, and two cell lines were identified that showed a 5- to 6-fold increase in fresh and dry weight in 40 days of culture in MS medium supplemented with 1 mg/L α-naphthalene acetic acid and 0.5 mg/L kinetin. In these two cell lines, bacosides A and B were detected from the 10^{th} day of culture, and they progressively accumulated up to the 40^{th} day. Bacoside A content was found to be higher than bacoside B all through the culture period, with a maximum of 1% dry weight basis (dwb) in 40-day cultures. Bacoside B accumulated up to 0.25-0.37% dwb in 40-day cultures (Jones, 1997).

Ali and coworkers (Rodriguez *et al*., 1998); Brown *et al*., 1975; Josse and Kornberg, 1962) used *B. monniera* grown *in vitro* as a model system for studying the effects of cadmium on plants, towards an effort to establish metal-tolerant plants, since cadmium is a widespread pollutant. It was found that the cultures could become acclimatized to cadmium through gradual exposure to increased concentrations of cadmium (Brown *et al*., 1975). Cadmium caused a reduction in photosynthetic rate, stomatal conductance, and internal carbon dioxide concentration, and supplementation with zinc improved these parameters under cadmium stress (Brown *et al*., 1975; Josse and Kornberg, 1962). There was a gradual increase in the accumulation of cadmium with an increase in concentration and duration of treatment, with maximum accumulation in root (Morera *et al*., 1999). Cadmium was shown to accumulate in the cell wall. Cadmium induced an increase in total protein content in the cultures (up to 50 µM cadmium), and caused the accumulation of proline (Rodriguez *et al*., 1998); Brown *et al*., 1975) it was suggested that these two parameters can serve as stress indicators (Josse and Kornberg, 1962). At higher concentrations (200 µM) cadmium caused oxidative damage with increased lipid peroxidation, decreased chlorophyll and protein content, and a concomitant induction in the activities of SOD, ascorbate peroxidase, and guicol peroxidase, perhaps to combat the oxidative stress caused by cadmium. However, there was a significant reduction in

catalase activity (Morera *et al*., 1999; Larivieri and Morera, 2002). Cadmium was found to induce phytochelatins in the root and leaf (Larivieri and Morera, 2002). Cadmium was shown to inhibit the growth of regenerants *in vitro*; copper partially alleviated this negative effect (Rodriguez *et al*., 1998).

In a study of the adaptation of *B. monniera* regenerants to salinity stress *in vitro*, the proline content of the plants was found to increase six-fold compared to the control, while photosynthetic rate, fresh mass, and root length of the regenerants decreased (Larivieri and Morera, 2004).

Tetraploid plants were generated by treating nodal explants from *in vitro*-grown plants with colchicine followed by culture in MS medium supplemented with 0.25 mg/L 6-benzyl amino purine. The tetraploid plants had larger leaves and flowers, but its growth was slow. The tetraploid obtained was named as a new variety, Ali INTA-JICA (Larivieri *et al*., 2003).

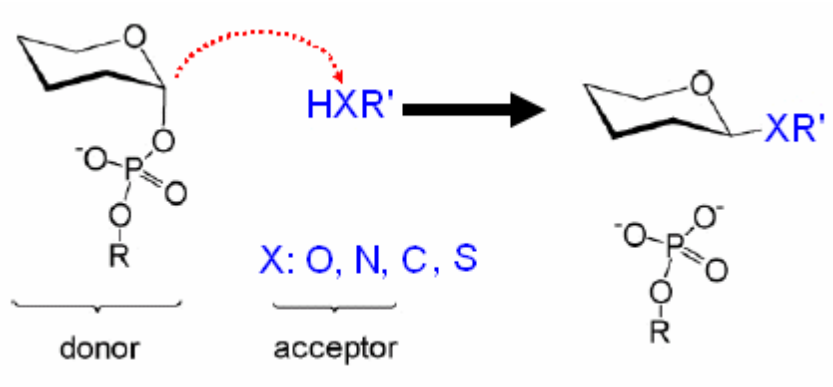
1.3. Glycosylation

Plants have evolved an extraordinary capacity to perceive changes in their environment and respond rapidly to maximize opportunities and minimize risks. The plasticity of these responses requires the integration of growth, development and metabolism, which in turn has led to the evolution of diverse mechanisms to regulate cellular homeostasis. Glycosylation is one of these mechanisms, with a large multigene family of glycosyltransferases (GTs) able to recognize lipophilic small molecules including hormones, secondary metabolites and both biotic and abiotic toxins in the environment (Lim and Bowles, 2004; Bowles *et al*., 2005). Glycosylation is often the final step in the biosynthesis of secondary plant products resulting in the formation of an overwhelming number of natural glucosides with numerous applications. The conjugation of structurally diverse compounds with glucose significantly enhances the water solubility necessary for compartmentalization (e.g., in the vacuole) and is thus one of the most important detoxification processes of the plant cell. One of the major groups of glucosides derives from phenolics, representing a significant part of the total glucoside pool of plants. The attachment of glucose is a rather plant-specific process and only rarely catalyzed by fungi and bacteria. In this respect, plants have remarkable biosynthetic potential, which is absent in microorganisms. The isolation and purification of such

glycosides from plant sources is, however, tedious and usually results in low yields. In organic chemistry glycosylation plays a crucial role, mainly in the synthesis of natural β-D-glycosides. The synthetic strategies available in most cases result in rather low yields, lead to side products, or deliver a mixture of α- and β-glycosides instead of the desired pure β-compounds. Nonetheless, it is this latter class of naturally occurring products that has the most widespread interest. An enzyme-catalyzed transfer of glucose to aglycones would be an alternative approach to chemical synthesis but requires the appropriate glycosyltransferases. Unfortunately, these enzymes occur in only minute amounts in plant material, are mostly unstable, and generally exhibit very limited substrate specificities. This severely limits the broad application of this enzymatic method in glycoside synthesis. Thus, simple approaches for the generation of such glycosides would be highly beneficial. Since glycosylation confers amphipathic properties on the molecule and is normally critical for biological activity, this is clearly an important area in which an effort should be invested.

1.3.1 Glycosyltransferases

Glycosyltransferases (GTs) (EC 2.4.x.y) are enzymes that catalyze the transfer of sugar moieties from donor molecules (usually nucleoside diphospho-sugars) to specific acceptor molecules that may be a lipid, a protein or a growing oligosaccharide thus forming glycosidic bonds.



Hydroxylated molecules are the most common acceptors, whilst UDP-glucose is the most common donor in the GT catalytic glycosyl-group transferring reactions, thus the name

"UGTs" for plant glycosyltransferases. Sometimes, UDP-galactose, UDP-rhamnose and UDP-glucuronic acid also act as donors in some cases. In mammals, glucuronosylation is well known to occur during Phase II conjugative drug metabolism, which is catalyzed by glucuronic acid transferases/glucuronosyltransferases (GATs). Mammalian GATs are generally considered to be physiological membrane bound enzymes (Senafi *et al*., 1994; King *et al*., 2000; Mackenzie *et al*., 2003). By contrast, plants evolved various soluble UGTs that have intrinsic sugar donor specificity:glucosyltransferase (GlcT), rhamnosyltransferase (RhaT), galactosyltransferase (GalT), arabinosyltransferase (AraT), and GAT (Miller *et al*.,1999; Sawada *et al*., 2005; Yonekura-Sakakibara *et al*., 2007, 2008). However, the physiological roles and biochemical activities of most plant UGTs remain elusive because higher-plant genomes have a high copy number of UGT genes (107 and 240 UGTs in *Arabidopsis thaliana* and grapevine [*Vitis vinifera*], respectively) (Li *et al*., 2001; Lim *et al*., 2003; Paquette *et al*., 2003; The French-Italian Public Consortium for Grapevine Genome Characterization, 2007). Moreover, the molecular structures of the sugar acceptors are extremely diverse.

For a long period of time, although glycosyltransferase activities and glycosylated products were known from a variety of plants (Knofel *et al*., 1984; Hughes and Hughes, 1994; Schneider and Schliemann, 1994), the enzymes and genes involved in the glycosylation failed to be isolated. Their roles in plant growth and development were also uncertain. Until recently, glycosyltransferases were thought to have only limited influence on the basic physiology of plants. Identification of glycosyltransferase genes and their functional characterization has changed that view. In recent years, dozens of glycosyltransferase genes have been identified, and among these, quite a few have been characterized functionally. It is now recognized that the glycosylation of low-molecular-weight compounds of plants, by adding a sugar moiety to the acceptors, usually changes acceptors in terms of their bioactivity, stability, solubility, subcellular localization and binding property to other molecules, and this possibly reduces the toxicity of endogenous and exogenous toxic substances (Bowles *et al*., 2005). Therefore, glycosyltransferases might have an important role in maintaining cell homeostasis and the regulation of plant growth, development and defense response to stress environments (Jones and Vogt 2001; Lim and Bowles, 2004). The study on glycosyltransferases and glycosylation of plant molecules has thus attracted researcher's

considerable interest, because understanding the catalytic mechanisms of the enzymes and their physiological roles would be of great significance for *in vitro* designing and synthesizing of valuable glycosides and for *in vivo* metabolic engineering of crops for important agronomic traits (Kristensen *et al*., 2005; Lim, 2005; Weis *et al*., 2008).

1.3.1.1 Observed Glycosyltransferase Fold Types

The modest degree of sequence homology within and sometimes among the various families has made the prediction of tertiary structures difficult; however, structural determinations in recent years have revealed that the catalytic domains of most glycosyltransferases display one of two fold types designated GT-A or GT-B (Coutinho PM *et al*., 2003; Bourne Y and Henrissat B (2001).

GT-A and GT-B fold types consists of two closely associated domains at least one of which contains a Rossmann-fold responsible for donor nucleotide recognition. The Rossmann fold is a nucleotide binding domain (Dodson E *et al*., 1966), and is a ubiquitous structural motif among enzymes with either GT-A or GT-B fold type where it constitutes a dominant portion of the catalytic center in a cleft between the two domains. The Rossmann fold often contains much of the limited sequence homology that is observed across many glycosyltransferase families due to a finite repertoire of donor nucleotides utilized (Campbell JA *et al*., 1997).

The GT-A type fold, Fig. 1.3a, is believed to be ancestral to enzymes with the GT-B type fold (Coutinho PM *et al*., 2003; Unligil UM *et al*., 2000). The donor usually binds the N-terminal domain's Rossmann fold. Typically one or two of this domain's β-sheets extend into the C-terminal domain, rendering definite separation of the folds difficult. The second domain is usually responsible for acceptor recognition and has greater sequence and structural variability among the different families than the nucleotide binding domain. This is presumably due to the limited number of different donor nucleotide sugars, but a vast number of different acceptor molecules. The two domains in the GT-A type fold cooperate to form the active- site cleft.

The GT-B type fold, Fig. 1.3b, has two somewhat homologous Rossmann or Rossmann-like folds in distinct domains which are usually separated by a deep wide crevice

while the enzyme is in an unliganded “open” conformation. Enzymes with the GT-B type fold have been suggested to have step-wise reaction mechanisms, where substrate binding in the “open” form can induce a conformational shift by a pair of main chain rotations to generate the “closed” form to align the nucleophile and substrates in the active site for catalysis (Coutinho PM *et al*., 2003; Unligil UM *et al*., 2000; Mulichak AM *et al*., 2003).

There have been two further fold types predicted using iterative BLAST searches and other methods (Rosen ML *et al*., 2004; Liu J *et al*., 2003; Kikuchi N *et al*., 2003). These so called GT-C and GT-D folds are thought to contain catalytic centers within transmembrane loops. As the catalytic domains themselves consists of integral-membrane proteins they are intrinsically difficult to crystallize, and there is currently no structurally characterized example of either the GT-C or GT-D fold types.

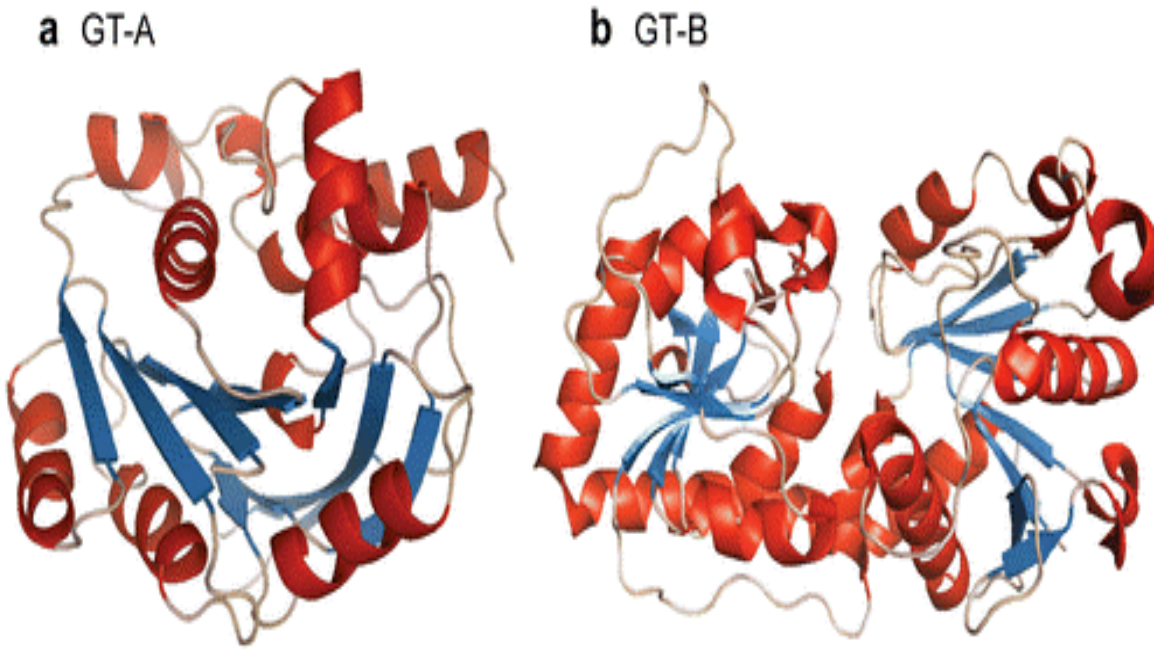


Fig. 1.3: a) GT-A-fold type with a single Rossmann fold on the right, and **b)** GT-B fold type with two Rossmann folds.

1.3.1.2 Multigene families of plant glycosyltransferases

To date, according to substrate recognition, sequence similarity and phylogenetic analysis, GTs that exist in the biosphere can be divided into 91 distinct families, and among these, family 1 has the most number of GTs which have a close relationship with plants. Substrates for GTs of family 1 are low-molecular-weight lipophilic compounds in which single or multiple glycosylations can take place at the -OH, -COOH, -NH2, -SH and C-C groups (Bowles *et al*., 2005).

In family 1 there are about 50% of members in which each contains a carboxy-terminal consensus sequence called the plant secondary product glycosyltransferase box (PSPG box) (Hughes and Hughes, 1994). This box consists of 44 amino acids close to the C-terminal part of the protein, and is believed to be involved in binding of the activated donor sugars (Fig. 1.4). The amino-terminal regions of GTs are more variable, supporting the hypothesis that the domain is involved in the recognition and binding of diverse acceptors.

Because of the completion of the whole genome sequence of *Arabidopsis thaliana*, multigene families of glycosyltransferases were first comprehensively analyzed in this plant species. In the *Arabidopsis* GT family 1, most of the GTs are UGTs and have the C-terminal consensus sequence except for three GTs (Paquette *et al*., 2003). Through a search of the *A. thaliana* genome with the PSPG motif, a very big glycosyltransferase superfamily consisting of 119 putative UGT genes was found. Further phylogenetic analysis of this superfamily showed that all of those UGTs could be divided into 14 distinct groups (Li *et al*., 2001). When comparing UGT sequences at amino acid level, it has been found that most UGTs from other plant species share a certain degree of homology with those of *Arabidopsis*. However, with some UGTs from different plant species, even though their activities and substrates are similar, their sequence similarities may be very low. For example, five UGTs responsible for the glycosylation of cytokinins were identified from *Arabidopsis* (Hou *et al*., 2004). When these sequences were compared to the cytokinin-glycosylating enzymes identified in *Zea* and *Phaseolus*, only low sequence similarity was observed, and the UGTs of *Zea* and *Phaseolus* form a unique branch on the phylogenetic tree containing *Arabidopsis* UGTs (Hou *et al*., 2004). It is likely that the *Zea* and *Phaseolus* UGTs evolved from a common ancestor distinct from those identified in *Arabidopsis*.

It has been found that the intron distribution in *Arabidopsis* UGT genes is diverse. Although some genes have one or two introns, over half of *Arabidopsis* UGT genes contain no intron. Analysis of the conserved regions and intron positions of the UGT genes of *Arabidopsis* suggests that at least ten independent intron-insertion events and either one or two intron-loss events have occurred during *Arabidopsis* UGT evolution (Ross *et al*., 2001).

UDP-sugar is the most commonly used donor for family 1 UGTs, but as for the types of monosaccharides, different UGTs use different monosaccharides. UDP-glucose (UDPGlc) is the most common sugar donor, whilst UDPrhamnose (UDP-Rha), UDP galactose (UDP-Gal), UDPxylose (UDP-Xyl) and UDP-glucuronic acid (UDPGlcUA) have also been used for some UGTs (Bowles *et al*., 2005).

The structural information of proteins could be very important for us to understand the evolution and catalytic mechanisms of family 1 GTs. Unfortunately, the protein structural information on family 1 GTs is very limited at the moment. Recently, Shao *et al*. (2005) and He *et al*. (2006) presented the crystal structures of a UDP-glucose flavonoid/triterpene glycosyltransferase (UGT) from *Medicago truncatula*. Their results show that plant glycosyltransferase contains two Rossmann folds and the acceptor binds to residues in the N-terminal half, whereas activated donor sugars bind to amino acids in the C-terminal region. In addition, their findings revealed the key residues involved in the recognition of donor substrates, and mutagenesis confirmed the roles of these key residues in enzyme activity, providing an initial structural basis for understanding the complex substrate specificity and regiospecificity underlying the glycosylation of plant molecules.

: .:: .: : . :.: :: :: ::: . : : ::. : ::

N-terminal < WAPQVEVLAHPAVGCFVTHCGWNSTLESISAGVPMVAWPFFADQ> C-terminal

Fig. 1.4: PSPG-box consensus sequence of plant glycosyltransferases. Highly conserved amino acids are indicated by two dots (identity **>** 80%) or one dot (identity **>** 50%) above the letters of amino acids. PSPG box: plant secondary product glycosyltransferase box.

1.3.1.3 Physiological roles of GTs in plants

Up to now, studies on glycosyltransferases and glycosylation towards low-molecular-weight plant compounds have been conducted mainly in *Arabidopsis* and several other plant species. Along with the availability of considerable biochemical data and genomic data on plant GTs, the analysis of the biological roles of GTs in plants has been possible by using the techniques of functional genomics, especially the strategies of gene over-expression and gene knock-out (or knock-down). Recent results obtained with functional characterization of plant GTs indicate that glycosyltransferases might play an important role in plant growth, development and interaction with the environment.

1.3.1.3.1 Taking part in hormone homeostasis

The homeostasis of hormones is crucial to plant growth, development, and adaptation responses to changes in the environment. Many mechanisms have evolved to control precisely the level of different hormones in plant cells and tissues. Glycosylation is thought to be one of these mechanisms, because glycosides of all classical hormones except ethylene have been identified in plant extracts (Creelman and Mullet, 1997; Mok and Mok, 2001; Fujioka and Yokota, 2003; Woodward and Bartel, 2005).

The bioassays of activities of hormone glycosides by exogenous application showed that the activities of hormones could generally be reduced or lost after glycosylation, although the inactivation mechanism is not very clear. It was assumed that glycosylation could alter recognition between acceptors and hormones or change properties of hormones (Kleczkowski and Schell, 1995).

At present, some studies have dealt with endogenous glycosylation of plant hormones. Jackson *et al*. (2002) cloned UGT84B1 in *Arabidopsis,* and they revealed that UGT84B1 could endogenously synthesize 1-O-indole acetyl glucose ester (IAGlc). Overexpression of that GT in *Arabidopsis* led to a phenotype of aerial parts similar to the auxin-deficient mutant as well as a root system losing geotropism (Jackson *et al*., 2002; Jackson *et al*., 2001). In another example, overexpression of UGT73C5 resulted in brassinosteroids (BR)-deficient phenotypes and reduced levels of active BRs in transgenic *Arabidopsis*, suggesting that

glucosylation of BRs reduces their bioactivities. On the other hand, silencing UGT73C5 led to a reduction of BR glucosylation, confirming that UGT73C5 is involved in the regulation of the active BR pool in plants (Poppenberger *et al*., 2005). Recently, Rodo *et al*. (2008) reported the effects of cytokinin O-glucosylation in maize. Their results showed that overexpressing ZOG1 gene (encoding a zeatin O-glucosyltransferase from *Phaseolus lunatus* L.) in the roots and leaves of the transgenic maize greatly increased the levels of zeatin-O-glucoside and produced a similar phenotype to those of cytokinin-deficient plants, leading to growth retardation and tassel seed formation. These results indicate that cytokinin O-glucosylation plays an important role in balancing cytokinin levels.

What are the effects of cytokinin N-glucosylation? It is believed that N-glycosylated cytokinins are irreversibly inactive forms (in contrast to reversibly inactive forms of O-glycosylated cytokinins), but the precise *in vivo* function of N-glycosylated cytokinins is unknown. Hou *et al*. (2004) made the first step towards answering this question. They screened the recombinant GTs from *Arabidopsis* and identified two UGTs responsible for the cytokinin N-glucosylation for the first time. The identification of cytokinin N-glucosylating GTs paved the way for the functional analysis of cytokinin N- glucosylation. Further, other GT genes capable of glycosylating ABA (Xu *et al*., 2002; Lim *et al*., 2005a) and SA (Taguchi *et al*., 2001; Lim *et al*., 2002) have also been identified. It is expected that these genes will further promote the study of the relationship between glycosylation and hormone homeostasis

1.3.1.3.2 Taking part in defense response

There are several examples that implicate the role of glycosylation of secondary metabolites in plant defense response to biotic stress. For instance, TOGTs are tobacco glycosyltransferases with the highest *in vitro* enzyme activity towards hydroxycoumarin, scopoletin, and hydroxycinnamic acids. It was found that reducing the expression of TOGTs in transgenic tobacco substantially decreased the levels of scopoletin glucoside and simultaneously impaired the resistance to Tobacco Mosaic Virus (TMV) (Chong *et al*., 2002). However, overexpression of TOGT1 in transgenic tobacco caused enhanced resistance to Potato Virus Y (Matros and Mock, 2004). Both of these studies suggested that

glycosyltransferases might play an important role in glycosylating scopoletin and enhancing plant resistance to pathogens by some unknown mechanisms.

Another example is the C-3 oligosaccharide chain of saponins. It is believed that the oligosaccharide chain consisting of Glc, Gal, arabinopyranose (Ara), GlcUA, Xyl or Rha attached to the C-3 of saponins may be crucial for resisting fungal pathogens. Removing these sugar residues often leads to loss of bioactivity. Interestingly, fungal pathogens could produce hydrolases to attack the C-3 oligosaccharide chain of saponins in order to detoxify saponins after their invasion. For example, the oat root infecting pathogen *Gaeumannomyces graminis* produces avenacinase, a kind of *β*-glucosidase, to detoxify the triterpenoid avenicin saponins by removing the Glc from the C-3 oligosaccharide chain. Tomato pathogens could also detoxify steroidal glycoalkaloids in the same way (Sandrock and VanEtten, 1998).

1.3.1.3.3 Taking part in detoxification

As described above, fungal pathogens could detoxify plant glycoside molecules by hydrolyzing the glycosidic bonds. However, in the battle between plants and pathogens, plants seem to take an opposite strategy, namely forming glycosidic bonds to detoxify the toxicity of pathogens. The evidence comes from the detoxification of trichothecene deoxynivalenol (DON) which is produced by the fungus *Fusarium*, one of the most familiar fungi of cereal species such as wheat, barley and maize. DON is not only harmful to plant growth but also to human health, and it is considered as a virulence factor in fungal pathogenesis. *In vitro* experiments indicate that the recombinant protein UGT73C5, a putative glycosyltransferase of *Arabidopsis*, could catalyze DON to form DON-3-O-glucoside. Glycosylated DON could lose toxicity, and overexpressing UGT73C5 in transgenic *Arabidopsis* could enhance the resistance of transgenics to DON (Poppenberger *et al*., 2003), suggesting a role of glycosylation in the detoxification of plants.

Further, GTs can also detoxify exogenous chemical compounds such as herbicides, insecticides, pollutants, as well as xenobiotics (Lim *et al*., 2002; Loutre *et al*., 2003). For example, overexpression of either GT BX8 or GT BX9 in *Arabidopsis* reduced the toxic effects of 2,4-dihydroxy-1,4-benzoxazin-3-one (DIBOA) and 2,4 dihydroxy-7-methoxy-1,4-benzoxazin-3-one (DIMBOA) applied on transgenic plants (Von Rad *et al*., 2001). Several

UGTs of *Arabidopsis* have also been confirmed to be involved in the detoxification of the xenobiotic 2,4,5-trichlorophenol (TCP) (Messner *et al*., 2003) or the pollutant 3,4-dichloroaniline (DCA) (Loutre *et al*., 2003).

1.3.1.3.4 Taking part in biosynthesis and storage of secondary metabolites

In the long-term process of evolution, plants have developed many metabolic pathways to synthesize a vast multitude of secondary metabolites for environmental adaptation. Many studies have shown that glycosyltransferases and glycosylation reactions could be involved in the biosynthesis, modification, transportation and storage of secondary metabolites.

In the lignin biosynthesis pathway, lignin monomers (coumaryl, coniferyl and sinapyl alcohols) need to be translocated from the cytosol to the cell wall, where they are polymerized into lignin. The glucosides of lignin monomers have been considered as the transport forms. The recombinant glycosyltransferases, UGT72E2 and UGT72E3, of *Arabidopsis* display strong activities of 4-O-glucosyltransferase to phenylpropanoids, in particular, the coniferyl and sinapyl alcohols (Lim *et al*., 2005b), suggesting their involvement in the biosynthesis of lignin. Down-regulating the expression of these glucosyltransferases in transgenic *Arabidopsis* plants could severely reduce the glucoside levels of lignin monomers (Lanot *et al*., 2006), but the relationship of monolignol glycosylation and lignin synthesis has not been clarified yet.

Flavonols, monoterpenoids and hydroxybenzoic acids usually accumulate as both aglycones and glycosides. In the case of flavonols, Jones *et al*. (2003) identified two GTs (UGT73C6 and UGT78D1) and Tohge *et al*. (2005) identified three GTs (UGT79B1, UGT75C1 and UGT78D2) involved in the biosynthesis of flavonol glycosides in *Arabidopsis*. Glycosylation is usually the last step of flavonol biosynthesis metabolism, probably indicating a requirement of stabilization, reactivity or translocation. Multiple additions of sugar moieties to a given compound, in parallel or in chains, give rise to a broad spectrum of secondary metabolites, thus contributing to their unique properties. For example, one single flavonol, quercetin, has 300 different glycosides naturally occurring in plants. The biosynthesis of some flavonol or anthocyanin glycosides plays an important role in the occurrence of fruit and flower color, fruit flavor and the protection of plants from damage due to ultraviolet radiation

(Bowles *et al*., 2005). Glycosylation also strongly affects the bioavailability of these dietary compounds, some of which display antioxidant or anticancer activity.

The biosynthesis of steviol glycosides is another example. *Stevia rebaudiana* accumulates a mixture of at least eight different steviol glycosides, kinds of compounds that are unique in the plant world because of their intense sweetness (300 times sweeter than sugar) and high concentration in leaf tissue. The majority of the glycosides are formed by four glucosylation reactions that start with steviol, producing the mixture of mono-, di-, tri- and tetraglycosides. The tri-glycoside stevioside and the tetraglycoside rebaudioside A represent the majority of the steviol glycosides present in *S. rebaudiana* leaves. The first step of stevioside biosynthesis takes place in plastids. After glucosylation of the C-4 carboxyl position of steviolbioside, the glycosides are transported into the vacuoles of leaf cells. The glycosylation of the C-4 carboxyl position appears to be critical for glycoside transport into the vacuole, because glycoside accumulation only occurs following that step. Using functional genomics strategy, Richman *et al*. (2005) found three glucosyltransferases involved in the synthesis of the major sweet glucosides of *S. rebaudiana*, and their in vitro regioselective glucosylating activities towards steviol were confirmed by the recombinant enzymes.

1.3.1.4. Potential applications of Plant GTs

As more and more plant glycosyltransferases are identified and their biological roles are revealed, increasing attention is paid to glycosyltransferases due to their potentials in practical applications. At the first stage, it appears that the use of plant glycosyltransferases is mainly in the *in vitro* biochemical synthesis of some valuable glycoconjugates. For example, Karim and Hashinaga (2002) cloned a GT from *pummelo* and by using enzyme immobilization, they have successfully converted the limonoids, the bitter components of lemon juice, into tasteless glycosides. ABA can exist as two enantiomers, i.e. (±)-ABA. Both enantiomers occur in chemical preparations. It was found that UGT71B6 of *Arabidopsis* enantioselectively glucosylated only (+)-ABA, thereby it could be used to separate (+)-ABA from (–)-ABA, offering an alternative to chemical synthesis for the production of pure (+)-ABA (Lim *et al*., 2005a). Weis et al. (2008) exchanged the N-terminal domain and the C-terminal domain of two *Arabidopsis* GTs, 71C1 and 71C3, to construct chimeric mutants. This engineering of

proteins provides a basis for creating novel GTs and for improving the substrate affinity of the enzymes. Enzymatic synthesis of glycosides has many advantages over conventional chemical methods, for example, synthesizing stereospecific glycosides without the use of hazardous chemicals as blocking and deblocking reagents, large scale production in fermentation by using the microbial whole-cell systems, with fewer synthetic steps and lower cost. Therefore, the potential application of glycosyltransferases as biocatalysts in *in vitro* glycoconjugates synthesis has attracted researchers' considerable interest in recent years. On this topic, Lim (2005) has given a detailed discussion in his review. The study of enzymatic activities and biological roles of plant GTs will also provide the basis for applications in crop improvement in the near future. Several GT-encoding genes, such as those involved in hormone homeostasis, defense response and detoxification may be suitable candidates for insertion into a variety of plants with the aim of improving crop plants. For example, the recent identification of UGTs acting on ABA, auxins, cytokinins and BRs opens new perspectives on how to manipulate the levels of active hormones in plants and on how to control growth and development of plants. Other applications may involve the metabolic engineering of GTs to detoxify pesticide residues and pollutants so as to increase food security, enhance plant resistance to biotic and abiotic stress and increase the content in food of glycosides with antioxidant or anticancer activities. There is also the possibility that over-expressing specific GTs in plants could change flower colors or fruit flavors.

1.3.1.5. Approaches for GT identification

Due to the development of new biotechnologies, the studies on plant GTs have advanced much in recent years. Several approaches, including through biochemistry, bioinformatics, molecular biology, and genetics, have been successfully used to clone, identify and analyze genes that encode plant GTs. More often however, identifying GT genes requires an integration of several methods. Therefore, the following approaches are divided basically according to their differences in the first steps of identifying GTs.

1.3.1.5.1. Classical biochemical methods

The isolation and purification of the target proteins directly from plants is usually the first step in these methods when identifying GT genes. However, these methods often encounter difficulties in getting target proteins with high purity because of the complexity of protein components in plants. Once the target proteins are purified, their enzyme activities toward glycosylation of specific substrates can be investigated and the corresponding genes can be cloned by the derived nucleotide sequences from the amino acids. By classical biochemical methods, plant GT gene from *Arabidopsis*, GT72B1, has been cloned and the gene product GT72B1 has been confirmed to be capable of glycosylating inorganic pollutant 3, 4-dichloroaniline *in vitro* (Loutre *et al*., 2003). Another plant GT gene encoding limonoid UDP-glucosyltransferase in Citrus has also been cloned by this method (Frydman *et al*., 2004).

1.3.1.5.2. Bioinformatics combined with biochemistry

The development of genomics and bioinformatics greatly facilitates the identification of plant GTs. As mentioned above, through the bioinformatical analysis of many GTs identified from plants, the PSPG conserved box was proposed. This consensus sequence provides a good starting point for searching new glycosyltransferases from a database. After the related expressed sequence tags (ESTs), cDNAs or genes are determined by means of bioinformatics, a series of investigations such as cloning the full length of cDNA, expressing and purifying recombinant proteins *in vitro*, analyzing the substrate specificity and enzyme activities of recombinant proteins can be sequentially performed. Using these methods, several glycosyltransferases for the plant hormone auxins, cytokinins, brassinosteroids (BR), abscisic acid (ABA) and salicylic acid (SA), have been successfully identified (Jackson *et al*., 2001; Lim *et al*., 2001; Hou *et al*., 2004; Poppenberger *et al*., 2005).

1.3.1.5.3. Genetic methods

The use of mutants is very important for gene identification and functional analysis. However, the lack of mutants of plant glycosyltransferases makes it difficult to use this method to

identify GT genes currently. Quiel and Bender (2003) luckily cloned a GT gene, UGT74F2, by screening blue fluorescent mutants of *Arabidopsis*. UGT74F2 could catalyze the glycosylation of anthranilate to form a blue fluorescent anthranilate glucose ester. Different from other methods, the identification of UGT74F2 was merely an accident, because the experiment was supposed to have no relation to GT in the beginning.

1.3.1.5.4. Molecular biological methods

Because of the rapid development of molecular biology, many methods could be used in the identification of plant GTs. For example, based on the conserved amino acid sequence of GTs, degenerate primers can be designed and RT-PCR can be carried out using plant RNA material to clone putative GTs. Using molecular biological methods, Moraga *et al*. (2004) cloned a saffron apocarotenoid crocetin GT, which provides the foundation for producing crocin in heterologous systems.

1.4. Rationale of the Thesis

The reputation of *Bacopa monniera* as a nervine tonic is well known in Indian traditional medicine. Recent pharmacological studies indeed confirmed the activity. It was also confirmed that the activity was due to the glycosides present in the alcoholic extract of the plant. Because of the presence of a large number of glycosidic constituents in the extract as an intimate mixture, several groups of workers could isolate and characterize only a part of the constituents chemically. At the same time purifying and characterizing GTs directly from the tissue has led to a wealth of biochemical information. The labile nature of the enzymes has made it difficult to purify the enzyme in quantities necessary for structure-function analysis. Establishment of glycosyltransferase gene (GTs) in the recombinant expression systems and characterization of the resultant protein has been shown to be an effective means to overcome this limitation.

In view of the increasing interest on this herb drug, this study involved isolation and characterization of GTs which are responsible for producing several medicinally important glycosides, since till now no study has been done on glycosyltransferase gene from *B.*

monniera. The study of this gene will help in understanding its role and specificity towards different aglycone moieties. The present work deals with the molecular study of GT gene from *B. monniera* and its over-expression in suitable heterologous system.

Chapter: 2

Materials & Methods

2. Materials and methods

This chapter includes general laboratory techniques routinely followed in the studies. Other important specific methodologies followed are discussed separately in the respective chapters.

2.1 Plant material

2.1.1 *Bacopa monniera*

Bacopa monniera (Linn.) Pennel (*Scrophulariaceae*) was collected from the Department of Horticulture, National Chemical Laboratory, Pune.

For genomic DNA isolation, young *Bacopa* shoots were harvested, washed thoroughly with water, dried and then crushed with liquid nitrogen. For RNA isolation, leaves of mature healthy plant were harvested as and when required. The tissue was ground using liquid nitrogen and RNA was isolated for normal cDNA preparations, RACE and QRT-PCR.

For Real time PCR experiments, *Bacopa monniera* plants were maintained as potted plant in the garden under moderate sunlight. They were watered at regular intervals. Shoots were collected from these healthy and disease free plants and were used for initiation of the cultures. Shoots from these *in vitro* grown plants were used for QRT-PCR experiments.

2.1.2 Surface sterilization

The shoots were cleaned under tap water and treated with savlon (10% v/v) for one minute followed by thorough washing with sterile distilled water for 2-3 times. They were treated with 70% alcohol in sterile flask for few seconds in laminar air flow cabinet. Alcohol was quickly decanted and shoots were washed thoroughly with sterile distilled water. After that washed shoots were treated with 0.01% $HgCl_2$ solution for 30 seconds with frequent shaking and rinsed with sterile distilled water several times to remove all the traces of $HgCl_2$. Explants were prepared by dissecting the shoots into pieces with 1-2 axillary buds.

2.1.3 Media used

All media were based on Murashige and Skoog basal medium (Murashige & Skoog, 1962). The pH of the media was adjusted to 5.8 prior to autoclaving (121 $^{\circ}$C and 1.05 kg/cm^2 pressure for 20 minutes). Initiation medium was MS basal medium supplemented with 0.5 mg/L 6-benzyl amino purine and 3% sucrose, solidified with 1% agar. Proliferation medium was plain MS medium with 3% sucrose.

2.1.4 Inoculation and Incubation

The surface sterilized explants were inoculated on initiation medium under aseptic conditions. Cultures were incubated at the temperature 25±1°C under 16 hour photoperiod at 11.7 µmol/m^2/sec light intensity/8 h dark cycles for about 20 days. The proliferating multiple adventitious shoots were separated and re-inoculated on proliferation medium for further growth.

2.1.5 Stress treatment

In vitro proliferated shoots were used for study of different stress treatments like salt, methyl jasmonate, salicylic acid, mannose, cold and heat by transferring them to MS basal medium supplemented with 50 mM NaCl, 20 µM methyl jasmonate, 30 µM salicylic acid and 2% mannitol. Approximately 0.1-0.2 gm of tissue was inoculated on the medium containing salt, methyl jasmonate, salicylic acid and mannose. The cultures were further incubated under the same conditions as described earlier. Sample tissue was collected from these cultures after 2 months. In case of cold (-20 °C) and heat (37 °C) treatment, cultures were reinoculated without any supplement in MS basal medium and grown till 2 months followed by heat treatment for one day and cold shock treatment for 2 hours. Tissue was harvested from heat and cold treated plants and were used for QRT-PCR.

For immuno-cytolocalization studies cultured plants were transferred to pots for hardening in greenhouse so as to get enough material for roots since roots of the *Bacopa* plant were very thin.

2.2 Glassware

Glassware used in all the experiments was 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 ^{o}C and 1.05 kg/cm^2 for 1 hour.

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), Bioworld (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. 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 and standards for enzyme assays i.e. naringenin, diadzein, genistein, apigenin, catechin, myricetin, isorhamnetin, luteolin, hesperetin, curcumin, Kaempferol, diosgenin, capsaicin, vanilic acid, salicylic acid and all the respective glycosides were also obtained from Sigma-Aldrich (USA) and Chromadex (USA).

2.5 Equipments: See appendix, Table 2.1 List of Equipments used.

2.6 Buffers and solutions

2.6.1 Buffers and solutions for DNA electrophoresis

See appendix, Table 2.2 Buffers and Solutions for DNA Electrophoresis.

2.6.2 Buffers and solutions for gDNA isolation

See appendix, Table 2.3 Buffers and Solutions for g-DNA isolation.

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

See appendix, Table 2.4 Stock solutions for *E. coli* transformation and selection.

2.6.4 Buffers and solutions for plasmid isolation (Alkaline lysis method)

See appendix, Table 2.5 Buffers and solutions for plasmid isolation (Alkaline lysis method).

2.6.5 Buffers and solutions for gel electrophoresis (PAGE)

See appendix, Table 2.6 Buffers and Solutions for gel electrophoresis (native/SDS-PAGE).

2.6.6 Buffers and solutions for protein purification under native conditions

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

2.6.7 Buffers and solutions for protein extraction under denaturing conditions

See appendix, Table 2.8 Buffers and solutions for protein extraction under denaturing conditions.

2.6.8 Buffers and solution for the glycosyltransferase (GT) enzyme assay

Glycosyltransferase activity was checked using 20 mM Tris-HCl, pH- 7.0, 5 mM glycosyl donor (UDP-Glucose, UDP-Galactose and UDP-Glucuronic acid), 600 µM glycosyl acceptors (see appendix Table 2.9) and enzyme.

2.6.9 Buffers and solutions for ELISA/ Immuno-cytolocalization

See appendix, Table 2.10 Buffers and Solutions used for ELISA, Immuno-cytolocalization.

2.6.10 Component of Murasighe and Skoog media/different inducing media and hormones

See appendix, Table 2.11 Component of Murasighe and Skoog media.

2.6.11 Different media used for studies

See appendix, Table 2.12 Different media use for bacterial studies.

2.7 Host cells

See appendix, Table 2.13 Bacterial cell lines used.

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(s).

2.8.2 Bacterial transformation

2.8.2.1 Preparation of competent cells using TB buffer

A single colony of *E. coli* XL1 Blue was inoculated in 5 mL of LB medium containing tetracycline (12.5 mg/L) and grown overnight at 37 ^{o}C in incubator shaker at 200 rpm. Overnight grown culture (500 µL) was added to 50 mL of SOB medium and grown for 23 hours at 37 ^{o}C in incubator shaker at 200 rpm. Cells were kept on ice for 10 min and harvested by centrifugation at 5,000 rpm for 10 min at 4 ^{o}C. The cell pellet was suspended in 5 mL ice-cold TB buffer containing 7% DMSO. This was then dispensed into aliquots of 100 µL in 1.5 mL eppendorf tubes, frozen in liquid nitrogen and stored at -80 ^{o}C.

2.8.2.2 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 A_{600} was 0.5. 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.

2.8.2.3 *E. coli* transformation

The competent *E. coli* cells were transformed according to Sambrook *et al.*, (1989). DNA (~50 ng in 10 µL or less) was added gently 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 2 min (heat shock) and immediately kept back on ice. LB broth (800 µL) was added to each tube and further incubated at 37 ºC for 1 h. Cells were pelleted by centrifugation at 5,000 rpm for 5 min. The supernatant media was discarded and pellet was dissolved in 100 µL 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	Stocks	Final concentration
IPTG	200 mg/mL in sterile distilled water.	40 µg/mL
X-gal	20 mg/mL in dimethylformamide.	40 µg/mL

2.8.3 Colony PCR for screening recombinant colonies after bacterial transformation

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 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 used in 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 gene 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 cloned gene of interest.

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 530 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 TEG buffer (Tab: 2.5) by vigorous pipetting, 200 µL of Soln. II (Tab: 2.5) was added, mixed by inversion till the solution becomes clear, 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 12,000 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, vortexed 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 at 60 °C and 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, 70% 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 1.0 g 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 polyvinylpolypyrrolidone (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: isoamyl alcohol (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) per 100 µL DNA solution 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's recommendations.

2.8.4.4 Extraction and purification of DNA from agarose gels

The restriction digested DNA or PCR amplified products were run on an agarose gel in 1X TAE buffer. The gel was stained with ethidium bromide (0.5 µg/mL) and viewed using a hand held long wavelength UV illuminator. The fragment of interest was excised from the gel and weighed. A 100 µg gel slice was transferred to a 1.5 mL microcentrifuge tube and 300 µL Buffer DE-A (Axygen^{TM} 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. Wash buffer 1 (500 µL) (provided by Axygen) was added and centrifuged at 12,000 g for 30 s, filtrate was discarded. Wash buffer 2 (700 µL) was added and spin at 12,000 g for 30 s, filtrate was discarded. It was repeated again with wash buffer 2. One 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 was stored at -20 ^{o}C and was used for subsequent reactions.

2.8.4.5 Total RNA Isolation

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, 1.0 mL TRIzol reagent was added and mixed thoroughly using a vortimix. Chloroform: isoamyl alcohol (300 µL) was added and mixed thoroughly using vortimix. The tubes were centrifuged at 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 checked by measuring OD at 260/280 nm by Agilent 2100 Bioanalyser 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 2100 (Agilent). The amount of RNA was in the range of 1 to 3 mg. Appropriate amount of OBB Buffer and Oligotex suspension was added according to manufacturer's 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 micro-centrifuge 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 micro-centrifuge 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 acids concentration

DNA concentration was determined by measurement of the absorption at 260 nm. Nanodrop was used to determine the concentration of 1:50 diluted RNA or DNA samples. 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 2100 (Agilent) 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 thereafter, analyses the various parameters in one go within 30 min.

2.8.4.8 cDNA first strand 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 ^{o}C 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 and concentration	**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 micro centrifuge 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 micro centrifuge tube on ice.

Reagent and concentration	**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 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 and concentration	Volume
Sterile deionized water	10.7 µL
Template (100 ng/µL)	1.0 µL
Forward primer (10 pmol)	1.0 µL
Reverse primer (10 pmol)	1.0 µL
dNTPs (0.2 mM)	4.0 µL
10 x Buffer (Mg^{2+} 1.5 mM)	2.0 µL
Taq DNA Polymerase (1 U/µL)	0.3 µL
Total volume	**20.0 µL**

PCR cycle conditions

95 ºC - 5 min
95 ºC - 1 min
45-65 ºC - 45-90 s
72 ºC - 1.5 min.

35 cycles
(Annealing temperature (Ta) was dependent on melting Temperature of the primer (Tm); Ta = Tm - 5 ºC)

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.

GeneRacer RACE Ready cDNA Synthesis

The pre-requisite 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 and concentration	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
Total Volume	**10.0 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 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 and concentration	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	**10.0 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 the mRNA. The reaction of ligating the RNA Oligo to decapped mRNA is as follows:

Reagent and concentration	Volume
Decapped RNA	7.0 µL
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	**11.0 µL**

7 µL of dephosphorylated, decapped RNA was added to the tube containing the prealiquoted, 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, 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 and concentration	Volume
5 X RT Buffer	4.0 µL
SuperScript™ III RT (200 U/µL)	1.0 µL

0.1 M DTT	1.0 µL
RNaseOut™ (40 U/µL)	1.0 µL
Sterile water	2.0 µL
Total Volume	**20.0 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 and concentration	**5′ RACE**	**3′ RACE**
GeneRacer 5′ Primer 10 µM	3.0 µL	-
Reverse GSP 10 µM	1.0 µL	-
GeneRacer™ 3′ Primer 10 µM	-	3.0 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 or 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
95 ºC	5 min	1 cycle
95 ºC	1 min	35 cycles
50 ºC	0.45 min	
72 ºC / 1 kb	1.30 min	
72 ºC / 1 kb	5 min	1 cycle

Nested PCR

Nested PCR was done to increase the specificity and sensitivity of RACE products for the 5′ or 3′ ends of gene. 0.5 µL of the original amplification reaction was used as template for nested PCR. Reactions were set up as described below:

Reagent	5′ RACE	3′ RACE
Gene Racer 5′Nested Primer10 µM	1.0 µL	-
Reverse Nested GSP 10 µM	1.0 µL	-
Gene Racer 3′ Nested Primer 10 µM	-	1.0 µL
Forward Nested GSP 10 µM	-	1.0 µL
Primary PCR Product	0.5 µL	0.5 µL
10 X High Fidelity PCR Buffer	2.0 µL	2.0 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	0.5 µL	0.5 µL
Sterile Water	13.5 µL	13.5 µL
Total Volume	20.0 µL	20.0 µL

Following cycling condition was used for the nested PCR reactions,

Temperature	Time	Cycles
95 ºC	5 min	1 cycle
95 ºC	1 min	35 cycles
55-58 ºC	0.45 min	
72 ºC / 1 kb	1.30 min	
72 ºC / 1 kb	5 min	1 cycle

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 band was 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). In the present study the SYBR Brilliant® II QPCR Master Mix (Stratagene, USA) was used. This kit supports quantitative amplification and detection systems. The kit supports PCR amplifications and detection of a variety of DNA targets, including genomic DNA, plasmid DNA, and cDNA. The SYBR 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.

The fluorescent dye SYBR Green I in the master mix enables the analysis of many different targets without having to synthesize target-specific labeled probes.

Real time PCR can be divided into four major phases: the linear ground phase, early exponential phase, log-linear (also known as exponential) phase, and plateau phase. During the linear ground phase (usually the first 10–15 cycles), PCR is just beginning, and fluorescence emission at each cycle has not yet risen above background. Baseline fluorescence is calculated at this time. At the early exponential phase, the amount of fluorescence has reached a threshold where it is significantly higher (usually 10 times the standard deviation of the baseline) than background levels. The cycle at which this occurs is known as Ct in ABI Prism® literature (Applied Biosystems, Foster City, CA, USA) or crossing point (CP) in LightCycler® literature (Roche Applied Science, Indianapolis, IN, USA). This value is representative of the starting copy number in the original template and is used to calculate experimental results. During the log-linear phase, PCR reaches its optimal amplification period with the PCR product doubling after every cycle in ideal reaction conditions. Finally, the plateau stage is reached when reaction components become limited and the fluorescence intensity is no longer useful for data calculation. In general, lower Ct value indicates higher initial copies.

2.8.5.1 QRT-PCR considerations

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 best concentrations of the upstream and downstream primers are not always of equal molarity. In this study, 100 nM was considered optimum. Reaction was standardized in such a way that there was no primer dimer formation. Acquisition of real-time data generated by SYBR Green 1 was done as recommended by the instrument manufacturer. Data collection was either at the annealing step (3- step cycling protocol) or extension step of each cycle.

Magnesium chloride concentration in the PCR reaction mix affects the specificity of the PCR primers and probe hybridization. The SYBR Green Brilliant® II QPCR Master Mix

kit contains $MgCl_2$ at a concentration of 5.5 mM (in the 1X solution), which is suitable for most targets.

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. The SYBR Green Brilliant® II QPCR Master Mix kit contains reference dye which is suitable for most targets.

2.8.5.2 Preparing the reactions

Real time PCR model Mx 3000P (Stratagene, USA) was used in the present study. The experimental reactions were prepared by adding the following components in order.

2.8.5.3 Reaction Mixture

Nuclease-free PCR-grade H_2O to adjust the final volume to 25 µL (including experimental DNA)

Reagent	Volume
2X master mix	12.5 µL
Upstream primer (optimized concentration)	1.0 µL
Downstream primer (optimized concentration)	1.0 µL
Total	25.0 µL

The reactions were mixed without creating bubbles and 1 µL of experimental genomic DNA, cDNA or plasmid DNA was added to each experimental reaction. Reaction was mixed gently and briefly spun.

2.8.5.4 PCR cycling programs

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

Two step cycling Protocol,

Temperature	Time	No. of cycles
95 ºC	10 min	1
95 ºC	30 sec	40
55- 57 ºC	30 sec	
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.

2.9 Expression and purification of recombinant Protein

Heterologous expression and activity of glycosyltransferase (*GT*) gene was attempted in *E. coli* (BL21) host cell. The transformants were grown at 37 ºC in Luria–Bertani medium containing kanamycin (50 μg/mL). A single isolated bacterial colony from freshly streaked plates (grown on LB agar medium containing 50 μg/mL kanamycin) was used to inoculate 5 mL liquid LB medium containing the same concentrations of the antibiotic. Culture was grown overnight with shaking at 225 rpm at 37 °C. One mL aliquot of each culture was used to inoculate 100 mL liquid LB containing 50 μg/mL kanamycin. Once the cultures reached A_{600} 0.4 - 0.5, recombinant protein expression was induced by the addition of isopropyl β-D-thiogalactopyranoside (IPTG), and the culture was grown for 4 to 6 h at 37 °C with shaking at 150 rpm. Cells were harvested by centrifugation at 6,000 rpm for 10 min at 4 °C. Pellets were resuspended in 6.25 mL lysis buffer. Cells were disrupted by sonication for 5 mins at 70 amplitude on an ultrasonic liquid processor, XL 2000 model (MESONIX). $MgSO_4$, final

concentration of 10 mM and lysozyme, final concentration 100 µg/mL was added to the disrupted cells and kept at 37 ^{o}C for ½ h. It was centrifuged at 10,000 rpm for 10 minute and supernatant was saved as lysate and pellet was resuspended in 2 mL sonication buffer. Suspension was again sonicated for 1 min at 70 amplitude to disrupt the inclusion bodies and the disrupted inclusion bodies were dissolved in 3 mL of dispersion buffer and an aliquot of 20 µL checked on SDS PAGE (Chapter 2: section 2. 9.3 and Table: 2.6) to check heterologous expression of *GT* genes.

2.9.1 Protein isolation from inclusion body

Cells were isolated by centrifugation and resuspended in 100 mM Tris-HCl buffer, pH 8, 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 ultrasonic liquid processor, XL 2000 model (MESONIX) for 5 min (5 seconds on and 5 sec off) at 70 amplitude. 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.

2.9.2 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 2.6.7). The washing efficiency may be improved by the addition of 20 mM imidazole and histidine-tagged proteins are then usually eluted with 150-300 mM imidazole. (Chapter 2: section 2.6.7 and Tab 2.8). The 6x His-tagged protein was eluted in 4 aliquots of elution buffer, 1 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. 9.3 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. 9.3.1 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. 9.3.2 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. 9.3.3 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. 9.3.4 Loading and running the polyacrylamide gel

Once well, have 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 polyacrylamide gel.

2. 9.3.5 Coomassie Blue staining of the gel

Coomassie blue staining solution: 45 mL Methanol, 10 mL acetic acid, 45 mL de-ionized water and 0.25% Coomassie blue (R 250). After running the gel it was transferred directly to a tray containing Coomassie blue staining solution and was kept for three to four hours at room temperature on rocker. Staining solution was poured off and de-staining solution (Table 2.5) was poured in. De-staining step was repeated two-three times till clear bands appeared.

2.10 Raising polyclonal antibody against *GT* in rabbit

The purified recombinant *Bacopa* BMGT1 protein was used for raising polyclonal antibodies in rabbit (New Zealand White).

2.10.1 Pre-treatment of serum

Rabbit immune serum (stored at -70 ºC), containing polyclonal antibodies against the antigen, was thawed overnight at 4 ºC. The Serum was kept at 55 °C for one hour and the immune serum was centrifuged at 12,000 g for 15 min at 4 ºC. Aliquots according to the requirements were made and stored at -70 ºC. Prior to use again antiserum was thawed overnight at 4 ºC.

2.10.2 Determination of titre of antibodies and ELISA

Buffers

1X Phosphate buffer saline (PBS): 8 gm NaCl; 1.44 g Na_2HPO_4; 0.24 g KCl; 0.200 g K_2HPO_4 in 1 L De-ionized water.

Wash Buffer (PBST): 1X PBS, 0.05% Tween 20 and 0.1% BSA

Antibody diluting buffer: 1X PBS and 0.25% BSA

Blocking reagent: 1% BSA in PBS

Substrate diluting buffer: 200 mM Tris-HCl, pH 9.5 and 10 mM $MgCl_2$

All reagents were prepared in sterile milliQ water.

ELISA was performed to determine the titre of first, second and third bleed of rabbit serum. Equal quantity of antigen *i.e* 100 ng was coated in triplicates on ELISA plates and kept overnight at 4 °C. Next morning, the plate was washed with 250 µL of PBST, three times for 5 min and 300 µL of blocking reagent was added. The plate was wrapped in aluminium foil and kept at 37 °C for 2 h. ELISA plate was washed again as described earlier and challenged to different dilution of serum such as, 1:10000, 1:20000, 1:30000, 1:40000, 1:60000, 1:80000 and 1:100000. All dilutions were in triplicates and plate was kept at 37 °C for 2 h again. Plate was washed as described earlier with PBST buffer and secondary antibodies (anti goat IgG against rabbit IgG) tagged with alkaline phosphatase was added to a dilution of 1:20000. Plate was incubated at 37 °C for two more hours. After 2 hours of incubation plate was washed with 250 µL of PBST, three times for 5 min and 100 µL of 1 mg/mL substrate (*p*- Nitro phenyl phosphate) was added and incubated for 45 min. Reaction was stopped by adding 10

mM EDTA. Once the antibody titre was determined then a fixed dilution of antibodies was used for rest of the experiments.

2.10.3 Histology and Immuno-cytolocalization

Free hand transverse sections of root, leaf and stem 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%, 50% ethanol and 0.5X SSC (Table 2.10) 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 1: 5000 dilution of primary antibody 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, 1:10000 dilution of 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 (Axioplan 2 Carl Zeiss) and microscopic image captured (AxioCamMRc5).

2.11 MALDI MS/MS

For MALDI analysis 10% SDS-PAGE was run for both recombinant glycosyltransferase proteins. Gel was then stained with coomassie Blue. Expected stained protein bands were excised from the gel. The gel pieces were then destained by destaining solution (50% acetonitrile / 50% 50 mM NH_4HCO_3) till colour was gone. Gel was then dehydrated by treating with 100% acetonitrile. After dehydration acetonitrile was completely removed by evaporating briefly in speedvac till noticeably shrunken and white. Then gel pieces were dissolved in 10 mM DTT in 100 mM NH_4HCO_3 and proteins were reduced for 45-60 min. at 56 ^{o} C. Cooled to room temperature, DTT solution removed and 55 mM iodoacetate in 100 mM NH_4HCO_3 was added. This mixture was vortexed, spun briefly and incubated for 45 min in dark place at room temperature. This was followed by iodoacetamide removal and gel pieces were washed with 100 mM NH_4HCO_3 for 5 min. Again gel pieces were washed twice with 50% acetonitrile / 50% 50 mM NH_4HCO_3) and dehydrate with 100% acetonitrile as mentioned above. Then enough trypsin solution was added to cover the gel pieces (usually around 20 µL) and the gel pieces were rehydrated at 4 ^{o} C for 30 min in buffer containing 50 mM NH_4HCO_3 and trypsin. Spun briefly and more NH_4HCO_3 was added to cover gel pieces (typically another 25 µL). This was followed by overnight digestion at 37 ^{o} C. The digested solution (supernatant) was transferred into clean 1.5 mL eppendorf tube. Add 50 % acetonitrile / 2% formic acid solution to the gel pieces which was incubated and vortexed for 20 min. This was spun and sonnicated for 5 mins in a water bath with no heat. Supernatant was then removed and combined with initial digestion solution (supernatant). Extracted digests was then vortexed, evaporated to reduce to 5-10 µL. The remaining 5-10 µL was spun at 14k rpm for at least 10 min. to remove any microparticulates. The supernatant was carefully transferred to a fresh 1.5 mL eppendorf tube. The sample was then ready for loading onto MALDI MS/MS. Simultaneously MALDI plate was also washed inorder to remove the particulate matter deposited if any. Then sample and the MALDI matrix (α cyano 4 hydroxy cinnamic acid) in the proportion of 2:1 ratio were loaded on to the MALDI plate. This was

followed by reading of MALDI plate with laser energy of about 280 Volts using MALDSNYPT equipment (Waters).

2.12 *GT* enzyme assay

The standard reaction mixture (500 µL) consisted of 5 mM glycosyl acceptor (see appendix Table 2.9), 600 µM glycosyl donor, 20 mM Tris-HCl buffer, pH 7.0, and enzyme. After a 10 minutes preincubation of the mixture without the enzyme at 30 ^{o} C, the reaction was initiated by addition of the enzyme. After incubation at 30 ^{o}C for 3 hours, the reaction was extracted thrice by the addition of equal volume of ethyl acetate. The extracted reaction was then completely vacuum dried and dissolved in methanol. The substrates and glycosylated products were analysed by LC-MS and HPLC.

2.12.1 LC-MS

LC-MS analysis was performed using a Q-TOF Premier mass spectrometer (Waters) outfitted with an electrospray ion source operated in the V-Optics negative mode. A Develosil C18 column was used for LC on a Shimadzu LC-20AD HPLC. For LC-MS analysis solvent system used was 20% water and 80% (v/v) acetonitrile (CH_3CN) for 5 min.

2.12.2 HPLC

Analysis of the substrates and their glycosylated products was done by using reversed-phase HPLC (Perkin Elmer series 200) on a supelco C18 column (5µm, 25cm x 4.6 mm). For Genistein, mobile phase consisted of 50 mM Tris buffer (pH 8.0). 20 µL of sample dissolved in methanol was injected in HPLC column and was programmed as follows: 10% acetonitrile for 0.5 min, 40% acetonitrile for 10 min and 60% acetonitrile for 5 min at a flow rate of 1 mL/min. A dual wavelength recorder set at 270 nm and 340 nm was used to detect the compounds eluting from the column.

For naringenin, mobile phase consisted of milliQ water. 20 µL of sample dissolved in methanol was injected in HPLC column and was programmed as follows: 20% acetonitrile for

0.5 min, 40% acetonitrile for 5 min, 70% acetonitrile for 5 min, 90% acetonitrile for 5 min and 20% acetonitrile for 5 min at a flow rate of 1 mL/min. A dual wavelength recorder set at 270 nm and 340 nm was used to detect the compounds eluting from the column.

For kaempferol mobile phase consisted of milliQ water. 20 µL of sample dissolved in methanol was injected in HPLC column and was programmed as follows: 10% acetonitrile for 0.5 min, 10% acetonitrile for 5 min, 30% acetonitrile for 5 min, 75% acetonitrile for 5 min and 95% acetonitrile for 5 min at a flow rate of 1 mL/min. Compounds eluted from the column were detected at 420 nm.

2.13 MTT cytotoxicity test on human cancer cell lines

In order to check the cytotoxicity effects of substrate (genistein) and the product (genistein 4-glycoside also known as sophoricoside) on MCF-7 (human breast cancer cell line) and THP-1 (human monocytes) cancer cells, 5 mg/mL stock solution of MTT (sigma) was prepared in cell culture media. Solution was filter sterilized using 0.22 µm filter (Millipore). Solution was stored at 4 ºC, protecting it from light. The cells were maintained at 37 ºC in a humidified atmosphere containing 5% CO_2. Approximately 1 $\times 10^5$ cells per mL were seeded in MEM (minimum essential medium) medium containing 10% heat inactivated fetal bovine serum (FBS). 100 µl of this cell suspension was added to each well of sterile 96-well plate and compound (genistein and sophoricoside) was added in the wells at the time of inoculation in a dose dependent manner and was kept at 37 ºC for 72 hours. All the samples were taken in triplicates. This was followed by addition of 10 µL of MTT dye solution (5 mg/mL) and was further incubated for another 1 hour at 37 ºC. Then 200 µL of isopropanol was added to the culture which was further incubated at 37 ºC for 4 hours. Then reading was taken at the absorbance of 490 nm in a xMark microplate spectrophotometer (Biorad) to examine their effect on proliferation.

Chapter 3

Isolation, Cloning and Characterization of *B. monniera* Glycosyltransferase gene(s)

3.1 Introduction

Plants have the ability to produce a large variety of secondary metabolites, such as terpenoids, phenylpropanoids, flavonoids and alkaloids, most of which are potentially useful not only as pharmaceuticals and their leads but also as pigments, flavors, insecticides, etc. In particular, flavonoids which share a common 15-carbon polyphenolic skeleton are one of the most characteristic classes of compounds in higher plants. They are distributed ubiquitously in plants, including in mosses and ferns, and are among the most intensely studied secondary product groups, with >6000 known species (Parker, 2001). Many of them play important roles as pigments in flowers and fruits, UV-B protectants, signaling molecules between plants and microbes, and regulators of auxin transport (Dixon *et al*., 1995; Harborne and Williams, 2000). The flavonoids also have antioxidant, antiallergenic, anti-inflammatory, anti-conceptive and neuropharmacological activities, thus contributing to human health (Lee *et al*., 2007; Bischoff, 2008). There is increasing evidence to suggest that flavonoids, in particular those belonging to the class of flavonols and isoflavonoids are potentially health-protecting components in the human diet as a result of their high antioxidant capacity (Dugas *et al*., 2000; Ross and Kasum, 2002). Therefore, flavonoids may offer protection against major diseases such as coronary heart diseases and cancer (Hertog & Hollman, 1996; Steinmetz & Potter, 1996). Although there are many differentially substituted flavonoid aglycones, a vast array of structural variations within the flavonoids come from sequential modifications of the flavonoid core structure by glycosylation, acylation, and methylation. Flavonoids are stabilized by glycosylation. Glycosylation is often the last step in the biosynthesis of natural products in plants which is catalysed by the enzyme glycosyltransferase.

Glycosyltransferase (GTs) has been already discussed in detail in Chapter one of this thesis. The present chapter deals in detail regarding isolation, cloning and characterization of two flavanoid glycosyltransferase gene(s) from *Bacopa monniera.* There is considerable interest in understanding the structural basis for glycosylation of natural products which will facilitate future metabolic engineering of these compounds to improve the health of plants (resistance to disease and herbivory), animals (via improved forage quality), and humans (through drug discovery).

Isolation of a gene is the pre-requisite for its thorough study or characterization. To isolate any gene of interest generally there are two major approaches as follows:-

1. Genomic/cDNA library screening- It is the classical method of gene isolation in which genomic/cDNA libraries are screened with homologous or heterologous probe. After 2-3 rounds of screenings, the plaques showing strong positive signals are excised and the DNA fragment sequenced.
2. PCR- It is one of the most popular approaches of gene isolation because of its simplicity and rapidity. Forward and reverse primers are designed on the basis of available sequences in the database. PCR is performed using these primers and genomic/cDNA as template. The amplicon is sequenced to confirm its identity.

In the present study, PCR based approach was followed to fish out the *GT* gene(s) from *B. monniera*.

3.2 Materials and methods

3.2.1 Genomic DNA extraction

Genomic DNA was extracted from *B. monniera* using the method given by Lodhi *et al*. (Chapter 2, Section 2.8.4.2).

3.2.2 Restriction digestion of plasmid DNA

Plasmid DNA restriction digestion was set up as per manufacturers (Promega, USA; NEB, UK; Amersham, USA) recommendations (Chapter 2, Section 2.8.4.3).

3.2.3 Bacterial strains and plasmids used in the study

Escherichia coli XL-1 blue (Stratagene, USA)
E. coli -XL1-Blue MRF (Stratagene, USA)
pGEM-T Easy Vector Cloning vector (Promega, USA)

3.2.4 RNA isolation and cDNA first strand synthesis

Total RNA was isolated from *B. monniera* leaves (Chapter 2, Section 2.8.4.5); mRNA was purified from total RNA using oligotex dT resins (Chapter 2, Section 2.8.4.6) and cDNA Reverse Transcription (RT) 1st strand was synthesized (Chapter 2, Section 2.8.4.8).

3.2.5 Polymerase Chain Reaction

PCR amplification was done using cDNA 1st strand or genomic DNA as template (Chapter 2, Section 2.8.4.9). Amplified PCR products were eluted from Agarose gel (Chapter 2, Section 2.8.4.4). The eluted PCR products were used for ligation into suitable vector.

3.2.6 Transformation and selection

The ligation mixture was used for transformation of *E. coli* XL1-Blue, host cell line. Putative transformants were selected using Blue-white screening by adding X-gal and IPTG. Colony PCR (Chapter 2, Section 2.8.3) was done to screen the putative transformants. Plasmid was isolated from clones showing positive colony PCR results. The plasmid was restriction digested to check for the insert size and the insert was further sequenced to confirm its identity.

3.2.7 Bioinformatic analysis

The *GT* gene/nucleotide sequences available in the NCBI GeneBank database were aligned and multiple sets of primers were designed from the conserved regions. Nucleotide and amino acid sequence analysis was done using software pDRAW 32, ClustalX 1.8 and online bioinformatics analysis facility available at www.justbio.com, www.expasy.org and www.ncbi.nlm.nih.gov. Multiple alignments of the amino acid sequences were carried out with the Clustal W1.8 program (http://www.ebi.ac.uk/clustalw/). The evolutionary history was inferred using the Neighbor-Joining method (Saitou and Nei, 1987). The bootstrap consensus tree inferred from 1000 replicates is taken to represent the evolutionary history of

the taxa analyzed (Felsenstein, 1985). Branches corresponding to partitions reproduced in less than 50% bootstrap replicates are collapsed. The evolutionary distances were computed using the Poisson correction method (Zuckerkandl and Pauling, 1965) and are in the units of the number of amino acid substitutions per site. All positions containing gaps and missing data were eliminated from the dataset (Complete deletion option). Phylogenetic analyses were conducted in MEGA4 (Tamura, 2007).

3.2.8 Rapid Amplification of cDNA Ends (RACE)

RACE (Chapter 2, Section 2.8.4.10) was done to fish out the full length gene with its 5' and 3' UTRs.

3.2.9 Sequencing

DNA sequencing was performed with the ABI Prism Big-Dye Terminator Cycle Sequencing Kit on the ABI Prism 3730 DNA analyzer (Applied Biosystems) at GenomeBio Biotech Pvt. Ltd., Pune, Maharashtra, India.

3.3 Results and discussion

3.3.1 PCR based approach for the isolation of *GT* gene(s).

3.3.1.1 Multiple sequence alignment of nucleotide sequences of reported glycosyltransferase gene(s) from different family members.

GT gene nucleotide sequences available at NCBI GenBank database were aligned using Clustal W (1.8) multiple sequence alignment software. Forward and reverse primers were designed from regions showing highest homology. The conserved regions selected for primer designing has been shown high-lighted (red & blue (Fig 3.1 & 3.2).

```
gi|62241066      ATGGGTTCCATTGGTGCTGAATTTACAAAGCCACATGCAGTTTGCATACCATATCCCGCC
gi|62241064      ATGGGTTCCATTGGTGCTGAATTAACAAAGCCACATGCAGTTTGCATACCATATCCCGCC
                 ********************** *************************************
```

NTMF1 forward Primer

```
gi|62241066      CAAGGCCATATTAACCCCATGTTAAAGCTAGCCAAAATCCTTCATCACAAAGGCTTTCAC
gi|62241064      CAAGGCCATATTAACCCCATGTTAAAGCTAGCCAAAATCCTTCATCACAAAGGCTTTCAC
                 ************************************************************

gi|62241066      ATCACTTTTGTCAATACTGAATTTAACCACAGACGTCTGCTTAAATCTCGTGGCCCTGAT
gi|62241064      ATCACTTTTGTCAATACTGAATTTAACCACCGACGTCTCCTTAAATCTCGTGGCCCTGAT
                 ****************************** ******* ********************

gi|62241066      TCTCTCAAGGGTCTTTCTTCTTTCCGTTTTGAGACAATTCCTGATGGACTTCCGCCATGT
gi|62241064      TCTCTCAAGGGTCTTTCTTCTTTCCGTTTTGAGACCATTCCTGATGGACTTCCGCCATGT
                 ********************************** ************************

gi|62241066      GATGCAGATGCCACACAAGATATACCTTCTTTGTGTGAATCTACAACCAATACTTGCTTG
gi|62241064      GAGGCAGATGCCACACAAGATATACCTTCTTTGTGTGAATCTACAACCAATACTTGCTTG
                 ** ************************************************************

gi|62241066      GGTCCTTTTAGGGATCTTCTTGCGAAACTCAATGATACTAACACATCTAACGTGCCACCC
gi|62241064      GCTCCTTTTAGGGATCTTCTTGCGAAACTCAATGATACTAACACATCTAACGTGCCACCC
                 * ************************************************************

gi|62241066      GTTTCGTGCATCATCTCAGATGGTGTCATGAGCTTCACCTTAGCCGCTGCACAAGAATTG
gi|62241064      GTTTCGTGCATCGTCTCGGATGGTGTCATGAGCTTCACCTTAGCCGCTGCACAAGAATTG
                 ************ **** ******************************************

gi|62241066      GGAGTCCCTGAAGTTCTGTTTTGGACCACTAGTGCTTGTGGTTTCTTAGGTTACATGCAT
gi|62241064      GGAGTCCCTGAAGTTCTGTTTTGGACCACTAGTGCTTGTGGTTTCTTAGGTTACATGCAT
                 ************************************************************

gi|62241066      TATTACAAGGTTATTGAAAAAGGATACGCTCCACTTAAAGATGCGAGTGACTTGACAAAT
gi|62241064      TACTGCAAGGTTATTGAAAAAGGATATGCTCCACTTAAAGATGCGAGTGACTTGACAAAT
                 ** * ********************** **********************************

gi|62241066      GGATACCTAGAGACAACATTGGATTTTATACCATGCATGAAAGACGTACGTTTAAGGGAT
gi|62241064      GGATACCTAGAGACAACATTGGATTTTATACCAGGCATGAAAGACGTACGTTTAAGGGAT
                 ********************************** ***************************

gi|62241066      CTTCCAAGTTTCTTGAGAACTACAAATCCAGATGAATTCATGATCAAATTTGTCCTCCAA
gi|62241064      CTTCCAAGTTTCTTGAGAACTACAAATCCAGATGAATTCATGATCAAATTTGTCCTCCAA
                 ************************************************************

gi|62241066      GAAACAGAGAGAGCAAGAAAGGCTTCTGCAATTATCCTCAACACATATGAAACACTAGAG
gi|62241064      GAAACAGAGAGAGCAAGAAAGGCTTCTGCAATTATCCTCAACACATTTGAAACACTAGAG
                 ********************************************** **************

gi|62241066      GCTGAAGTTCTTGAATCGCTCCGAAATCTTCTTCCTCCAGTCTACCCCATTGGGCCCTTG
gi|62241064      GCTGAAGTTCTTGAATCGCTCCGAAATCTTCTTCCTCCAGTCTACCCCATAGGGCCCTTG
                 ************************************************* **********

gi|62241066      CATTTTCTAGTGAAACATGTTGATGATGAGAATTTGAAGGGACTTAGATCCAGCCTTTGG
gi|62241064      CATTTTCTAGTGAAACATGTTGATGATGAGAATTTGAAGGGACTTAGATCCAGCCTTTGG
                 ************************************************************

gi|62241066      AAAGAGGAACCAGAGTGTATACAATGGCTTGATACCAAAGAACCAAATTCTGTTGTTTAT
gi|62241064      AAAGAGGAACCAGAGTGTATACAATGGCTTGATACCAAAGAACCAAATTCTGTTGTTTAT
                 ************************************************************

gi|62241066      GTTAACTTTGGAAGCATTACTGTTATGACTCCTAATCAACTTATTGAATTTGCTTGGGGA
gi|62241064      GTTAACTTTGGAAGCATTACTGTTATGACTCCTAATCAGCTTATTGAGTTTGCTTGGGGA
                 ************************************ ******** ************

gi|62241066      CTTGCAAACAGCCAACAATCATTCTTATGGATCATAAGACCTGATATTGTTTCAGGTGAT
gi|62241064      CTTGCAAACAGCCAGCAAACATTCTTATGGATCATAAGACCTGATATTGTTTCAGGTGAT
                 ************** *** ****************************************

gi|62241066      GCATCGATTCTTCCCCCCGAATTCGTGGAAGAAACGAAGAAGAGAGGTATGCTTGCTAGT
```

```
gi|62241064      GCATCGATTCTTCCACCCGAATTCGTGGAAGAAACGAAGAACAGAGGTATGCTTGCTAGT
                 ************** ************************* ******************

gi|62241066      TGGTGTTCACAAGAAGAAGTACTTAGTCACCCTGCAATAGGAGGATTCTTGACTCACAGT
gi|62241064      TGGTGTTCACAAGAAGAAGTACTTAGTCACCCTGCAATAGTAGGATTCTTGACTCACAGT
                 ************************************** ********************

                 NTMR2 reverse primer
gi|62241066      GGATGGAATTCGACACTCGAAAGTATAAGCAGTGGGGTGCCTATGATTTGCTGGCCATTT
gi|62241064      GGATGGAATTCGACACTCGAAAGTATAAGCAGTGGGGTGCCTATGATTTGCTGGCCATTT
                 ************************************************************

gi|62241066      TTCGCTGAACAGCAAACAAATTGTTGGTTTTCCGTCACTAAATGGGATGTTGGAATGGAG
gi|62241064      TTCGCTGAACAGCAAACAAATTGTTGGTTTTCCGTCACTAAATGGGATGTTGGAATGGAG
                 ************************************************************

gi|62241066      ATTGACTGTGATGTGAAGAGGGATGAAGTGGAAAGCCTTGTAAGGGAATTGATGGTTGGG
gi|62241064      ATTGACAGTGATGTGAAGAGAGATGAAGTGGAAAGCCTTGTAAGGGAATTGATGGTTGGG
                 ****** ************* ************************************

gi|62241066      GGAAAAGGCAAAAAGATGAAGAAAAAGGCAATGGAATGGAAGGAATTGGCTGAAGCATCT
gi|62241064      GGAAAAGGCAAAAAGATGAAGAAAAAGGCAATGGAATGGAAGGAATTGGCTGAAGCATCT
                 ************************************************************

gi|62241066      GCTAAAGAACATTCAGGGTCATCTTATGTGAACATTGAGAAGGTGGTCAATGATATTCTT
gi|62241064      GCTAAAGAACATTCAGGGTCATCTTATGTGAACATTGAAAAGTTGGTCAATGATATTCTT
                 ************************************ *** ******************

gi|62241066      CTTTCGTCCAAACATTAA
gi|62241064      CTTTCATCCAAACATTAA
                 ***** ************
```

Fig. 3.1: Multiple sequence alignment of Glycosyltransferase (*GT*) nucleotide sequences of *Solanaceae* member *Nicotiana tabacum.* Highlighted are conserved regions considered for primer synthesis.

Forward primer NTMF1- 5’ TCC CGC CCA AGG CCA TAT TAA CC 3’

Reverse primer NTMR2- 5’ ACT TTC GAG TGT CGA ATT CCA TCC AC 3’

```
>Linaria vulgaris      ATGGGAGAAGAATACAAGAAAACACACACAATAGTCTTTCACACTTCAGAAGAACACCTC
>Antirrhinum majus     ATGGGAGAAGAATACAAGAAAACACACACAATAGTCTTTCACACTTCAGAAGAACACCTC
>Antirrhinum majus     ------------------ATGGAAGATACCATCGTATTTTACACTCCAAGCGATCACAGT
                                          *    * * ** ** ** *** ***** **   ** ***

                                          PTF1 forward primer
>Linaria vulgaris      AACTCTTCAATAGCCCTTGCAAAGTTCATAACCAAACACCACTCTTCAATCTCCATCACT
>Antirrhinum majus     AACTCTTCAATAGCCCTTGCAAAGTTCATAACCAAACACCACTCTTCAATCTCCATCACT
>Antirrhinum majus     CAACCCACAATAGCGTTGGCAAAGTTCATCAGCAAACACCACCCTTCCATCTCCATGACA
                        *  *  *******  * *********** * ********** **** ******** **

>Linaria vulgaris      ATCATCAGCACTGCCCCCGCCGAATCTTCTGAAGTGGCCAAAATTATTAATAATCCGTCA
>Antirrhinum majus     ATCATCAGCACTGCCCCCGCCGAATCTTCTGAAGTGGCCAAAATTATTAATAATCCGTCA
>Antirrhinum majus     ATCATCAGCACCGCCGCATTCCCTTCGTCCGCAGCGGT---G-CTGCCTAAAA-----CA
                       *********** *** *   *   ** ** * ** **       *    * **     **

>Linaria vulgaris      ATAACTTACCGCGGCCTCACCGCGGTAGCGCTCCCTGAAAATCTCACCAGTAACATTAAT
>Antirrhinum majus     ATAACTTACCGCGGCCTCACCGCGGTAGCGCTCCCTGAAAATCTCACCAGCAACATTAAT
>Antirrhinum majus     ATAAGTTACCACCCCCTCCCCGCCGTGCCCATGCCCCCGAACCTCTCCTCCAAT------
                       **** ***** *  **** **** **  *  * **    ** *** **   **
```

```
>Linaria vulgaris       AAAAACCCCGTCGAACTTTTCTTCGAAATCCCTCGTCTACAAAACGCCAACCTTCGAGAG
>Antirrhinum majus      AAAAACCCCGTCGAACTTTTCTTCGAAATCCCTCGTCTACAAAACGCCAACCTTGGAGAG
>Antirrhinum majus      ------CCCGTGGAATTCCTCTTCGAAATCCCCCGACTCCACAACACTAAACTCCGCGAA
                              ***** *** *  ************* ** ** ** *** * ** **  * **

>Linaria vulgaris       GCTTTACTAGATATTTCGCGAAAATCCGATATCAAAGCATTAATCATCGATTTCTTCTGC
>Antirrhinum majus      GCTTTACTAGATATTTCGCGAAAATCCGATATCAAAGCATTAATCATCGATTTCTTCTGC
>Antirrhinum majus      GCACTCGAAAGAATCTCCGAGACATCAAAGATCAAGGCGTTGGTTATCGATTTCTTTTGC
                        **  *   *   ** **    * ***  * ***** ** **  * *********** ***

>Linaria vulgaris       AATGCGGCATTTGAAGTATCCACCAGCATGAACATACCCACTTACTTCGACGTCAGTGGC
>Antirrhinum majus      AATGCGGCATTTGAAGTATCCACCAGCATGAACATACCCACTTACTTCGACGTCAGTGGC
>Antirrhinum majus      AACTCCGCTTTCGAAGTTTCCAGGAGCTTGAACATTCCGACATTCTTCGAAGCCAGCCTC
                        **  * ** ** ***** ****  *** ******* ** ** * ****** * ***   *

>Linaria vulgaris       GGCGCTTTTCTCCTCTGCACGTTTCTCCACCACCCGACACTACACCAAACTGTTCGTGGA
>Antirrhinum majus      GGCGCTTTTCTCCTCTGCACGTTTCTCCACCACCCGACACTACACCAAACTGTTCGTGGA
>Antirrhinum majus      GGCGCGTCCGGGCTCTGCGAGTTTCTCTACCACCCGACATTTCACAAAACCGTCCCCGGA
                        ***** *     ******  ******* *********** * *** **** ** *  ***

>Linaria vulgaris       GACATTGCGGATTTGAACGATTCTGTTGAGATGCCCGGGTTCCCATTGATTCACTCCTCT
>Antirrhinum majus      GACATTGCGGATTTGAACGATTCTGTTGAGATGCCCGGGTTCCCATTAATTCACTCCTCT
>Antirrhinum majus      GACATCGCGGACTTCAACGACTTTCTTGAAATCCCGGGGTGCCCTCCGCTTCACTCGGCT
                        ***** ***** ** ***** * * **** ** ** **** ***     *******  **

>Linaria vulgaris       GATTTACCAATGAGTTTGTTTTATCGTAAGACTAATGTTTACAAACACTTTCTAGACACT
>Antirrhinum majus      GATTTACCAATGAGTTTGTTTTATCGTAAGAGTAATGTTTACAAACACTTTCTAGACACT
>Antirrhinum majus      GATGTCCCTAAGGGTTTGTTCCGACGCAAGACTATTGCTTACAAACACTTCCTCGACACT
                        *** * ** * * *******    ** **** ** ** ************ ** ******

>Linaria vulgaris       TCCTTAAACATGCGCAAATCGAGTGGGATACTCGTGAACACGTTTGTTGCGCTCGAGTTT
>Antirrhinum majus      TCCTTAAACATGCGCAAATCGAGTGGGATACTCGTGAACACGTTTGTTGCGCTCGAGTTT
>Antirrhinum majus      GCCAACAACATGCGGATGTCGAGTGGAATCCTCTTACACGCGTTCGATGCGCTTGAATAC
                         **   ******** *  ******** ** *** *  ** **** * ****** ** *

>Linaria vulgaris       CGAGCTAAGGAAGCTTTGTCCAACGGTTTGT---------ACGGTCCAACTCCGCCTCTT
>Antirrhinum majus      CGAGCTAAGGAAGCTTTGTCCAACGGTTTGT---------ACGGTCCAACTCCGCCTGTT
>Antirrhinum majus      CGAGCTAAGGAAGCTTTGTCCAACGGCTTGTGCAATCCGGACGGGCCAACTCCGCCTGTT
                        ************************** ****         **** ************ **

>Linaria vulgaris       TATTTACTTTCACATACAATTGCCGAACCC----CACGACACTAAAGTGTTGGTAAAC--
>Antirrhinum majus      TATTTACTTTCACATACAATTGCCGAACCC----CACGACACTAAAGTGTTGGTAAAC--
>Antirrhinum majus      TACTTTGTTTCGCCTACTGTGGCTGAAACATTGGCATACAGGGAAAACACCGCCGCCTTG
                        ** **  **** * ***  * ** *** *     **       ***     *

>Linaria vulgaris       CAACACGAATGCCTATCATGGCTTGATTTGCAGCCTAGTAAAAGCGTGATTTTCCTTTGT
>Antirrhinum majus      CAACACGACTGCCTATCATGGCTTGATTTGCAGCCTAGTAAAAGCGTGATTTTCCTTTGT
>Antirrhinum majus      CGGCATGAATGCTTGACGTGGCTTGATTTGCAGCCTGATAAAAGCGTTATCTTCCTTTGT
                        *  ** ** *** *  * ******************  ********* ** *********

>Linaria vulgaris       TTCGGAAGAAGAGGAGCGTTCTCAGCACAACAGTTGAAAGAAATTGCGATAGGGTTGGAG
>Antirrhinum majus      TTCGGAAGAAGAGGAGCGTTCTCAGCACAACAGTTGAAAGAAATTGCCATAGGGTTGGAG
>Antirrhinum majus      TTTGGAAGGAGGGGAACATTCTCCATGCAACAGTTGCATGAAATTGCTGTCGGTCTTGAA
                        ** ***** ** *** * *****    ********* * ********  * **  * **

>Linaria vulgaris       AAGAGTGGATGTCGATTTCTTTGGTTGGCCCGCATTTC----ACCGGAGATGG-------
>Antirrhinum majus      AAGAGTGGATGTCGATTTCTTTGGTTGGCCCGCATTTC----ACCGGAGATGG-------
>Antirrhinum majus      CGGAGCGGGCGAAGATTTCTCTGGGCCATCCGCAGTAGTGGGGCAGGGAACGGTGAGCCT
                          *** **  *  ******* ***     ***** *       * **  * **

>Linaria vulgaris       -ACTTAAATGCGCTTCTGCCGGAGGGTTTTCTATCGAGAACTAAAGGAGTAGGGTTTGTG
>Antirrhinum majus      -ACTTAAATGCGCTTCTGCCGGAGGGTTTTTTATCGAGAACTAAAGGAGTAGGGTTTGTG
>Antirrhinum majus      GACTTGAGCGTGGTGCTGCCGGAGGGTTTCTTGGAGAGAACCAAAGATATTGGGCTGGTG
                         **** *  * * * **************  *   ****** ****   * *** * ***

>Linaria vulgaris       ACAAACACATGGGTGCCGCAAAAAGAGGTGTTGAGTCATGATGCAGTGGGGGGGTTTGTG
>Antirrhinum majus      ACAAACACATGGGTGCCGCAGAAAGAGGTGTTGAGTCATGATGCAGCGGGGGGGTTTGTG
>Antirrhinum majus      ATAACGACATGGGCGCCGCAGAAAGAGGTGTTAAGCCATGTGGCCGTGTGTGGATTTGTG
                        * **  ******* ****** *********** ** ****  ** * * * ** ******
```

```
>Linaria vulgaris       ACTCATTGCGGGTGGAGTTCGGTTCTTGAAGCGCTGTCGTTCGGTGTCCCGATGATTGGT
>Antirrhinum majus      ACTCATTGTGGGTGGAATTCTGTTCTTGAAGCGCTGTCGTTCGGTGTCCCGATGATTGGT
>Antirrhinum majus      ACGCACTGCGGCTGGAACTCAGTTCTCGAGGCGGTGTCGTTTGGGGTTCCGATGATTGGG
                        ** ** ** ** ****  ** ***** ** *** ******* ** ** ***********

                                                                    PTR4 reverse primer
>Linaria vulgaris       TGGCCGTTGTACGCAGAGCAGAGGATCAATAGGGTGTTCATGGTGGAGGAAATAAAGGTG
>Antirrhinum majus      TGGCCGTTGTACGCAGAGCAGAGGATCAATAGGGTGTTCATGGTGGAGGAAATAAAGGTG
>Antirrhinum majus      TGGCCGCTGTACGCAGAGCAGAGGATGAATCGGGTGTTTATGGTGGAGGAAATAAAGGTG
                        ****** ****************** *** ******* *********************

>Linaria vulgaris       GCGCTGCCATTGGATGAGGAAG---ATGGATTTGTGACGGCGATGGAGTTGGAGAAGCGC
>Antirrhinum majus      GCACTGCCATTGGATGAGGAAG---ATGGATTTGTGACGGCGATGGAGTTGGAGAAGCGC
>Antirrhinum majus      GCATTGCCTTTGGAGGAGGAGGCGGATGGGTTGGTGAGGGCGACAGAATTGGAGAAGCGG
                        **  **** ***** ***** *   **** ** **** *****  ** ***********

>Linaria vulgaris       GTCAGGGAGTTGATGGAGTCGGTAAAGGGGAAAGAAGTGAAGCGCCGTGTGGCGGAATTG
>Antirrhinum majus      GTCAGGGAGTTGATGGAGTCGGTAAAGGGGAAAGAAGTGAAGCGCCGTGTGGCGGAATTG
>Antirrhinum majus      GTGAGAGAGTTGACCGAGTCCGTGAGGGGAAAAGCGGTAAGCCGGCGGGTGGAGGAAATG
                        ** ** *******  ***** ** * *** ****  ** *  ** ** **** **** **

>Linaria vulgaris       AAAATCTCTACAAAGGCAGCCGTGAGTAAAGGTGGATCGTCCTTGGCTTCTTTGGAGAAG
>Antirrhinum majus      AAAATCTCTACAAAGGCAGCCGTGAGTAAAGGTGGATCGTCCTTGGTTGCTTTGGAGAAG
>Antirrhinum majus      AGACTCTCGGCAGAGAAGGCCGTGAGCAAGGGTGGAACGTCGCTGATTGCATTGGAGAAA
                        * * ****  ** **   ******** ** ****** ****  **  * * ********

>Linaria vulgaris       TTCATCAACTCGGTCACTCGTTAA
>Antirrhinum majus      TTCATCAACTCGGTCACTCGTTAA
>Antirrhinum majus      TTCATGGACTCGATTACTCTATAA
                        *****  ***** * ****  ***
```

Fig. 3.2: Multiple sequence alignment of Glycosyltransferase (*GT*) nucleotide sequences of *Plantaginaceae* members i.e, *Linaria vulgaris* and *Antirrhinum majus.* Highlighted are conserved regions considered for primer synthesis.

Forward primer PTF1- *5' GCA AAG TTC ATC AGC AAA CAC CAC 3'*

Reverse primer PTR4- 5' GCC ACC TTT ATT TCC TCC ACC A 3'

3.3.1.2 Genomic DNA extraction

Genomic DNA was isolated (Fig 3.3) by using the protocol of Lodhi *et al.* (1994) as mentioned in Chapter 2, Section 2.8.4.2. The concentration of the gDNA was measured by spectrophotometer as mentioned in Chapter 2, Section 2.8.4.7. PCR was performed using 100 ng of good quality of gDNA as the template. All possible primer combinations and cycling parameters were tried. None of the reactions gave a positive result.

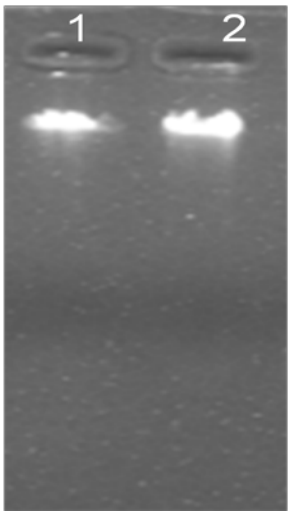


Fig. 3.3: *B. monniera* genomic DNA (lane 1 & 2) resolved on 0.8% agarose gel.

3.3.1.3 Isolation of total RNA from *B. monniera*, its cDNA synthesis and PCR

RNase free environment was created and maintained as described, the glassware and plasticware were treated overnight with DEPC (0.1% in water) and autoclaved. The pestle and mortar were also DEPC treated and then baked at 300 ºC for 6 h.

Total RNA (Fig.3.4) was isolated from leaf tissue of mature *B. monniera*. The quantity and integrity of the isolates was checked by Agilent 2100 Bioanalyser system and also by agarose gel electrophoresis. First strand cDNA was synthesized using 1 µg of total RNA by AMV reverse transcriptase based ImPromII™ Reverse Transcription System (Promega, USA) (as mentioned in Chapter 2, Section 2.8.4.8).

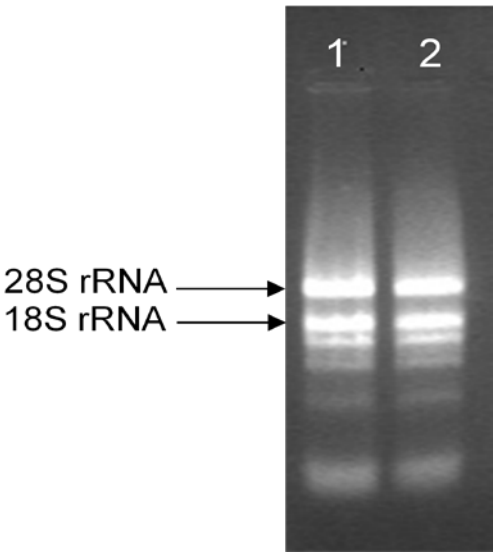


Fig. 3.4: Total RNA isolated from leaf tissue (lanes 1 to 2) of *B. monniera*, resolved on 1% denaturing gel. Arrows indicates both 28S rRNA and 18S rRNA bands

3.3.1.4 PCR amplification of partial cDNA fragment of *GT* gene from *Bacopa monniera*

Two Primer sets i.e. NTMF1-NTMR2 and PTF1-PTR4 (as mentioned in section 3.3.1.1) were used to amplify ~1 kb and ~1.1 kb fragments of *GT* gene using cDNA 1st strand as a template (Fig. 3.5 A & B). The fragments were cloned in pGEM-T easy vector (Fig. 3.6) and confirmed by restriction analysis (Fig. 3.7 A & B) and sequencing (Fig. 3.8 A & B). These partial clones will be referred to as GT1 and GT2 in the subsequent discussion. The partial clone of GT1 showed maximum homology (49%) at amino acid level with the *Lamiaceae* member, *Lycium barbarum* where as partial clone of GT2 showed maximum homology (72%) at amino acid level with *Sesamum indicum* of *Lamiaceae* family.

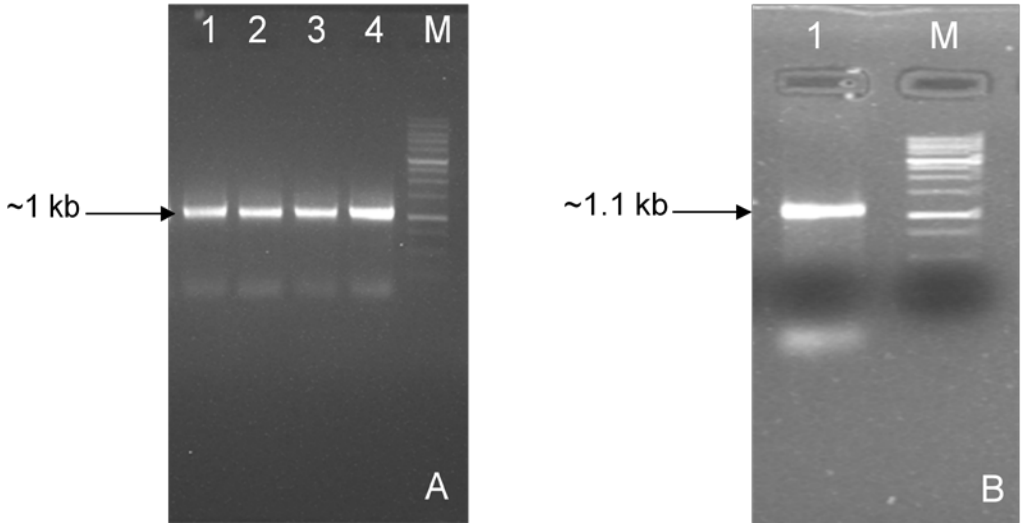


Fig. 3.5 A: PCR amplification product of ~1 kb fragment of GT1 gene separated on 1% agarose gel. Lane 1-4: Amplified product and lane M- marker. **B:** PCR Amplicon of ~1.1 kb of GT2 gene on 1% agarose gel. Lane 1- amplified product and lane M- marker.

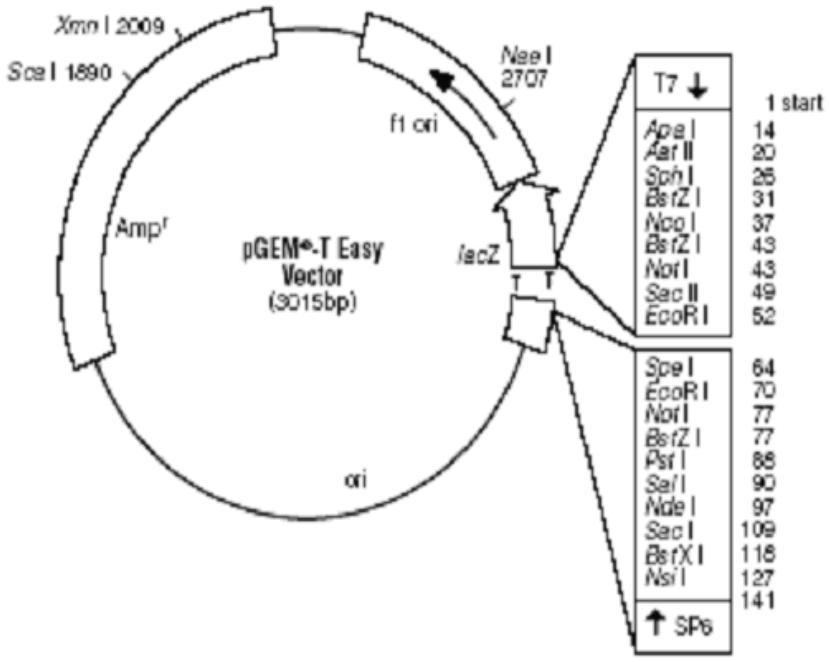


Fig. 3.6 Map of pGEM-T Easy vector.

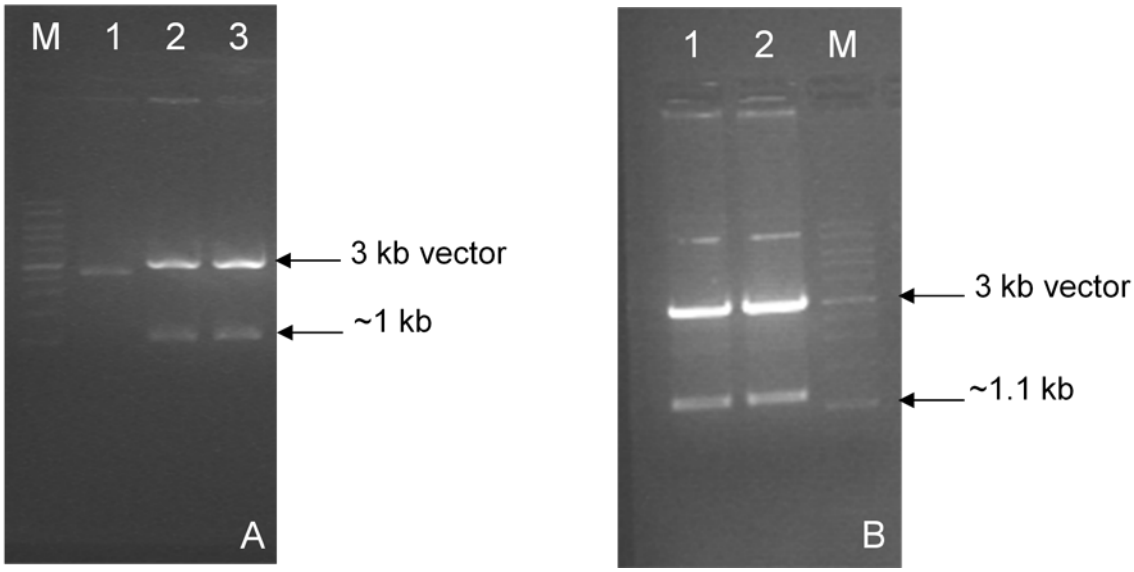


Fig. 3.7 A: Partial clone of GT1 gene releasing 1 kb insert when digested with *Eco* RI, lane M: marker, lane 1-3: cloned ~1 kb insert. **B:** Partial clone of GT2 gene releasing ~1.1 kb insert when digested with *Eco* R1, lane 1 & 2- Digested product, lane M- marker.

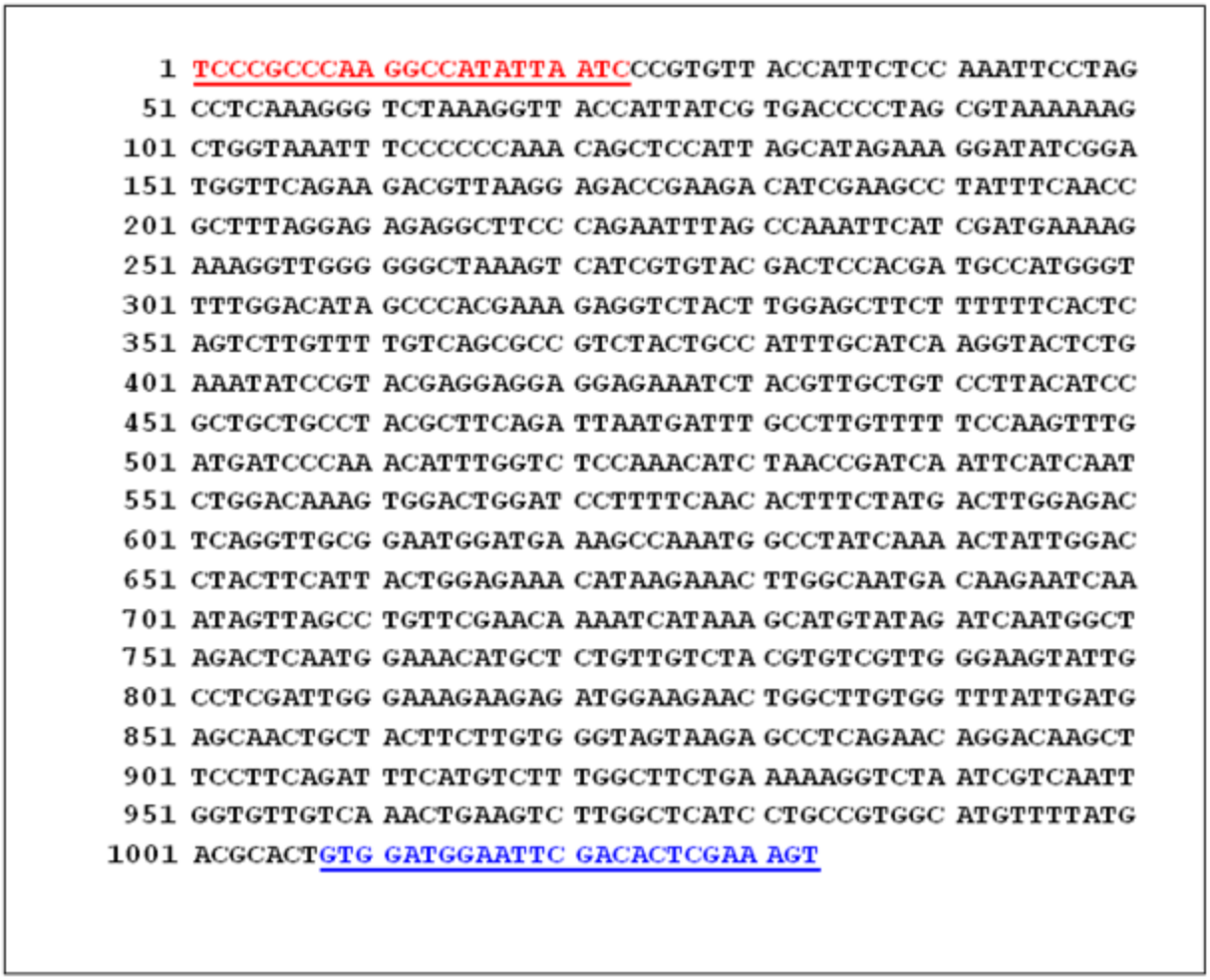


Fig. 3.8 A: Nucleotide sequence of partial clone of GT1 gene. Forward and reverse primers are highlighted in red and blue respectively.

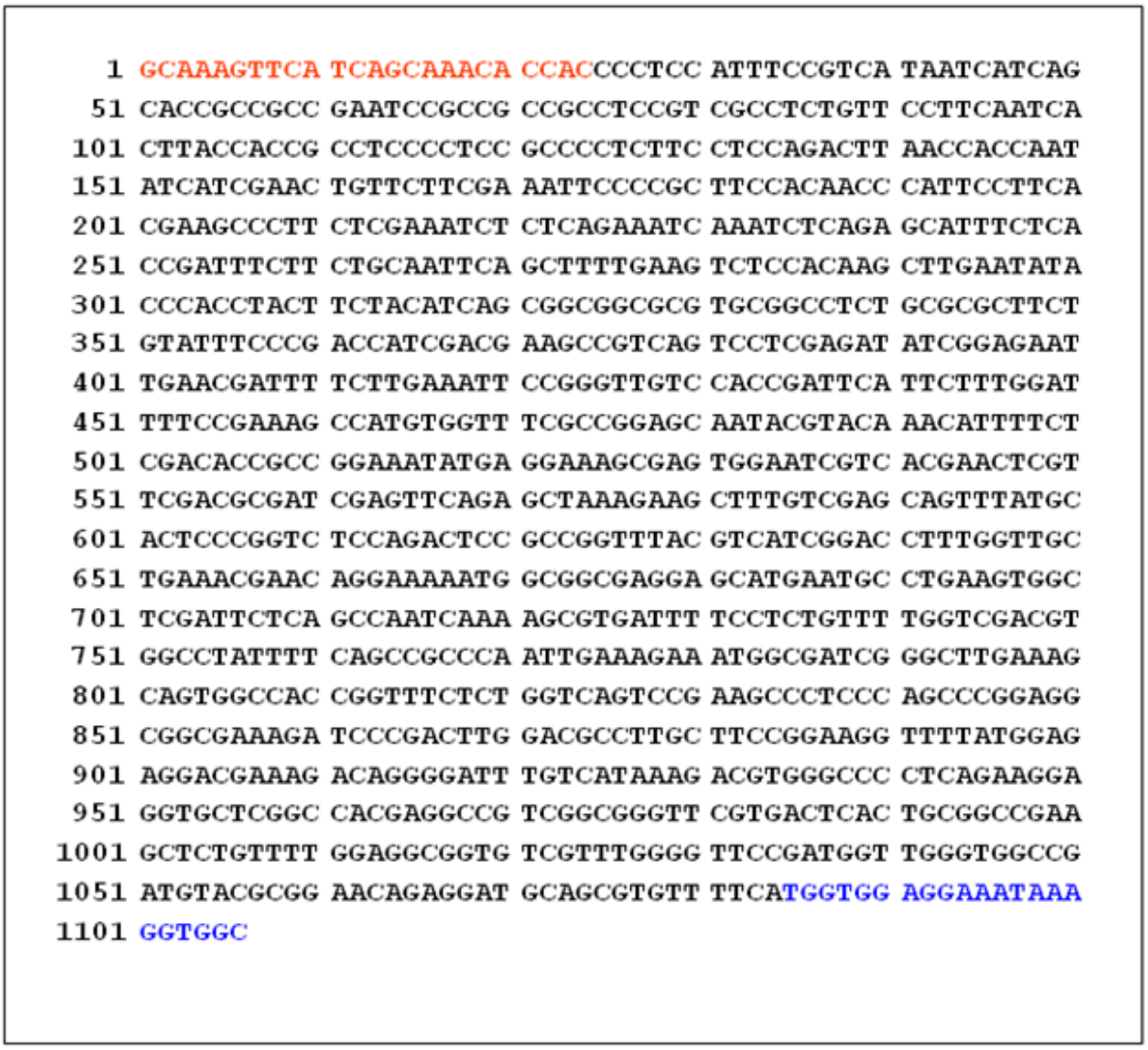


Fig. 3.8 B: Nucleotide sequence of partial clone of GT2 gene. Forward and reverse primers are highlighted in red and blue respectively.

3.3.2 Rapid Amplification of cDNA Ends (RACE)

RACE is an important tool to obtain the UTRs (Un-Translated Regions) of a particular gene and is also useful to obtain full length gene when a partial fragment is known.

3.3.2.1 Primer designing for RACE:

Based on the sequence of partial GT1 gene, one 5'Gene specific primer (GSP) and one 5'Nested GSP were designed for 5'RACE as shown below highlighted and underlined with blue and red respectively. For 3'RACE, NTMF1 was used as Gene specific primer and 3'Nested GSP was designed as shown below highlighted (with green) and underlined.

NTMF1 primer (3'Gene specific primer)
TCCCGCCCAAGGCCATATTAATCCCGTGTTACCATTCTCCAAATTCCTAGCCTCAAAGGGTC
TAAAGGTTACCATTATCGTGACCCCTAGCGTAAAAAAGCTGGTAAATTTCCCCCCAAACAGC
TCCATTAGCATAGAAAGGATATCGGATGGTTCAGAAGACGTTAAGGAGACCGAAGACATCGA
AGCCTATTTCAACCGCTTTAGGAGAGAGGCTTCCCAGAATTTAGCCAAATTCATCGATGAAA
AGAAAGGTTGGGGGGCTAAAGTCATCGTGTACGACTCCACGATGCCA

5'Nested GSP
TGGGTTTTGGACATAGCCCACGAAAGAGGTCTACTTGGAGCTTCTTTTTTCACTCAGTC
3' Nested GSP
TTGTTTTGTCAGCGCCGTCTACTGCCATTTGCATCAAGGTACTCTGAAATATCCGTACG

AGGAGGAGGAGAAATCTACGTTGCTGTCCTTACATCCGCTGCTGCCTACGCTTCAGATTAAT
GATTTGCCTTGTTTTTCCAAGTTTGATGATCCCAAACATTTGGTCTCCAAACATCT
AACCGATCAATTCATCAATCTGGACAAAGTGGACTGGATCCTTTTCAACACTTTCTATG

5'Gene specific primer
ACTTGGAGACTCAGGTTGCGGAATGGATGAAAGCCAAATGGCCTATCAAAACTATTGGA

CCTACTTCATTACTGGAGAAACATAAGAAACTTGGCAATGACAAGAATCAAATAGTTAGCCT
GTTCGAACAAAATCATAAAGCATGTATAGATCAATGGCTAGACTCAATGGAAACATGCTCTG
TTGTCTACGTGTCGTTGGGAAGTATTGCCTCGATTGGGAAAGAAGAGATGGAAGAACTGGCT
TGTGGTTTATTGATGAGCAACTGCTACTTCTTGTGGGTAGTAAGAGCCTCAGAACAGGACAA
GCTTCCTTCAGATTTCATGTCTTTGGCTTCTGAAAAAGGTCTAATCGTCAATTGGTGTTGTC
AAACTGAAGTCTTGGCTCATCCTGCCGTGGCATGTTTTATGACGCACTGTGGATGGAATTCG
ACACTCGAAAGT

Similarly using partial sequence of GT2 gene, one 5'Gene specific primer and one 5'Nested gene specific primers were designed for 5'RACE as shown below in red letters and highlighted with green respectively. For 3'RACE, one 3'Gene specific primer (underlined in pink letters) and one 3'Nested gene specific primers (highlighted with turquoise) were designed as shown below.

GCAAAGTTCATCAGCAAACACCACCCCTCCATTTCCGTCATAATCATCAGCACCGCCGCCGA
ATCCGCCGCCGCCTCCGTCGCCTCTGTTCCTTCAATCACTTACCACCGCCTCCCCTCCGCCC
CTCTTCCTCCAGACTTAACCACCAATATCATCGAACTGTTCTTCGAAATTCCCCGCTTCCAC
AACCCATTCCTTCACGAAGCCCTTCTCGAAATCTCTCAGAAATCAAATCTCAGAGCATTTCT
CACCGATTTCTTCTGCAATTCAGCTTTTGAAGTCTCCACAAGCTTGAATATACCCACCTACT
TCTACATCAGCGGCGGCGCGTGCGGCCTCTGCGCGCTTCTGTATTTCCCGACCATCGACGAA
GCCGTCAGTCCTCGAGATATCGGAGAATTGAACGATTTTCTTGAAATTCCGGGTTGTCCACC
GATTCATTCTTTGGATTTTCCGAAAGCCATGTGGTTTC

3'Gene specific primer

GCCGGAGCAATACGTACAAACATTTTCTCGACACCGCCGCAAATATGAGGAAAGCGAGT

5'Nested gene specific primer 3'Nested gene specific primer

GGAATCGTCACGAACTCGTTCGACGCGATCGAGTTCAGAGCTAAAGAAGCTTTGTCGAG

CAGTTTATGCACTCCCGGTCTCCAGACTCCGCCGGTTTACGTCATCGGACCTTTGGTTG

5'Gene specific primer

CTGAAACGAACAGGAAAAATGGCGGCGAGGAGCATGAATGCCTGAAGTGGCTCGATTCT
CAGCCAATCAAAAGCGTGATTTTCCTCTGTTTTGGTCGACGTGGCCTATTTTCAGCCGC
CCAATTGAAAGAAATGGCGATCGGGCTTGAAAGCAGTGGCCACCGGTTTCTCTGGTCAGTCC
GAAGCCCTCCCAGCCCGGAGGCGGCGAAAGATCCCGACTTGGACGCCTTGCTTCCGGAAGGT
TTTATGGAGAGGACGAAAGACAGGGGATTTGTCATAAAGACGTGGGCCCCTCAGAAGGAGGT
GCTCGGCCACGAGGCCGTCGGCGGGTTCGTGACTCACTGCGGCCGAAGCTCTGTTTTGGAGG
CGGTGTCGTTTGGGGTTCCGATGGTTGGGTGGCCGATGTACGCGGAACAGAGGATGCAGCGT
GTTTTCATGGTGGAGGAAATAAAGGTGGC

3.3.2.2 3' RACE reaction

To isolate the 3' end sequence of both the glycosyltransferase genes, 3' RACE reaction was performed (as described in Chapter 2, Section 2.8.4.10). In case of GT1 gene, 3'RACE was performed using forward Gene specific primer NTMF1 (5' TCC CGC CCA AGG CCA TAT TAA TC 3') and a reverse 3'GeneRacer primer provided with the kit. 3' RACE ready cDNA was used as the template. The RACE reaction resulted in an amplified product of approximately 1.2 kb; which was further re-amplified using 3' Nested GSP (5'CCG TCT ACT GCC ATT TGC ATC AAG GT 3') and 3' Nested GeneRacer primer (3'NGRP) provided with kit. The resultant PCR product (Fig. 3.9 A) was cloned into pGEMT-Easy vector (Promega, USA) and the recombinant construct was confirmed by *Eco* RI restriction analysis (Fig. 3.9 B) and by sequencing.

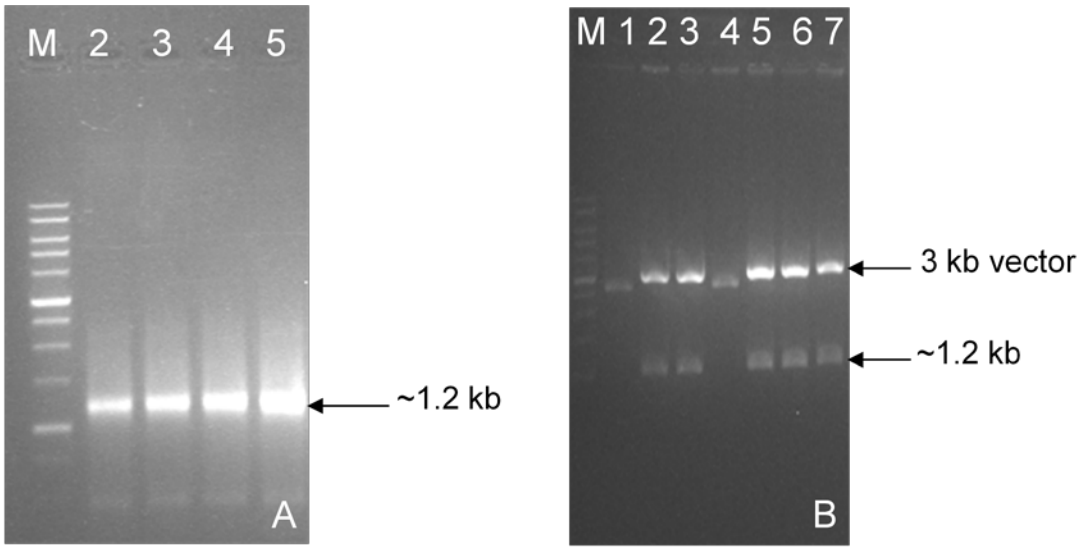


Fig. 3.9: A 1% agarose gel showing 3'RACE product of GT1 gene. Lane M- marker, lane 2-5 is ~1.2 kb 3'RACE product. **B** Lane M- marker, *Eco* RI digested recombinant pGEMT-Easy vector releasing ~ 1.2 kb fragment in lanes 2, 3, 5-7.

The analysis of the sequenced 3'RACE product of GT1 gene revealed 140 bp 3' UTR region, which is shown highlighted in blue (Fig. 3.10) and the stop codon (TGA) is highlighted in pink. The poly-A tail of mRNA is shown highlighted in red and the 3'NGRP is shown in green letters (Fig. 3.10).

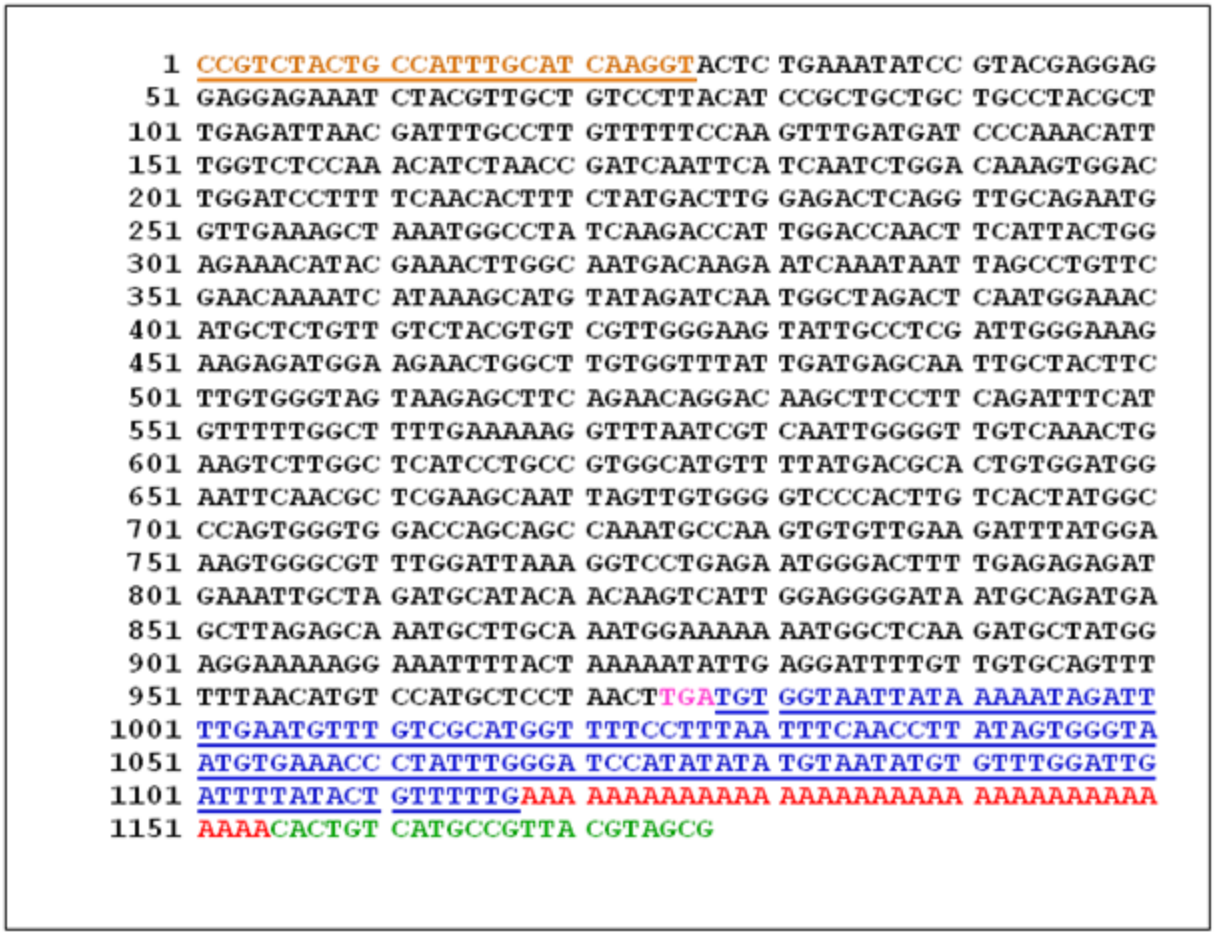


Fig. 3.10: Sequence analysis of 3' RACE product of GT1 gene: The 3'UTR is represented by 140 bp and it extends from stop codon till end of poly-A tail. Stop codon (TGA) is highlighted in pink, poly-A tail is highlighted in red and 3'NGRP is represented by green letters.

Similarly in case of GT2 gene, 3'RACE was performed using 3'Gene specific primer (5' CCG GAA ATA TGA GGA AAG CGA GT 3') and the 3'Gene Racer primer provided with the kit. 3'RACE cDNA was used as the template. The RACE reaction resulted in an amplification of approximately 1 kb fragment; which was further re-amplified using 3'Nested GSP (5' ATG AGG AAA GCG AGT GGA ATC GTC A 3') and 3' Nested GeneRacer primer provided with kit. The resultant PCR product (Fig. 3.11 A) was cloned into pGEMT-Easy vector and the recombinant construct was confirmed by *Eco* RI restriction analysis (Fig. 3.11 B) and by sequencing.

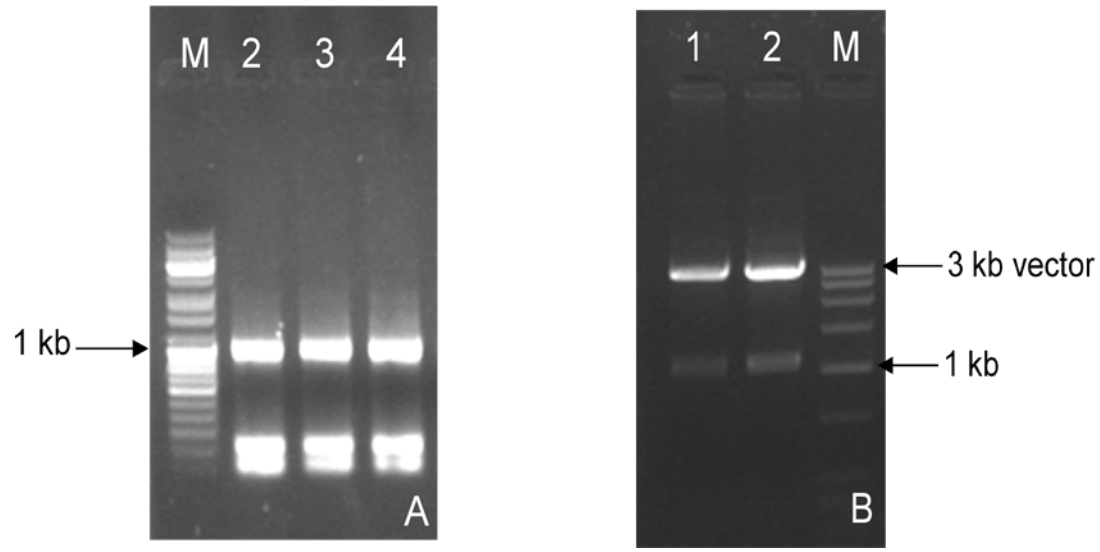


Fig. 3.11: A 1% agarose gel showing 3'RACE product of GT2 gene. Lane M- marker, lane 2-4 is 1 kb 3'RACE product. **B** Lane 1&2- *Eco* RI digested recombinant pGEMT-Easy vector releasing ~ 1 kb fragment Lane M- marker.

The analysis of the sequenced 3'RACE product of GT2 gene revealed 151 bp 3' UTR region, which is shown highlighted in pink (Fig. 3.12) and the stop codon (TAA) is highlighted in sky-blue. The poly-A tail of mRNA is shown highlighted in blue and the 3'NGRP is shown in green letters (Fig. 3.12).

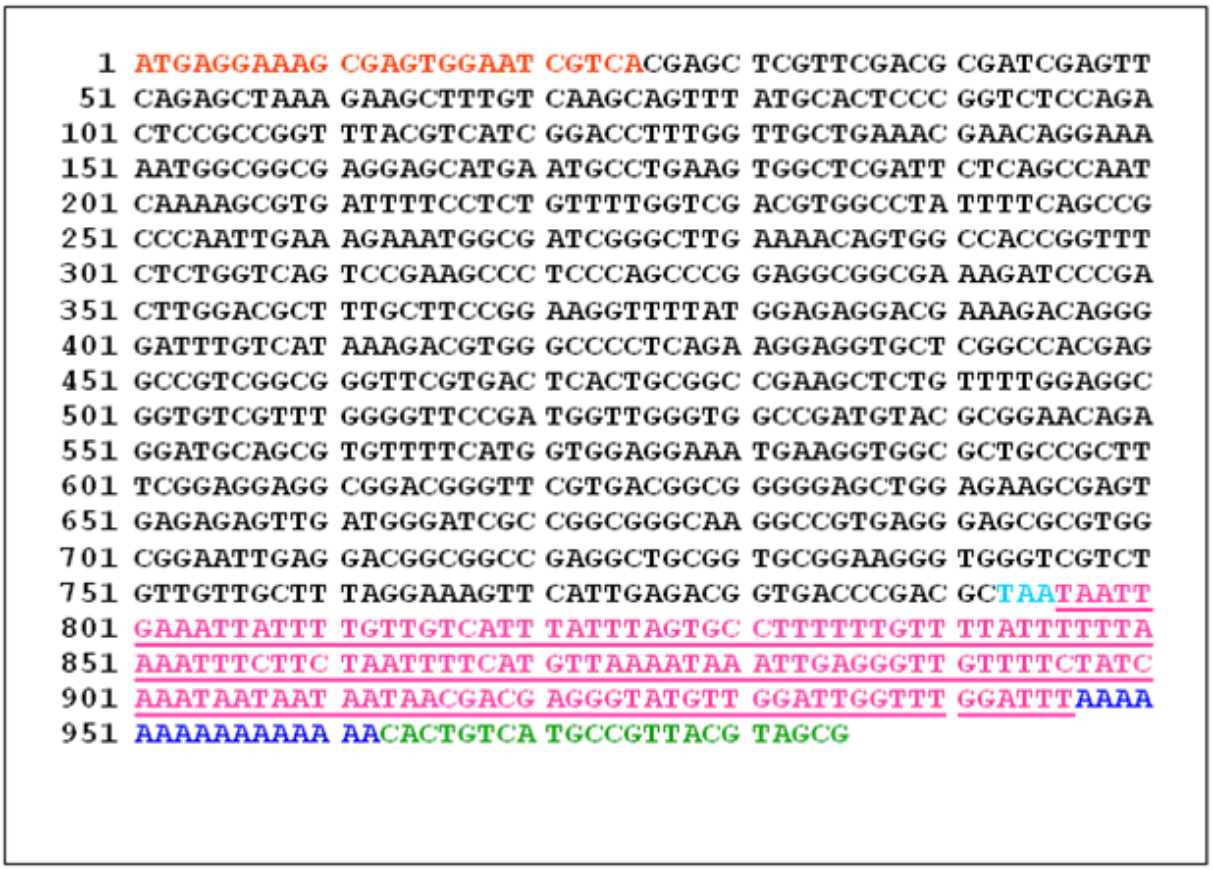


Fig. 3.12: Sequence analysis of 3' RACE product of GT2 gene: The 3'UTR is represented by 151 bp and it extends from stop codon till end of poly-A tail. Stop codon (TAA) is highlighted in skyblue, poly-A tail is highlighted in blue and 3'NGRP is represented by green letters.

3.3.2.3 5' RACE reaction

In order to get the 5'region of both the glycosyltransferase genes, 5' Rapid amplification of cDNA ends was performed as described in Chapter 2, Section 2.8.4.10. In case of GT1 gene 5'RACE was performed using gene specific reverse primer 5'GSP (5'TTC CGC AAC CTG AGT CTC CAA GTC AT 3') and 5'GeneRacer forward primer provided with the kit,. In this reaction 5' RACE ready cDNA was used as the template. The above RACE product was diluted (1:50) and secondary PCR was done using 5'nested GSP (5' GCG CTG ACA AAA CAA GAC TGA GTG 3') and a nested 5' GeneRacer primer provided with the kit. Agarose

gel electrophoresis analysis revealed an amplification product of approximately 500 bp as depicted in figure 3.13. This amplicon was cloned in pGEMT-Easy vector (Promega, USA) and the resulting construct was confirmed by *Eco*RI restriction analysis and by sequencing.

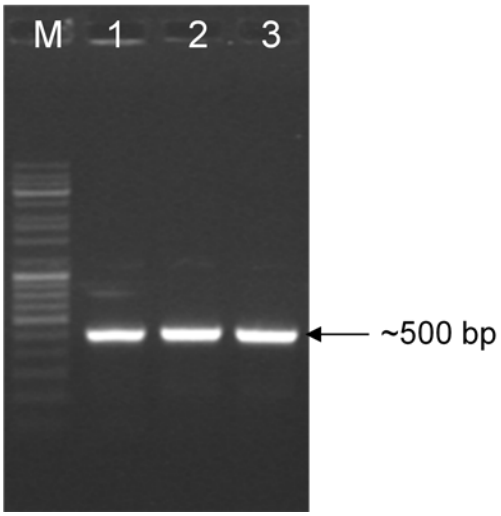


Fig. 3.13: 1% agarose gel showing 5'RACE product of GT1 gene. Lane M- marker, lane 1-3 is ~500 bp 5'RACE product.

The analysis of the sequenced product revealed the exact size of the amplicon to be 437 bp. The 5' GeneRacer nested primer and gene specific reverse primer sequences are underlined (Fig. 3.14). The partial 5' coding sequence of *GT* gene obtained from 5' RACE reaction is 410 bp. 5'GeneRacer primer got probably annealed to the start codon of the *GT* gene which could only provide extra 45 bps at the 5'end of the gene without confirming start codon.

Fig. 3.14: Sequence analysis of 5' RACE product of GT1 gene: Primers (5' NGRP & 5'NGSP shown highlighted in red and blue respectively.

Similarly 5'RACE reaction was performed for GT2 gene using 5'gene specific primer (5' AGA ATC GAG CCA CTT CAG GCA TT 3') and 5'GeneRacer primer supplied with the kit. 5'cDNA was used as template for the above reaction. The above 5'RACE reaction was diluted (1:50) times and the secondary RACE reaction was performed using 5'Nested GSP (5' TTC CGG CGG TGT CGA GAA AAT GTT TG 3') and 5'Nested GRP which gave an amplification of about 700 bp on 1% agarose gel (Fig. 3.15 A). This 700 bp fragment was cloned in pGEMT-easy vector and confirmed by *Eco*R1 restriction analysis (Fig. 3.15 B) and sequencing.

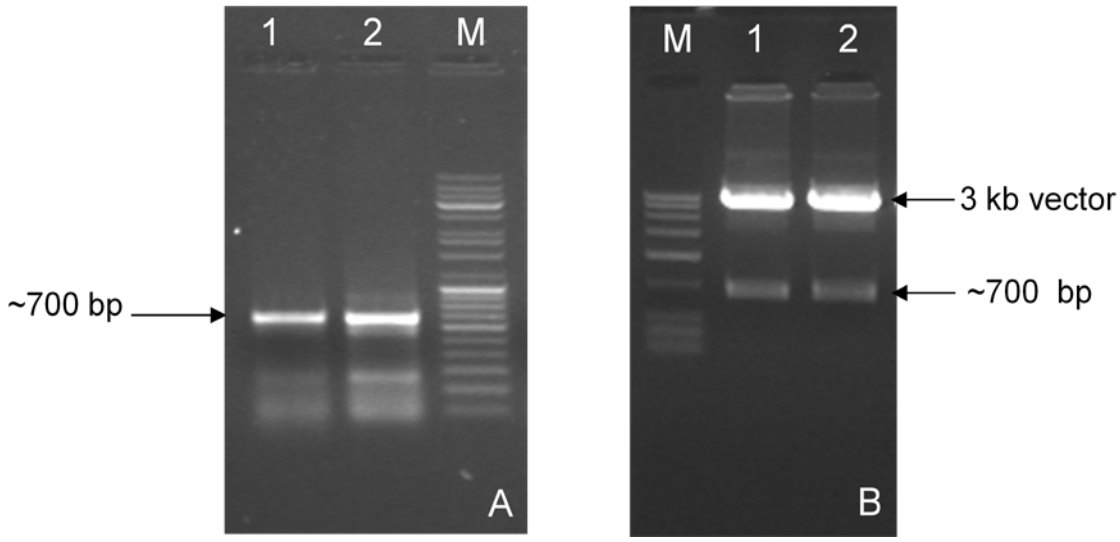


Fig. 3.15 A: 1% agarose gel showing 5'RACE product of GT2 gene. Lane 1 & 2: ~700 bp 5'RACE product, lane M- marker. **B** Lane M- marker, Lane 1&2- *Eco* RI digested recombinant pGEMT-Easy vector releasing ~ 700 bp fragment

The analysis of the sequenced product revealed the exact size of the amplicon to be 647 bp. The 5' GeneRacer nested primer and gene specific reverse primer sequences are underlined (Fig 3.16). The 5' UTR region is shown underlined (46 bp) and the start codon (ATG) is highlighted red (Fig 3.16). The partial 5' coding sequence of GT2 gene obtained from 5' RACE reaction is 574 bp.

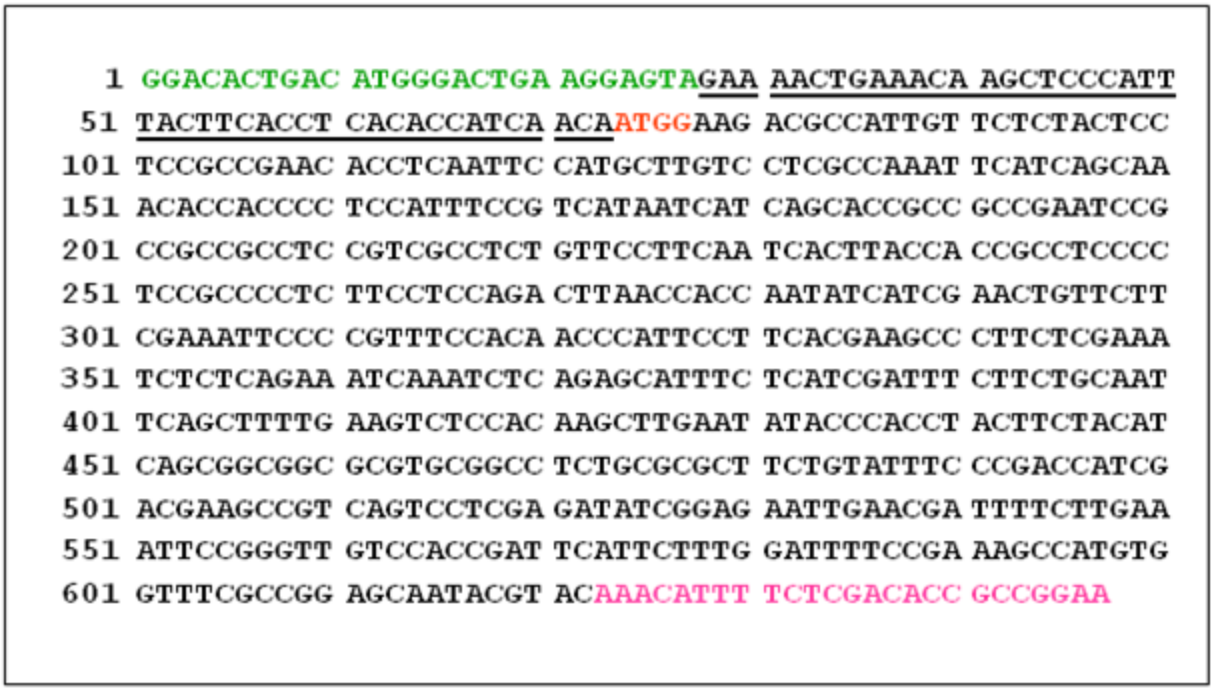


Fig 3.16: Sequence analysis of 5' RACE product of GT2 gene: Primers (5' NGRP & 5'NGSP shown highlighted in green and pink respectively.

3.3.2.4 Isolation of full length coding region of both the GT genes from cDNA of *B. monniera.*

3.3.2.4.1 Full length clone isolation of GT1 gene

5'RACE could provide only partial sequence from the 5'end of the GT1 gene which was due to annealing of the 5'GeneRacer primer supplied with the kit to the probable start codon of GT1 gene. In order to get the initiation codon of the GT1 gene, forward primer was designed from the extreme forward end of the 5'cDNA clone of the GT1 gene and reverse primer was designed from the stop codon of the 3'cDNA clone of the GT1 gene.

Forward primer (GT1F1) - 5'AAA GGA ACA GGG AAG GAG GCT CAT 3'
Reverse primer (BMGTR) - 5'TCA AGT TAG GAG CAG GGA CAT GTT AAA 3'

PCR was performed as mentioned in chapter 2, section 2.8.4.9 and an amplicon of around 1.4 kb was obtained (Fig. 3.17) which was cloned in pGEMT easy vector and confirmed by sequencing.

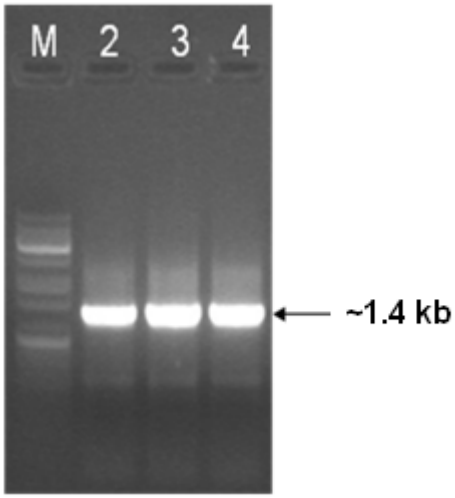


Fig. 3.17: 1% Agarose gel showing ~1.4 kb fragment of RGT1 gene.

The analysis of the sequence revealed the exact size of the amplicon to be 1384 bps (Fig. 3.18). This clone was named RGT1 in subsequent discussion. This RGT1 clone showed 50% identity at amino acid level with *Vitis vinefera* and 49% with *Ricinus communis* and *Lycium barbarum*.

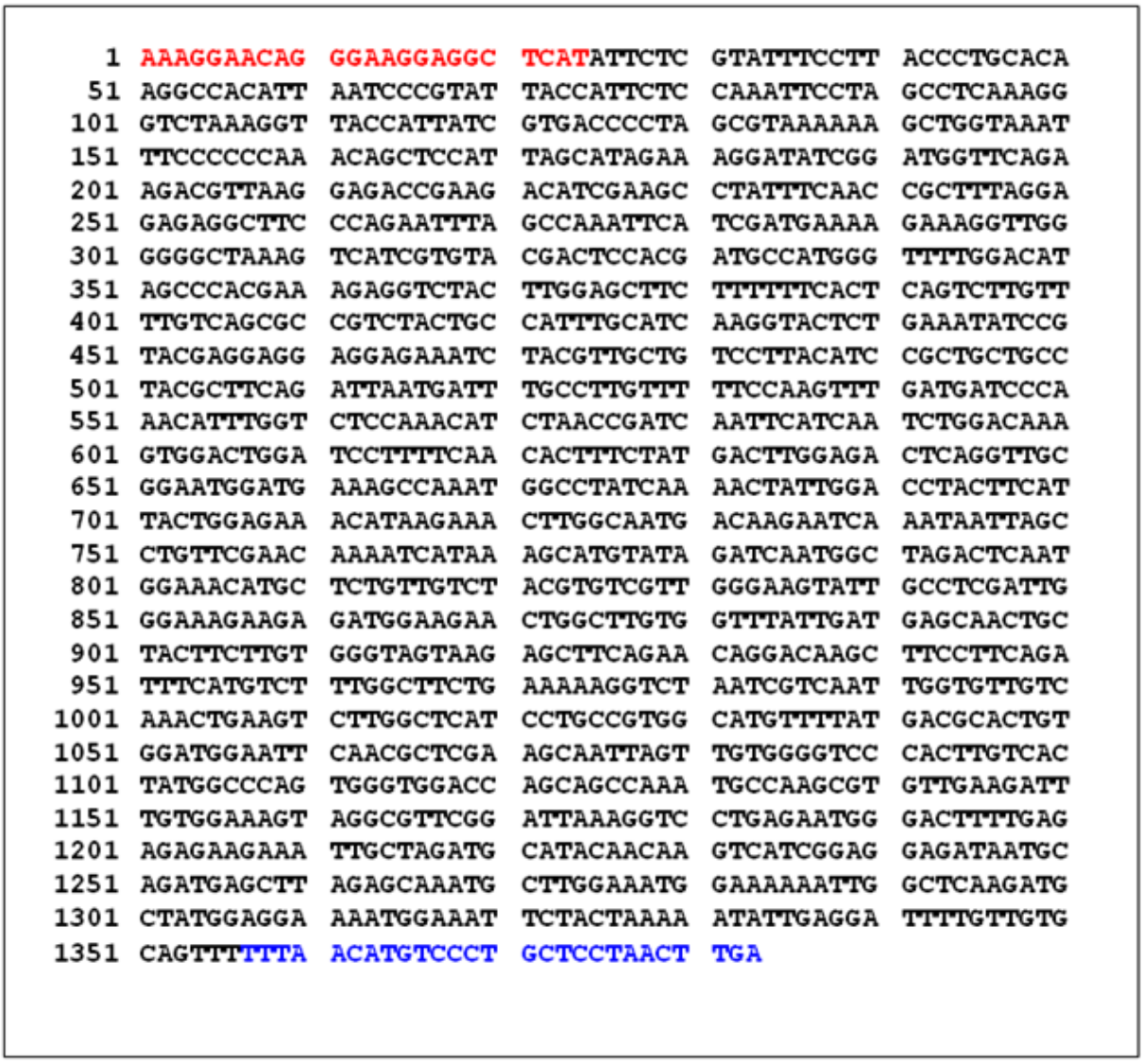


Fig. 3.18: Nucleotide sequence of the RGT1 clone. Forward primer is highlighted in red and reverse primer in blue letters.

In order to fish out the full length glycosyltransferase (GT1) gene from *B. monniera,* three different forward primers were designed from the initiation codon of the sequences to which RGT1 clone gave maximum homology i.e, *Vitis vinefera*, *Ricinus communis* and *Lycium barbarum*. Of the three forward primers, the forward primer designed from *Vitis vinefera* and reverse primer from the termination codon of RGT1 clone gave an amplification of around 1.4 kb (Fig. 3.19 A) which was expected to be the full length GT1 gene.

Forward Primer- 5'ATG GAG AGC AAA GGA ACA GGG AAG GA 3'
Reverse primer- 5'TCA AGT TAG GAG CAG GGA CAT GTT AAA 3'

This full length GT1 gene was cloned into pGEMT easy vector (Promega, USA) and was mobilized into *E. coli* XL-1 Blue cells. Plasmids were isolated from a few of the white recombinant colonies which survived on LB ampicillin plates containing X-gal – IPTG and were digested with *Eco* R1 to confirm the presence of the cloned insert (Fig. 3.19 B). The recombinant clone was further confirmed by sequencing.

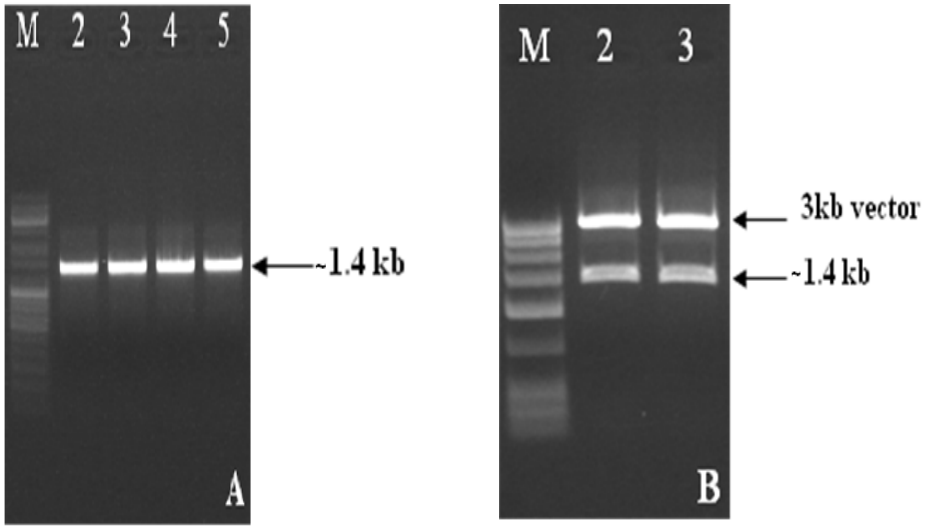


Fig. 3.19: A 1% Agarose gel showing full length GT1 gene PCR amplified from cDNA in lanes 2 to 5 (~1.4 kb) and lane M is marker. **B.** Restriction analysis of recombinant clones releasing ~1.4 kb insert (Lane 2-3) and lane M- marker.

Analysis of the sequenced product revealed that the exact size of the clone is 1392 bp (Fig. 3.20). This full length GT1 clone will be referred to as BMGT1 clone in all subsequent discussions. The BMGT1 clone showed 49% identity at amino acid level with the UDP-glucose:glycosyltransferase gene from *Lycium barbarum.*

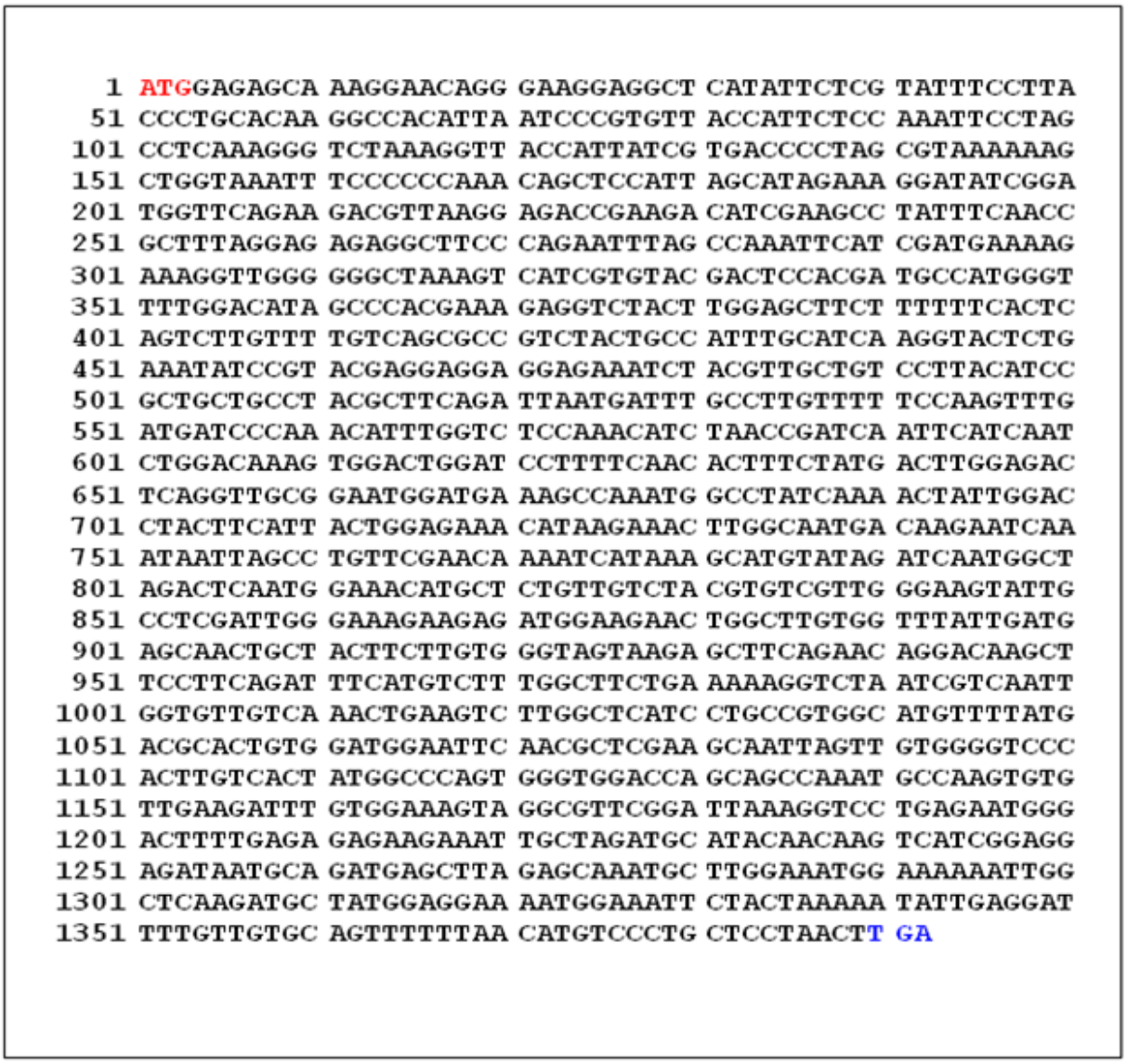


Fig. 3.20: Nucleotide sequence of BMGT1 gene.

3.3.2.4.1.1 Deduced amino acid sequence of BMGT1 gene

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

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

3.3.2.4.1.2 CLUSTAL W (1.8) multiple sequence alignment of amino acid sequences of BMGT1 with *Lycium barbarum*

```
BMGT1           MESKGTGKEAHILVFPYPAQGHINPVLPFSKFLASKGLKVTIIVTPSVKKLVNFPPNS-S 59
L.barbarum      MEEITN--KAHVLLLPYPLQGHINPMVQFSKRLASRGVKVTLVTIDNVSK--NMPKESGS 56
                **.  .  :**:*::*** ******:: *** ***:*:***::.  .*.*  *:* :* *

BMGT1           ISIERIS-DGSEDVKETEDIEAYFNRFRREASQNLAKFIDEKKGWGAKVIVYDSTMPWVL 118
L.barbarum      IKIESIPHDEAPPQSVDESLEWYFNLISKNLGAIVEKLSNSE--FPVKVLVFDSIGSWAL 114
                *.** *. * :   .  *.:* *** : :: .  : *: :.:  : .**:*:**  .*.*

BMGT1           DIAHERGLLGASFFTQSCFVSAVYCHLHQGTLKYPYEEEEKSTLLSLHPLLPTLQINDLP 178
L.barbarum      DLAHQLGLKGAAFFTQPCSLSAIFYHMDPETSKVPFDG----SVVTL-PSLPLLEKKDLP 169
                *:**: ** **:****.* :**:: *:.  * * *::     ::::* * ** *: :***

BMGT1           CFSKFDDPKHLVSKHLTDQFINLDKVDWILFNTFYDLETQVAEWMKAKWPIKTIGPTS-L 237
L.barbarum      TF-IYDDLYPSLAKLIFSQNIHFKKADWLLFNTFDVLEKEVVNWLRTQYPIKTIGPTIPS 228
                 *  :**    ::* : .* *::.*.**:*****  **.:*.:*:::::********

BMGT1           LEKHKKLGNDKNQIISLFEQNHKACIDQWLDSMETCSVVYVSLGSIASIGKEEMEELACG 297
L.barbarum      MYLDKRLKEDKEYGLSLFKPNGETCV-KWLDSREIGSVVYVSFGTLASLGEQQMEELAWG 287
                :  .*:* :**:  :***: * ::*: :**** *  ******:*::**:*:::***** *

BMGT1           LLMSNCYFLWVVRASEQDKLPSDFMSLASEKGLIVNWCCQTEVLAHPAVACFMTHCGWNS 357
L.barbarum      LMTSNCHFLWVVRTSEENKLPNEFMSKLSEKGLIVNWCPQLDVLAHQSVGCFFTHCGWNS 347
                *: ***:******:**::***.:***  ********** * :**** :*.**:*******

BMGT1           TLEAISCGVPLVTMAQWVDQQPNAKCVEDLWKVGVRIKGPENGTFEREEIARCIQQVIGG 417
L.barbarum      TLEALCLGVPMVAMPQWSDQPTNAKFISDVWQTGIRVKAGEDGVVNRDEIASSIREVMEE 407
                ****:. ***:*:*.** ** .*** :.*:*:.*:*:*. *:*..:*:*** .*::*:

BMGT1           DNADELRANAWKWKKLAQDAMEENGNSTKNIEDFVVQFFNMSLLLT 463
L.barbarum      EKGIMLKENAIKWKQLAKAAIDEGGSSDKNIEEFLSNL-------- 445
                ::.  *: ** ***:**: *::*.*.* ****:*: ::
```

3.3.2.4.2 Full length clone isolation of GT2 gene

The full length open reading frame of GT2 gene was achieved by designing a forward primer from the initiation codon of the 5’RACE product of the GT2 gene and reverse primer from the stop codon of the 3’RACE product of the GT2 gene. cDNA from *Bacopa monniera* was used as the template.

Forward primer- 5’ ATG GAA GAC GCC ATT GTT CTC TAC TC 3’
Reverse primer- 5’ TTA GCG TCG GGT CAC CGT CTC AAT 3’

PCR was performed as mentioned in chapter 2, section 2.8.4.9 and an amplicon of ~1.35 kb was obtained as expected with the full length GT2 gene (Fig. 3.21 A). The amplicon was excised from the agarose gel and purified using Axygen^{TM} GEL elution kit as mentioned in chapter 2, section 2.8.4.4. The full length PCR product was cloned into pGEMT- easy vector and was mobilized into *E. coli* XL-1 Blue cells. Plasmids were isolated from a few of the white recombinant colonies which survived on LB ampicillin plates containing X-gal – IPTG and were digested with EcoR1 to confirm the presence of the cloned insert (Fig. 3.21 B). The recombinant clone was further confirmed by sequencing.

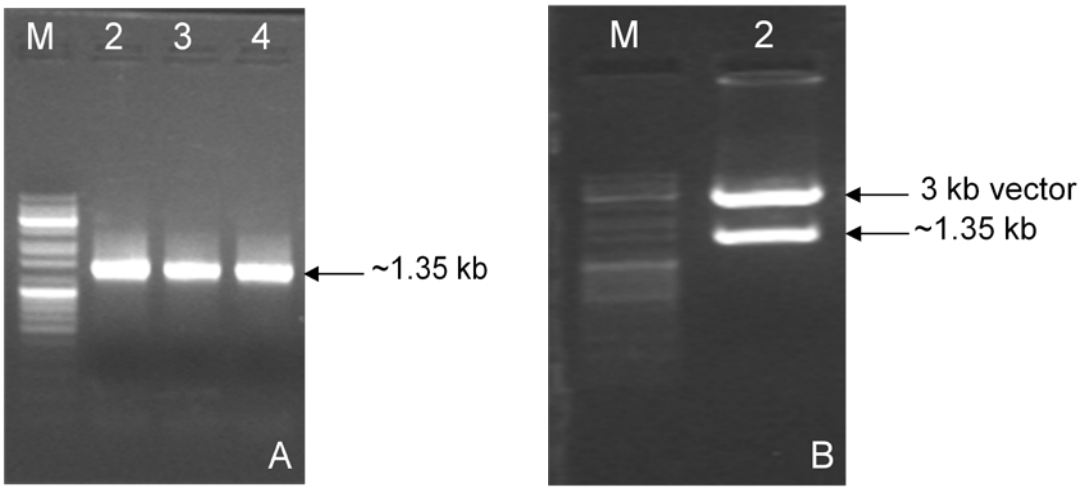
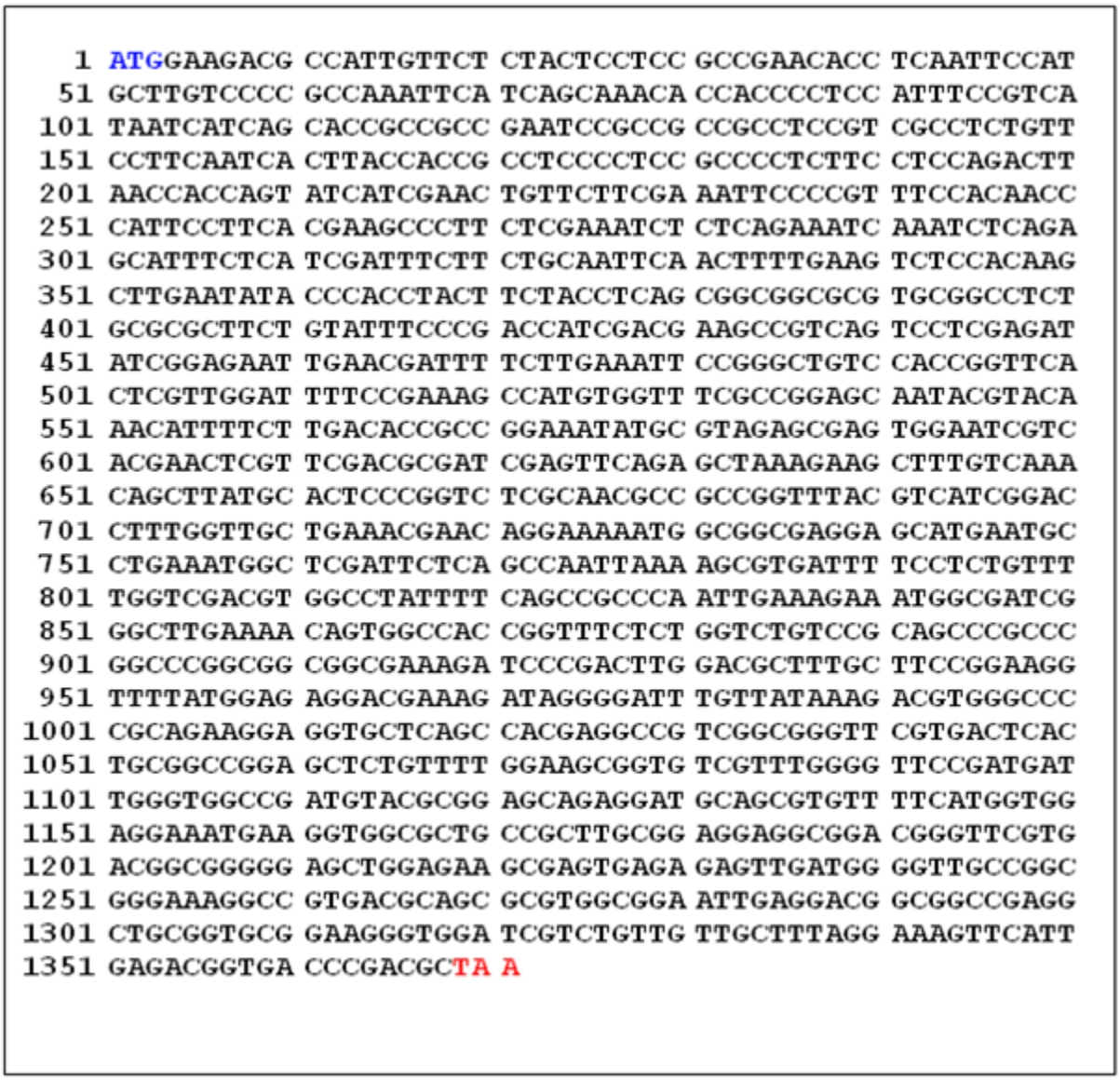


Fig. 3.21: A 1% Agarose gel showing full length GT2 gene PCR amplified from cDNA in lanes 2 to 4 (~1.35 kb) and lane M is marker. **B.** Restriction analysis of recombinant clone releasing ~1.35 kb insert (Lane 2) and lane M- marker.

Analysis of the sequenced product revealed that the exact size of the clone is 1371 bp (Fig. 3.22). This full length GT2 clone will be referred to as BMGT2 clone in all subsequent

discussions. The BMGT2 clone showed 72% identity at amino acid level with the UDP-glucose:glycosyltransferase gene from *Sesamum indicum.*



```
   1 ATGGAAGACG CCATTGTTCT CTACTCCTCC GCCGAACACC TCAATTCCAT
  51 GCTTGTCCCC GCCAAATTCA TCAGCAAACA CCACCCCTCC ATTTCCGTCA
 101 TAATCATCAG CACCGCCGCC GAATCCGCCG CCGCCTCCGT CGCCTCTGTT
 151 CCTTCAATCA CTTACCACCG CCTCCCCCTCC GCCCCTCTTC CTCCAGACTT
 201 AACCACCAGT ATCATCGAAC TGTTCTTCGA AATTCCCCGT TTCCACAACC
 251 CATTCCTTCA CGAAGCCCTT CTCGAAATCT CTCAGAAATC AAATCTCAGA
 301 GCATTTCTCA TCGATTTCTT CTGCAATTCA ACTTTTGAAG TCTCCACAAG
 351 CTTGAATATA CCCACCTACT TCTACCTCAG CGGCGGCGCG TGCGGCCTCT
 401 GCGCGCTTCT GTATTTCCCG ACCATCGACG AAGCCGTCAG TCCTCGAGAT
 451 ATCGGAGAAT TGAACGATTT TCTTGAAATT CCGGGCTGTC CACCGGTTCA
 501 CTCGTTGGAT TTTCCGAAAG CCATGTGGTT TCGCCGGAGC AATACGTACA
 551 AACATTTTCT TGACACCGCC GGAAATATGC GTAGAGCGAG TGGAATCGTC
 601 ACGAACTCGT TCGACGCGAT CGAGTTCAGA GCTAAAGAAG CTTTGTCAAA
 651 CAGCTTATGC ACTCCCGGTC TCGCAACGCC GCCGGTTTAC GTCATCGGAC
 701 CTTTGGTTGC TGAAACGAAC AGGAAAAATG GCGGCGAGGA GCATGAATGC
 751 CTGAAATGGC TCGATTCTCA GCCAATTAAA AGCGTGATTT TCCTCTGTTT
 801 TGGTCGACGT GGCCTATTTT CAGCCGCCCA ATTGAAAGAA ATGGCGATCG
 851 GGCTTGAAAA CAGTGGCCAC CGGTTTCTCT GGTCTGTCCG CAGCCCGCCC
 901 GGCCCCGGCGG CGGCGAAAGA TCCCGACTTG GACGCTTTGC TTCCGGAAGG
 951 TTTTATGGAG AGGACGAAAG ATAGGGGATT TGTTATAAAG ACGTGGGCCC
1001 CGCAGAAGGA GGTGCTCAGC CACGAGGCCG TCGGCGGGTT CGTGACTCAC
1051 TGCGGCCGGA GCTCTGTTTT GGAAGCGGTG TCGTTTGGGG TTCCGATGAT
1101 TGGGTGGCCG ATGTACGCGG AGCAGAGGAT GCAGCGTGTT TTCATGGTGG
1151 AGGAAATGAA GGTGGCGCTG CCGCTTGCGG AGGAGGCGGA CGGGTTCGTG
1201 ACGGCGGGGG AGCTGGAGAA GCGAGTGAGA GAGTTGATGG GGTTGCCGGC
1251 GGGAAAGGCC GTGACGCAGC GCGTGGCGGA ATTGAGGACG GCGGCCGAGG
1301 CTGCGGTGCG GAAGGGTGGA TCGTCTGTTG TTGCTTTAGG AAAGTTCATT
1351 GAGACGGTGA CCCGACGCTA A
```

Fig. 3.22: Nucleotide sequence of BMGT2 gene.

3.3.2.4.2.1 Deduced amino acid sequence of BMGT2 gene

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

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

3.3.2.4.2.2 CLUSTAL W (1.8) multiple sequence alignment of amino acid sequences of BMGT2 with *Sesamum indicum*

```
BMGT2           MEDAIVLYSSAEHLNSMLVPAKFISKHHPSISVIIISTAAESAAASVASVPSITYHRLPS 60
S.indicum       MEDTVVLYTSAEHLNSMLVLAKFISKHYPSIPLLILCSAPESAAASVATVPSITYHRLPP 60
                ***::***:********** *******:***.::*:.:*.********:**********.

BMGT2           APLPPDLTTSIIELFFEIPRFHNPFLHEALLEISQKSNLRAFLIDFFCNSTFEVSTSLNI 120
S.indicum       PALPPNLTTNPLELLFEIPRLNNPNVSKALQEISQKSRIKAFVIDFFCNPVFEVSTGLNI 120
                ..***:***. :**:*****::** : :** ******.::**:******..*****.***

BMGT2           PTYFYLSGGACGLCALLYFPTIDEAVSPRDIGELNDFLEIPGCPPVHSLDFPKAMWFRRS 180
S.indicum       PTYFYISSGAFGLCPFLNFPTIEETV-PGDLADLNDFVEIPGCPPVHSSDFPEAMIHRKS 179
                *****:*.** ***.:* ****:*:* * *:.:****:********** ***:** .*:*

BMGT2           NTYKHFLDTAGNMRRASGIVTNSFDAIEFRAKEALSNSLCTPGLATPPVYVIGPLVAETN 240
S.indicum       NIYKHFMDAARNMAKSTGNLVNAFDALEFRAKEALINGLCIPNAPTPPVYLVGPLVGDSN 239
                * ****:*:* ** :::* :.*:***:******** *.** *. .*****::****.::*

BMGT2           RKNGGEEHECLKWLDSQPIKSVIFLCFGRRGLFSAAQLKEMAIGLENSGHRFLWSVRSPP 300
S.indicum       RNNGCIQHECLKWLDSQPSKSVIFLCFGRRGLFSVEQLKEMALGLENSGYRFLWSVRSPP 299
                *:**  :*********** ***************. ******:******:**********

BMGT2           GP--AAAKDPDLDALLPEGFMERTKDRGFVIKTWAPQKEVLSHEAVGGFVTHCGRSSVLE 358
S.indicum       GKQNSAAAEPDLDELLPKGFLERTKDRGFIIKSWAPQTEVLSHDSVGGFVTHCGRSSILE 359
                *   :** :**** ***:**:********:**:****.*****::************:**

BMGT2           AVSFGVPMIGWPMYAEQRMQRVFMVEEMKVALPLAEEADGFVTAGELEKRVRELMGLPAG 418
S.indicum       AVSLGVPMIGWPLYAEQRMNRVFMVEEMKVALPLEETADGLVTAVELEKRVRQLMDSQTG 419
                ***:********:******:************** * ***:*** *******:**.  :*

BMGT2           KAVTQRVAELRTAAEAAVRKGGSSVVALGKFIETVTRR 456
S.indicum       RAVRHRVTELKSSAAAAVRKNGSSLVALQNFIASVTRV 457
                :** :**:**:::* *****.***:*** :** :***
```

3.3.2.5 Characterization of BMGT1 and BMGT2 cDNA sequences

3.3.2.5.1 Characterization of BMGT1 cDNA sequences (FJ586244)

The full length cDNA sequence of BMGT1 has been submitted to NCBI GenBank database under accession no. FJ586244. The *Bacopa GT* gene (BMGT1) sequence contains a 1392 bp open-reading frame, which codes for a protein of 463 amino acids with a molecular weight 52.3 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 (TGA) and has yielded a 140 bp 3'UTR and a poly-A tail

3.3.2.5.1a Nucleotide sequence of BMGT1

```
   1 ATGGAGAGCA AAGGAACAGG GAAGGAGGCT CATATTCTCG TATTTCCTTA
  51 CCCTGCACAA GGCCACATTA ATCCCGTGTT ACCATTCTCC AAATTCCTAG
 101 CCTCAAAGGG TCTAAAGGTT ACCATTATCG TGACCCCTAG CGTAAAAAAG
 151 CTGGTAAATT TCCCCCCAAA CAGCTCCATT AGCATAGAAA GGATATCGGA
 201 TGGTTCAGAA GACGTTAAGG AGACCGAAGA CATCGAAGCC TATTTCAACC
 251 GCTTTAGGAG AGAGGCTTCC CAGAATTTAG CCAAATTCAT CGATGAAAAG
 301 AAAGGTTGGG GGGCTAAAGT CATCGTGTAC GACTCCACGA TGCCATGGGT
 351 TTTGGACATA GCCCACGAAA GAGGTCTACT TGGAGCTTCT TTTTTCACTC
 401 AGTCTTGTTT TGTCAGCGCC GTCTACTGCC ATTTGCATCA AGGTACTCTG
 451 AAATATCCGT ACGAGGAGGA GGAGAAATCT ACGTTGCTGT CCTTACATCC
 501 GCTGCTGCCT ACGCTTCAGA TTAATGATTT GCCTTGTTTT TCCAAGTTTG
 551 ATGATCCCAA ACATTTGGTC TCCAAACATC TAACCGATCA ATTCATCAAT
 601 CTGGACAAAG TGGACTGGAT CCTTTTCAAC ACTTTCTATG ACTTGGAGAC
 651 TCAGGTTGCG GAATGGATGA AAGCCAAATG GCCTATCAAA ACTATTGGAC
 701 CTACTTCATT ACTGGAGAAA CATAAGAAAC TTGGCAATGA CAAGAATCAA
 751 ATAATTAGCC TGTTCGAACA AAATCATAAA GCATGTATAG ATCAATGGCT
 801 AGACTCAATG GAAACATGCT CTGTTGTCTA CGTGTCGTTG GGAAGTATTG
 851 CCTCGATTGG GAAAGAAGAG ATGGAAGAAC TGGCTTGTGG TTTATTGATG
 901 AGCAACTGCT ACTTCTTGTG GGTAGTAAGA GCTTCAGAAC AGGACAAGCT
 951 TCCTTCAGAT TTCATGTCTT TGGCTTCTGA AAAAGGTCTA ATCGTCAATT
1001 GGTGTTGTCA AACTGAAGTC TTGGCTCATC CTGCCGTGGC ATGTTTTATG
1051 ACGCACTGTG GATGGAATTC AACGCTCGAA GCAATTAGTT GTGGGGTCCC
1101 ACTTGTCACT ATGGCCCAGT GGGTGGACCA GCAGCCAAAT GCCAAGTGTG
1151 TTGAAGATTT GTGGAAAGTA GGCGTTCGGA TTAAAGGTCC TGAGAATGGG
1201 ACTTTTGAGA GAGAAGAAAT TGCTAGATGC ATACAACAAG TCATCGGAGG
1251 AGATAATGCA GATGAGCTTA GAGCAAATGC TTGGAAATGG AAAAAATTGG
1301 CTCAAGATGC TATGGAGGAA AATGGAAATT CTACTAAAAA TATTGAGGAT
1351 TTTGTTGTGC AGTTTTTTAA CATGTCCCTG CTCCTAACTT GA
```

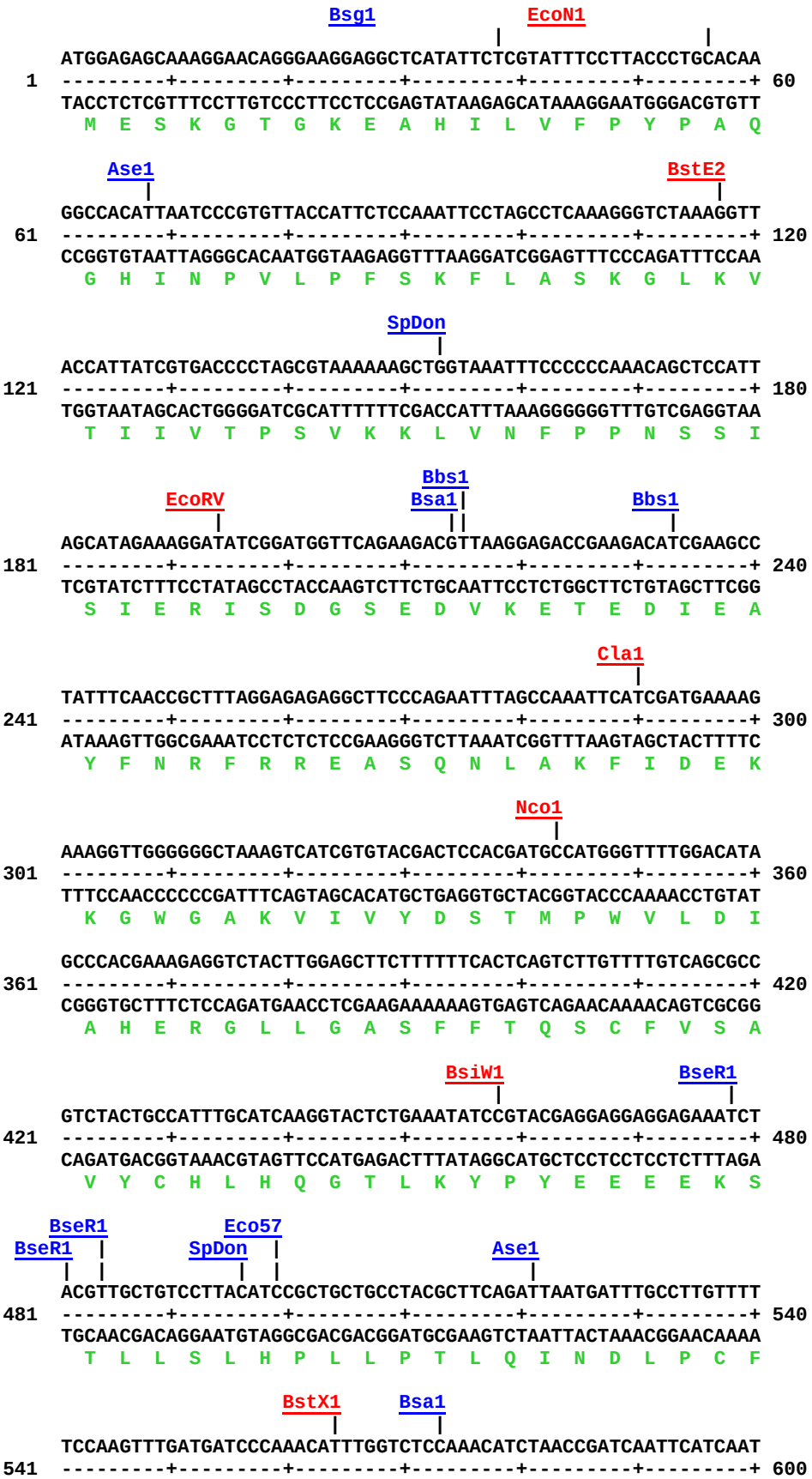
3.3.2.5.1b Deduced amino acid sequence of BMGT1

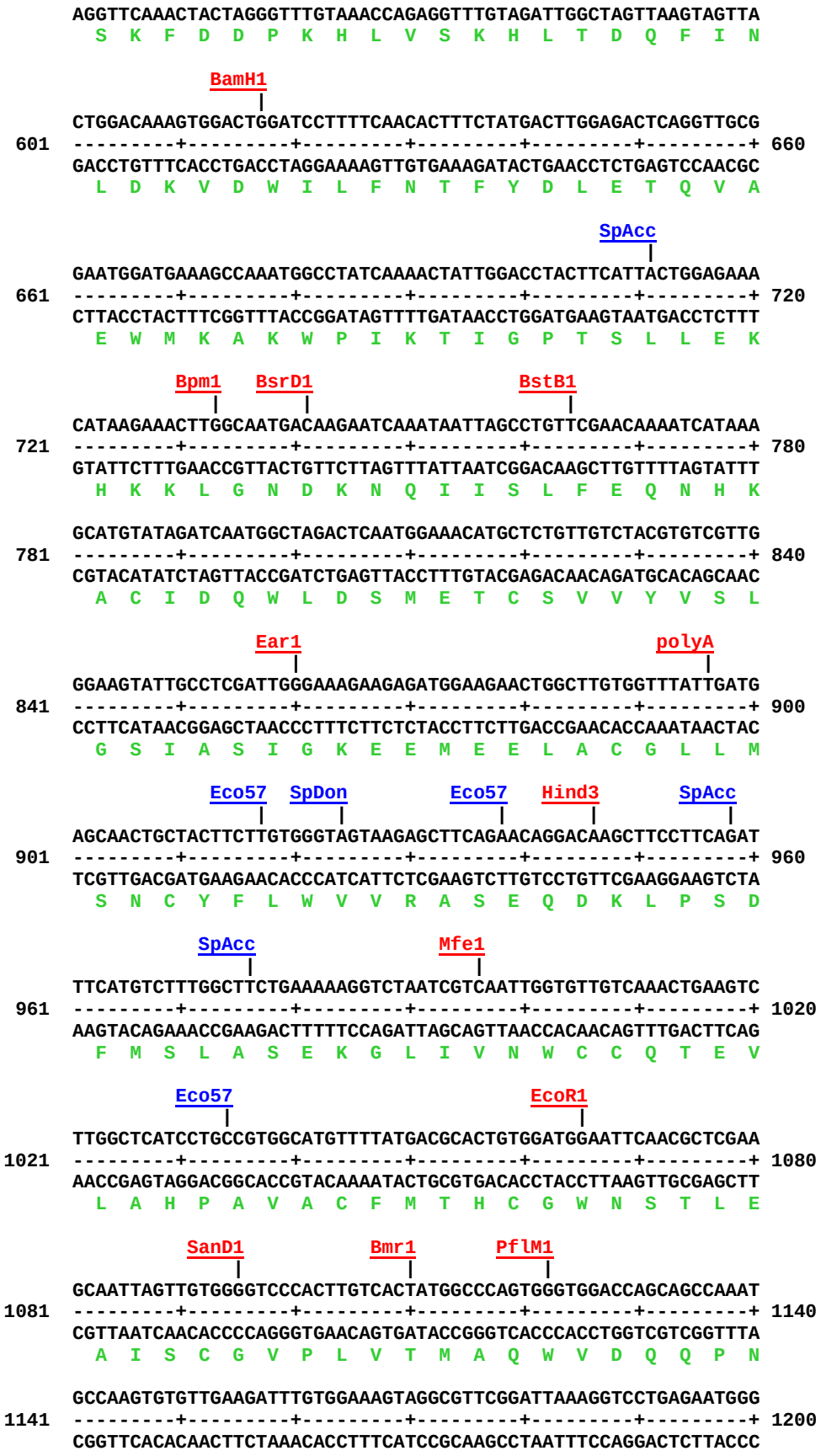
M E S K G T G K E A H I L V F P Y P A Q G H I N P V L P F S K F L A S K G L K V T I I V T P S V K K L V N F P P N S S I S I E R I S D G S E D V K E T E D I E A Y F N R F R R E A S Q N L A K F I D E K K G W G A K V I V Y D S T M P W V L D I A H E R G L L G A S F F T Q S C F V S A V Y C H L H Q G T L K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

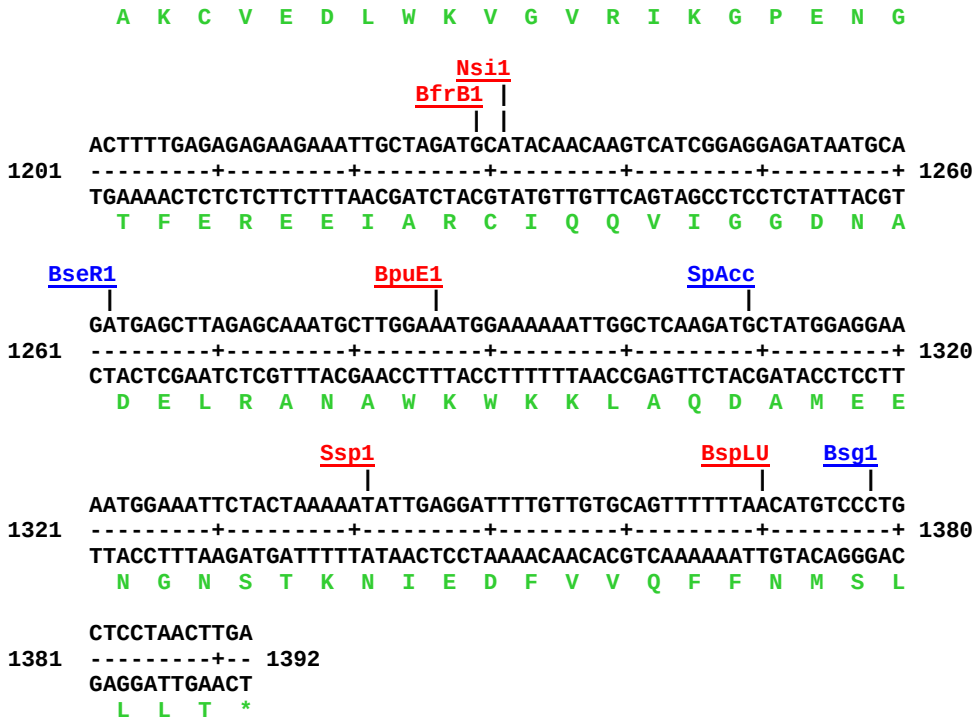
Analysis of the deduced amino acid sequence showed it contained the donor binding site (bold and underlined) and the 46 bp PSPG box which is conserved in all plant glycosyltransferase genes (highlighted in turquoise). Most conserved region of PSPG box i.e, **HCGWNS** is also present in the sequence. It also contained the tryptophan residue in the most conserved region of PSPG box (**HCGWNS**) which is found to be responsible for glucose specificity.

3.3.2.5.1c Restriction map of BMGT1 gene

Restriction map of BMGT1 gene was generated by using an online bioinformatic tool, Webmap DNA. Map created at http://pga.mgh.harvard.edu/web_apps/web_map/start. Editor at http://pga.mgh.harvard.edu/web_apps/dna_utilities.html Sites protected by *E. coli* methylation are not shown in the map.


Bsg1
EcoN1
ATGGAGAGCAAAGGAACAGGGAAGGAGGCTCATATTCTCGTATTTCCTTACCCTGCACAA
1 --------+---------+---------+---------+---------+---------+ 60
TACCTCTCGTTTCCTTGTCCCTTCCTCCGAGTATAAGAGCATAAAGGAATGGGACGTGTT
M E S K G T G K E A H I L V F P Y P A Q
Ase1
BstE2
GGCCACATTAATCCCGTGTTACCATTCTCCAAATTCCTAGCCTCAAAGGGTCTAAAGGTT
61 --------+---------+---------+---------+---------+---------+ 120
CCGGTGTAATTAGGGCACAATGGTAAGAGGTTTAAGGATCGGAGTTTCCCAGATTTCCAA
G H I N P V L P F S K F L A S K G L K V
SpDon
ACCATTATCGTGACCCCTAGCGTAAAAAAGCTGGTAAATTTCCCCCCAAACAGCTCCATT
121 --------+---------+---------+---------+---------+---------+ 180
TGGTAATAGCACTGGGGATCGCATTTTTTCGACCATTTAAAGGGGGGTTTGTCGAGGTAA
T I I V T P S V K K L V N F P P N S S I
Bbs1
EcoRV
Bsa1
Bbs1
AGCATAGAAAGGATATCGGATGGTTCAGAAGACGTTAAGGAGACCGAAGACATCGAAGCC
181 --------+---------+---------+---------+---------+---------+ 240
TCGTATCTTTCCTATAGCCTACCAAGTCTTCTGCAATTCCTCTGGCTTCTGTAGCTTCGG
S I E R I S D G S E D V K E T E D I E A
Cla1
TATTTCAACCGCTTTAGGAGAGAGGCTTCCCAGAATTTAGCCAAATTCATCGATGAAAAG
241 --------+---------+---------+---------+---------+---------+ 300
ATAAAGTTGGCGAAATCCTCTCTCCGAAGGGTCTTAAATCGGTTTAAGTAGCTACTTTTC
Y F N R F R R E A S Q N L A K F I D E K
Nco1
AAAGGTTGGGGGGCTAAAGTCATCGTGTACGACTCCACGATGCCATGGGTTTTGGACATA
301 --------+---------+---------+---------+---------+---------+ 360
TTTCCAACCCCCCGATTTCAGTAGCACATGCTGAGGTGCTACGGTACCCAAAACCTGTAT
K G W G A K V I V Y D S T M P W V L D I
GCCCACGAAAGAGGTCTACTTGGAGCTTCTTTTTTCACTCAGTCTTGTTTTGTCAGCGCC
361 --------+---------+---------+---------+---------+---------+ 420
CGGGTGCTTTCTCCAGATGAACCTCGAAGAAAAAAGTGAGTCAGAACAAAACAGTCGCGG
A H E R G L L G A S F F T Q S C F V S A
BsiW1
BseR1
GTCTACTGCCATTTGCATCAAGGTACTCTGAAATATCCGTACGAGGAGGAGGAGAAATCT
421 --------+---------+---------+---------+---------+---------+ 480
CAGATGACGGTAAACGTAGTTCCATGAGACTTTATAGGCATGCTCCTCCTCCTCTTTAGA
V Y C H L H Q G T L K Y P Y E E E E K S
BseR1
BseR1
Eco57
SpDon
Ase1
ACGTTGCTGTCCTTACATCCGCTGCTGCCTACGCTTCAGATTAATGATTTGCCTTGTTTT
481 --------+---------+---------+---------+---------+---------+ 540
TGCAACGACAGGAATGTAGGCGACGACGGATGCGAAGTCTAATTACTAAACGGAACAAAA
T L L S L H P L L P T L Q I N D L P C F
BstX1
Bsa1
TCCAAGTTTGATGATCCCAAACATTTGGTCTCCAAACATCTAACCGATCAATTCATCAAT
541 --------+---------+---------+---------+---------+---------+ 600


AGGTTCAAACTACTAGGGTTTGTAAACCAGAGGTTTGTAGATTGGCTAGTTAAGTAGTTA
S K F D D P K H L V S K H L T D Q F I N
BamH1
CTGGACAAAGTGGACTGGATCCTTTTCAACACTTTCTATGACTTGGAGACTCAGGTTGCG
601 --------+--------+--------+--------+--------+--------+ 660
GACCTGTTTCACCTGACCTAGGAAAAGTTGTGAAAGATACTGAACCTCTGAGTCCAACGC
L D K V D W I L F N T F Y D L E T Q V A
SpAcc
GAATGGATGAAAGCCAAATGGCCTATCAAAACTATTGGACCTACTTCATTACTGGAGAAA
661 --------+--------+--------+--------+--------+--------+ 720
CTTACCTACTTTCGGTTTACCGGATAGTTTTGATAACCTGGATGAAGTAATGACCTCTTT
E W M K A K W P I K T I G P T S L L E K
Bpm1 BsrD1 BstB1
CATAAGAAACTTGGCAATGACAAGAATCAAATAATTAGCCTGTTCGAACAAAATCATAAA
721 --------+--------+--------+--------+--------+--------+ 780
GTATTCTTTGAACCGTTACTGTTCTTAGTTTATTAATCGGACAAGCTTGTTTTAGTATTT
H K K L G N D K N Q I I S L F E Q N H K
GCATGTATAGATCAATGGCTAGACTCAATGGAAACATGCTCTGTTGTCTACGTGTCGTTG
781 --------+--------+--------+--------+--------+--------+ 840
CGTACATATCTAGTTACCGATCTGAGTTACCTTTGTACGAGACAACAGATGCACAGCAAC
A C I D Q W L D S M E T C S V V Y V S L
Ear1 polyA
GGAAGTATTGCCTCGATTGGGAAAGAAGAGATGGAAGAACTGGCTTGTGGTTTATTGATG
841 --------+--------+--------+--------+--------+--------+ 900
CCTTCATAACGGAGCTAACCCTTTCTTCTCTACCTTCTTGACCGAACACCAAATAACTAC
G S I A S I G K E E M E E L A C G L L M
Eco57 SpDon Eco57 Hind3 SpAcc
AGCAACTGCTACTTCTTGTGGGTAGTAAGAGCTTCAGAACAGGACAAGCTTCCTTCAGAT
901 --------+--------+--------+--------+--------+--------+ 960
TCGTTGACGATGAAGAACACCCATCATTCTCGAAGTCTTGTCCTGTTCGAAGGAAGTCTA
S N C Y F L W V V R A S E Q D K L P S D
SpAcc Mfe1
TTCATGTCTTTGGCTTCTGAAAAAGGTCTAATCGTCAATTGGTGTTGTCAAACTGAAGTC
961 --------+--------+--------+--------+--------+--------+ 1020
AAGTACAGAAACCGAAGACTTTTTCCAGATTAGCAGTTAACCACAACAGTTTGACTTCAG
F M S L A S E K G L I V N W C C Q T E V
Eco57 EcoR1
TTGGCTCATCCTGCCGTGGCATGTTTTATGACGCACTGTGGATGGAATTCAACGCTCGAA
1021 --------+--------+--------+--------+--------+--------+ 1080
AACCGAGTAGGACGGCACCGTACAAAATACTGCGTGACACCTACCTTAAGTTGCGAGCTT
L A H P A V A C F M T H C G W N S T L E
SanD1 Bmr1 PflM1
GCAATTAGTTGTGGGGTCCCACTTGTCACTATGGCCCAGTGGGTGGACCAGCAGCCAAAT
1081 --------+--------+--------+--------+--------+--------+ 1140
CGTTAATCAACACCCCAGGGTGAACAGTGATACCGGGTCACCCACCTGGTCGTCGGTTTA
A I S C G V P L V T M A Q W V D Q Q P N
GCCAAGTGTGTTGAAGATTTGTGGAAAGTAGGCGTTCGGATTAAAGGTCCTGAGAATGGG
1141 --------+--------+--------+--------+--------+--------+ 1200
CGGTTCACACAACTTCTAAACACCTTTCATCCGCAAGCCTAATTTCCAGGACTCTTACCC



3.3.2.5.2 Characterization of BMGT2 cDNA sequences (FJ586245)

The full length cDNA sequence of BMGT2 gene has been submitted to NCBI GenBank database under accession no. FJ586245. The *Bacopa GT* gene (BMGT2) sequence contained a 1371 bp open-reading frame, which codes for a protein of 456 amino acids with a molecular weight of 50 kDa. The sequence analysis of BMGT2 gene confirmed the 46 bp 5'UTR, the putative initiation codon ATG, stop codon (TAA), 150 bp 3'UTR and a poly-A tail.

3.3.2.5.2a Nucleotide sequence of BMGT2

```
  1 ATGGAAGACG CCATTGTTCT CTACTCCTCC GCCGAACACC TCAATTCCAT
 51 GCTTGTCCCC GCCAAATTCA TCAGCAAACA CCACCCCTCC ATTTCCGTCA
101 TAATCATCAG CACCGCCGCC GAATCCGCCG CCGCCTCCGT CGCCTCTGTT
151 CCTTCAATCA CTTACCACCG CCTCCCCTCC GCCCCTCTTC CTCCAGACTT
201 AACCACCAGT ATCATCGAAC TGTTCTTCGA AATTCCCCGT TTCCACAACC
251 CATTCCTTCA CGAAGCCCTT CTCGAAATCT CTCAGAAATC AAATCTCAGA
301 GCATTTCTCA TCGATTTCTT CTGCAATTCA ACTTTTGAAG TCTCCACAAG
351 CTTGAATATA CCCACCTACT TCTACCTCAG CGGCGGCGCG TGCGGCCTCT
401 GCGCGCTTCT GTATTTCCCG ACCATCGACG AAGCCGTCAG TCCTCGAGAT
451 ATCGGAGAAT TGAACGATTT TCTTGAAATT CCGGGCTGTC CACCGGTTCA
```

```
 501 CTCGTTGGAT TTTCCGAAAG CCATGTGGTT TCGCCGGAGC AATACGTACA
 551 AACATTTTCT TGACACCGCC GGAAATATGC GTAGAGCGAG TGGAATCGTC
 601 ACGAACTCGT TCGACGCGAT CGAGTTCAGA GCTAAAGAAG CTTTGTCAAA
 651 CAGCTTATGC ACTCCCGGTC TCGCAACGCC GCCGGTTTAC GTCATCGGAC
 701 CTTTGGTTGC TGAAACGAAC AGGAAAAATG GCGGCGAGGA GCATGAATGC
 751 CTGAAATGGC TCGATTCTCA GCCAATTAAA AGCGTGATTT TCCTCTGTTT
 801 TGGTCGACGT GGCCTATTTT CAGCCGCCCA ATTGAAAGAA ATGGCGATCG
 851 GGCTTGAAAA CAGTGGCCAC CGGTTTCTCT GGTCTGTCCG CAGCCCGCCC
 901 GGCCCGGCGG CGGCGAAAGA TCCCGACTTG GACGCTTTGC TTCCGGAAGG
 951 TTTTATGGAG AGGACGAAAG ATAGGGGATT TGTTATAAAG ACGTGGGCCC
1001 CGCAGAAGGA GGTGCTCAGC CACGAGGCCG TCGGCGGGTT CGTGACTCAC
1051 TGCGGCCGGA GCTCTGTTTT GGAAGCGGTG TCGTTTGGGG TTCCGATGAT
1101 TGGGTGGCCG ATGTACGCGG AGCAGAGGAT GCAGCGTGTT TTCATGGTGG
1151 AGGAAATGAA GGTGGCGCTG CCGCTTGCGG AGGAGGCGGA CGGGTTCGTG
1201 ACGGCGGGGG AGCTGGAGAA GCGAGTGAGA GAGTTGATGG GGTTGCCGGC
1251 GGGAAAGGCC GTGACGCAGC GCGTGGCGGA ATTGAGGACG GCGGCCGAGG
1301 CTGCGGTGCG GAAGGGTGGA TCGTCTGTTG TTGCTTTAGG AAAGTTCATT
1351 GAGACGGTGA CCCGACGCTA A
```

3.3.2.5.2b Deduced amino acid sequence of BMGT2

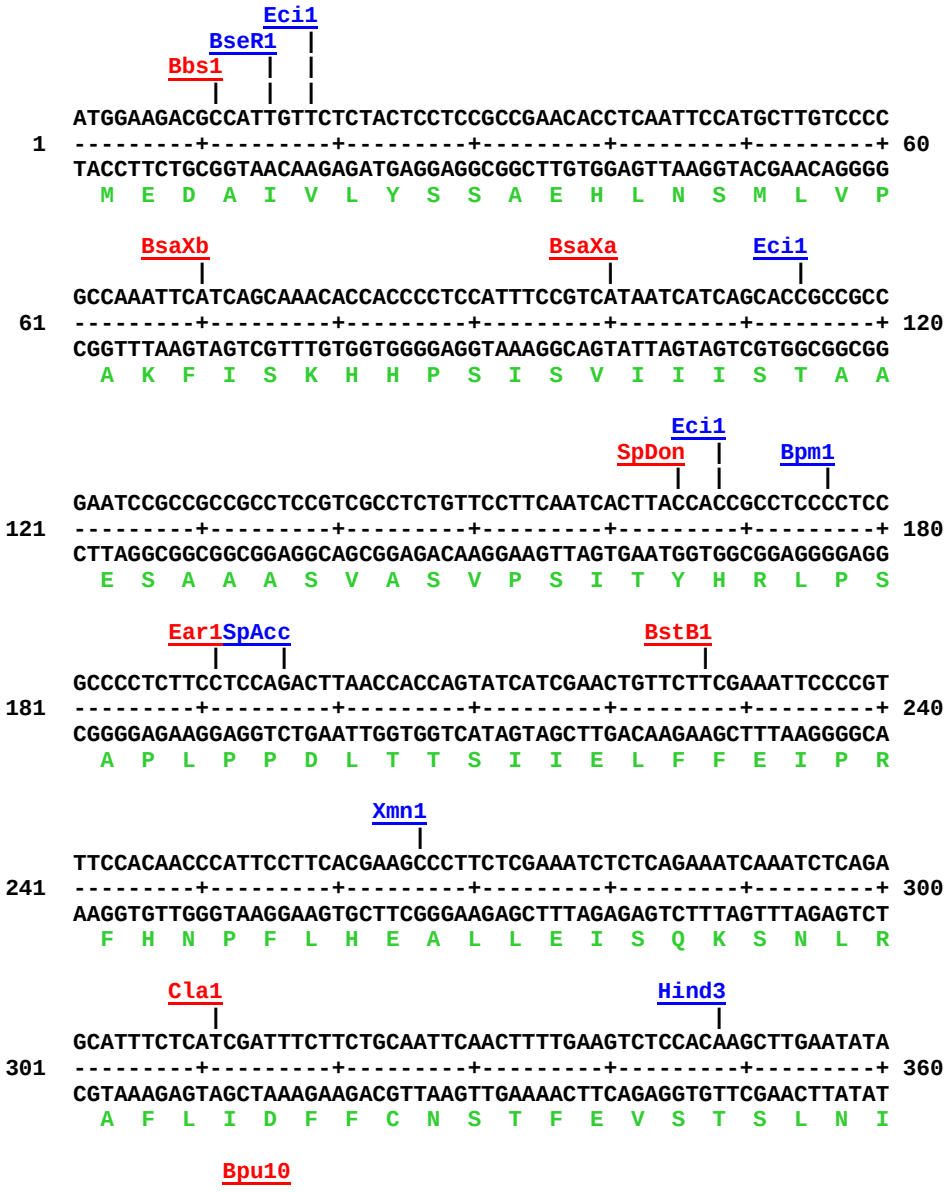
M E D A I V L Y S S A E H L N S M L V P A K F I S K H H P S I S V I I I S T A A E S A
A A S V A S V P S I T Y H R L P S A P L P P D L T T S I I E L F F E I P R F H N P F L H
E A L L E I S Q K S N L R A F L I D F F C N S T F E V S T S L N I P T Y F Y L S G G A
C G L C A L L Y F P T I D E A V S P R D I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

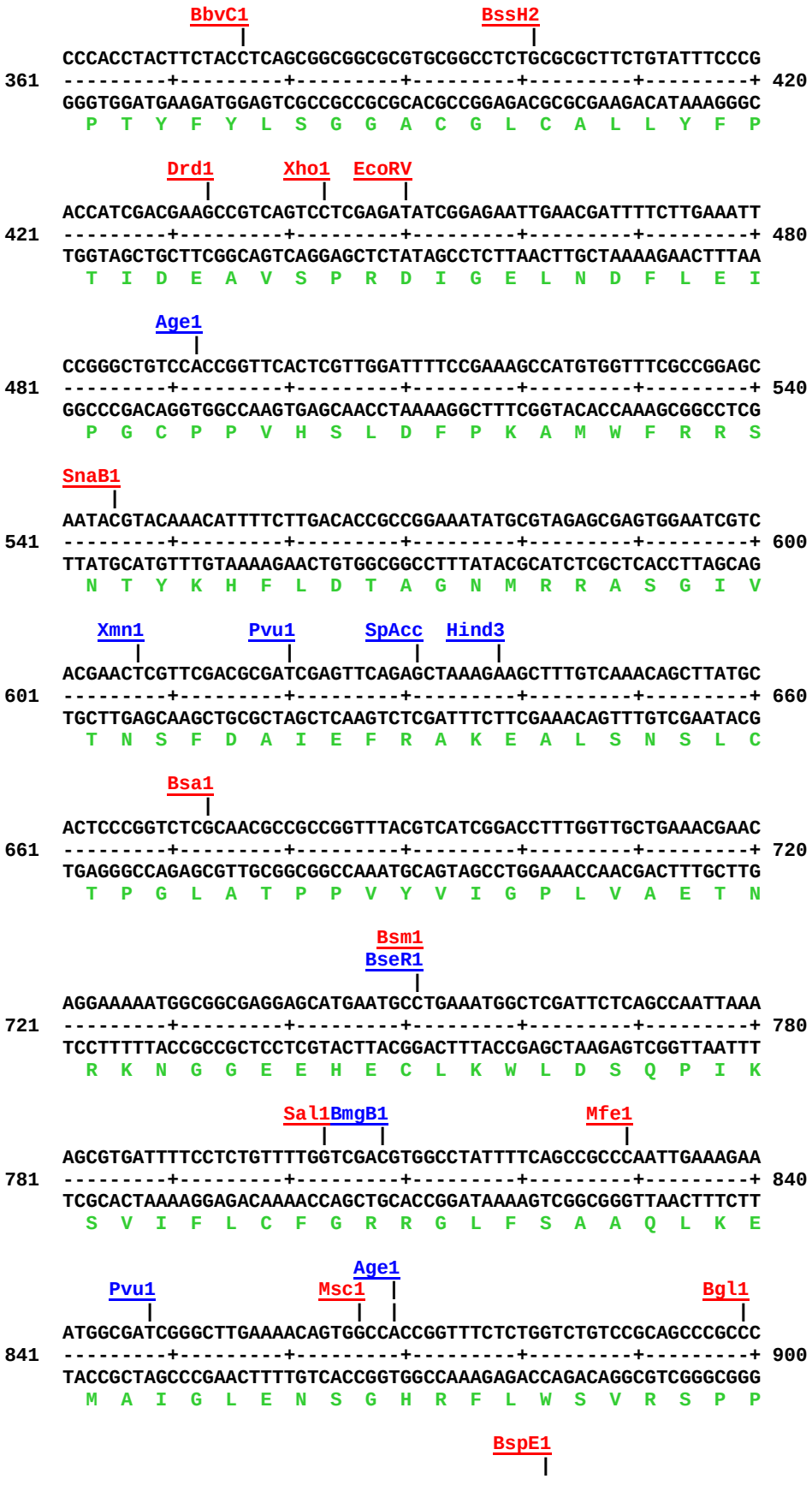
Analysis of the deduced amino acid sequence of BMGT2 gene showed that it contained the donor binding site (bold and underlined) and the 46 bp PSPG box (highlighted in green). It

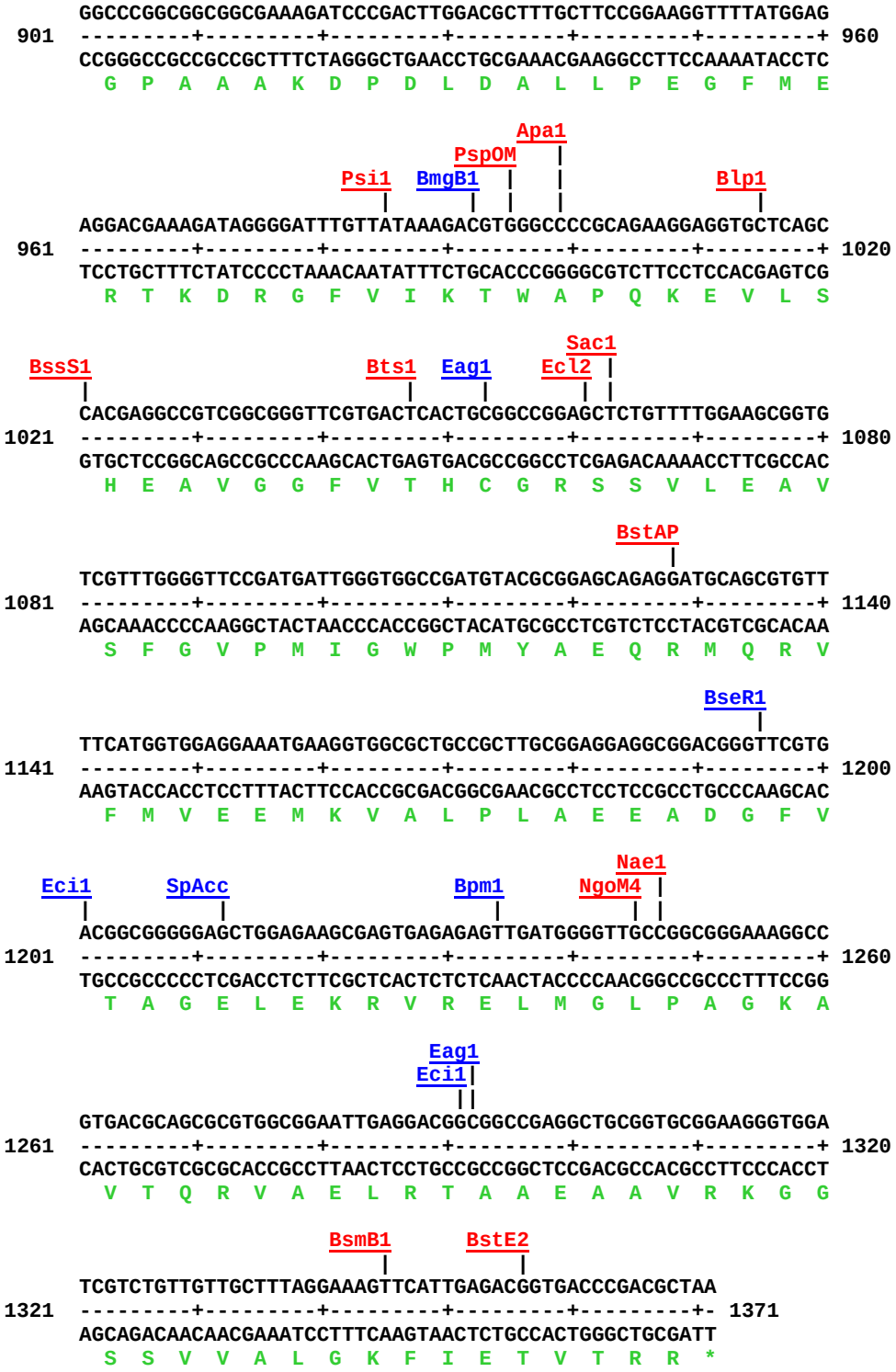
also contained the arginine residue in place of tryptophan residue in the conserved region of PSPG box (HCGRSS) which is responsible for glucuronic acid specificity instead of glucose.

3.3.2.5.2c Restriction map of BMGT2 gene

Restriction map of BMGT2 gene was created by using an online bioinformatic tool, Webmap DNA. Map created at http://pga.mgh.harvard.edu/web_apps/web_map/start . Editor at http://pga.mgh.harvard.edu/web_apps/dna_utilities.html Sites protected by *E. coli* methylation are not shown in this map.




BbvC1
BssH2
CCCACCTACTTCTACCTCAGCGGCGGCGCGTGCGGCCTCTGCGCGCTTCTGTATTTCCCG
361 --------+---------+---------+---------+---------+---------+ 420
GGGTGGATGAAGATGGAGTCGCCGCCGCGCACGCCGGAGACGCGCGAAGACATAAAGGGC
P T Y F Y L S G G A C G L C A L L Y F P
Drd1
Xho1
EcoRV
ACCATCGACGAAGCCGTCAGTCCTCGAGATATCGGAGAATTGAACGATTTTCTTGAAATT
421 --------+---------+---------+---------+---------+---------+ 480
TGGTAGCTGCTTCGGCAGTCAGGAGCTCTATAGCCTCTTAACTTGCTAAAAGAACTTTAA
T I D E A V S P R D I G E L N D F L E I
Age1
CCGGGCTGTCCACCGGTTCACTCGTTGGATTTTCCGAAAGCCATGTGGTTTCGCCGGAGC
481 --------+---------+---------+---------+---------+---------+ 540
GGCCCGACAGGTGGCCAAGTGAGCAACCTAAAAGGCTTTCGGTACACCAAAGCGGCCTCG
P G C P P V H S L D F P K A M W F R R S
SnaB1
AATACGTACAAACATTTTCTTGACACCGCCGGAAATATGCGTAGAGCGAGTGGAATCGTC
541 --------+---------+---------+---------+---------+---------+ 600
TTATGCATGTTTGTAAAAGAACTGTGGCGGCCTTTATACGCATCTCGCTCACCTTAGCAG
N T Y K H F L D T A G N M R R A S G I V
Xmn1
Pvu1
SpAcc
Hind3
ACGAACTCGTTCGACGCGATCGAGTTCAGAGCTAAAGAAGCTTTGTCAAACAGCTTATGC
601 --------+---------+---------+---------+---------+---------+ 660
TGCTTGAGCAAGCTGCGCTAGCTCAAGTCTCGATTTCTTCGAAACAGTTTGTCGAATACG
T N S F D A I E F R A K E A L S N S L C
Bsa1
ACTCCCGGTCTCGCAACGCCGCCGGTTTACGTCATCGGACCTTTGGTTGCTGAAACGAAC
661 --------+---------+---------+---------+---------+---------+ 720
TGAGGGCCAGAGCGTTGCGGCGGCCAAATGCAGTAGCCTGGAAACCAACGACTTTGCTTG
T P G L A T P P V Y V I G P L V A E T N
Bsm1
BseR1
AGGAAAAATGGCGGCGAGGAGCATGAATGCCTGAAATGGCTCGATTCTCAGCCAATTAAA
721 --------+---------+---------+---------+---------+---------+ 780
TCCTTTTTACCGCCGCTCCTCGTACTTACGGACTTTACCGAGCTAAGAGTCGGTTAATTT
R K N G G E E H E C L K W L D S Q P I K
Sal1
BmgB1
Mfe1
AGCGTGATTTTCCTCTGTTTTGGTCGACGTGGCCTATTTTCAGCCGCCCAATTGAAAGAA
781 --------+---------+---------+---------+---------+---------+ 840
TCGCACTAAAAGGAGACAAAACCAGCTGCACCGGATAAAAGTCGGCGGGTTAACTTTCTT
S V I F L C F G R R G L F S A A Q L K E
Age1
Pvu1
Msc1
Bgl1
ATGGCGATCGGGCTTGAAAACAGTGGCCACCGGTTTCTCTGGTCTGTCCGCAGCCCGCCC
841 --------+---------+---------+---------+---------+---------+ 900
TACCGCTAGCCCGAACTTTTGTCACCGGTGGCCAAAGAGACCAGACAGGCGTCGGGCGGG
M A I G L E N S G H R F L W S V R S P P
BspE1


GGCCCGGCGGCGGCGAAAGATCCCGACTTGGACGCTTTGCTTCCGGAAGGTTTTATGGAG
901 --------+---------+---------+---------+---------+---------+ 960
CCGGGCCGCCGCCGCTTTCTAGGGCTGAACCTGCGAAACGAAGGCCTTCCAAAATACCTC
G P A A A K D P D L D A L L P E G F M E
Apa1
PspOM
Psi1 BmgB1 Blp1
AGGACGAAAGATAGGGGATTTGTTATAAAGACGTGGGCCCCGCAGAAGGAGGTGCTCAGC
961 --------+---------+---------+---------+---------+---------+ 1020
TCCTGCTTTCTATCCCCTAAACAATATTTCTGCACCCGGGGCGTCTTCCTCCACGAGTCG
R T K D R G F V I K T W A P Q K E V L S
Sac1
BssS1 Bts1 Eag1 Ecl2
CACGAGGCCGTCGGCGGGTTCGTGACTCACTGCGGCCGGAGCTCTGTTTTGGAAGCGGTG
1021 --------+---------+---------+---------+---------+---------+ 1080
GTGCTCCGGCAGCCGCCCAAGCACTGAGTGACGCCGGCCTCGAGACAAAACCTTCGCCAC
H E A V G G F V T H C G R S S V L E A V
BstAP
TCGTTTGGGGTTCCGATGATTGGGTGGCCGATGTACGCGGAGCAGAGGATGCAGCGTGTT
1081 --------+---------+---------+---------+---------+---------+ 1140
AGCAAACCCCAAGGCTACTAACCCACCGGCTACATGCGCCTCGTCTCCTACGTCGCACAA
S F G V P M I G W P M Y A E Q R M Q R V
BseR1
TTCATGGTGGAGGAAATGAAGGTGGCGCTGCCGCTTGCGGAGGAGGCGGACGGGTTCGTG
1141 --------+---------+---------+---------+---------+---------+ 1200
AAGTACCACCTCCTTTACTTCCACCGCGACGGCGAACGCCTCCTCCGCCTGCCCAAGCAC
F M V E E M K V A L P L A E E A D G F V
Nae1
Eci1 SpAcc Bpm1 NgoM4
ACGGCGGGGGAGCTGGAGAAGCGAGTGAGAGAGTTGATGGGGTTGCCGGCGGGAAAGGCC
1201 --------+---------+---------+---------+---------+---------+ 1260
TGCCGCCCCCTCGACCTCTTCGCTCACTCTCTCAACTACCCCAACGGCCGCCCTTTCCGG
T A G E L E K R V R E L M G L P A G K A
Eag1
Eci1
GTGACGCAGCGCGTGGCGGAATTGAGGACGGCGGCCGAGGCTGCGGTGCGGAAGGGTGGA
1261 --------+---------+---------+---------+---------+---------+ 1320
CACTGCGTCGCGCACCGCCTTAACTCCTGCCGCCGGCTCCGACGCCACGCCTTCCCACCT
V T Q R V A E L R T A A E A A V R K G G
BsmB1 BstE2
TCGTCTGTTGTTGCTTTAGGAAAGTTCATTGAGACGGTGACCCGACGCTAA
1321 --------+---------+---------+---------+---------+- 1371
AGCAGACAACAACGAAATCCTTTCAAGTAACTCTGCCACTGGGCTGCGATT
S S V V A L G K F I E T V T R R *

3.3.2.5.3 CLUSTAL W (1.8) multiple sequence alignment of nucleotide sequences of BMGT1 (FJ586244) and BMGT2 (FJ586245)

The start codon and stop codon of both the sequences is highlighted in red.

```
FJ586244        ATGGAGAGCAAAGGAACAGGGAAGGAGGCTCATATTCTCGTATTTCCTTACCCTGCAC-- 58
FJ586245        ------------------ATGGAAGACGCCATTGTTCTC-TACTCCTCCGCCGAACACCT 41
                                  * * ** **   * ***** ** * *    **   ***

FJ586244        -AAGGCCACATTAATCCC-GTGTTACCATTCTCCAAATTCCTAGCCTCAAAGGGTCTAAA 116
FJ586245        CAATTCCATGCTTGTCCCCGCCAAATTCATCAGCAAACACCACCCCTCCATTTCCGTCAT 101
                 **  ***   *  **** *    *    **  ****  **   **** *      * *

FJ586244        GGTTACCATTATCGTGACC---CCTAGCGTAAAAAAGCTGGTAAATTTCCCCCCAAACAG 173
FJ586245        AATCATCAGCACCGCCGCCGAATCCGCCGCCGCCTCCGTCGCCTCTGTTCCTTCAATCAC 161
                  * * **  * **   **    *   **         * *    * * **  *** **

FJ586244        CTCCATTAGCATAGAAAGGATATCGGATGGTTCAGAAGACGTTAAGGAGACCGAAGACAT 233
FJ586245        TTACCACCGCCTCCCCTCCGCCCCTCTTCCTCCAGAC----TTAACCA--CCAGTATCAT 215
                 * *    ** *           *   *  * ****     ****  *  **     ***

FJ586244        CGAAGCCTATTTCAA---CCGCTTTAGGAGAGAGGCTTCCCAGAATTTAGCCAAATTC-- 288
FJ586245        CGAACTGTTCTTCGAAATTCCCCGTTTCCACAACCCATTCCTTCACGAAGCCCTTCTCGA 275
                ****   *  *** *    * *  *       *  * * **   *   ****    **

FJ586244        -ATCGATGAAAAGAAAGGTTGGGGGGCTAAAGTCATCGTGTACGACTCCACGATGCCATG 347
FJ586245        AATCTCTCAGAAATCAAATCTCAGAGCATTTCTCATCGATTTC--TTCTGCAATTCAAC- 332
                 ***  * * **   *  *    * **     ******  * *   **  * ** * *

FJ586244        GGTTTTGGACATAGCCCACGAAAGAGGTCTACTTGGAGCTTCTTTTTTCACTCAGTCTTG 407
FJ586245        --TTTTGAAGTCTCCACAAGCTTGAATATACCCACCTACTTCTACCTCAGCGGCGGCGCG 390
                  ***** *     * ** *   **      *      *****   *   *   * *  *

FJ586244        TTTTGTCAGCGCCGTCTACTGCCATTTGCATCAAGGTACTCTGAAATATCCGTACGAGGA 467
FJ586245        TGCGGCCTCTGCGCGCTTCTGTATTTCCCGACCATCGACG--AAGCCGTCAGTCCTCGAG 448
                *   * *   **   ** ***   **  *  * *   **    *    ** ** *  *

FJ586244        GGAGGAGAAATCTACGTTGCTGTCCTTACATCCGCTGCTGCCTAC-GCTTCAGATTAATG 526
FJ586245        ATATCGGAGAATTGAACGATTTTCTTGAAATTCCGGGCTGTCCACCGGTTCACTCGTTGG 508
                  *   ** *  *       * ** * * ** *   **** * ** * ****       *

FJ586244        ATTTGCCTTGTTTTTCCAAGTTTGATGATCCCAAACATTTGGTCTCCAAACATCTAACCG 586
FJ586245        ATTTTCCGAAAGCCATGTGGTTTCG----CCGGAGCAATACGTA--CAAACATTTTCTTG 562
                **** **           ****      **  * ** *  **   ******* *    *

FJ586244        ATCAATTCATCAATCTGGACAAAGTGGACTGGATCCTTTTCAACACTTTCTATGACTTGG 646
FJ586245        ACACCGCCGGAAATATGCGTAGAGCGAGTGGAATCGTCACGAACTCGTTCGACGCGATCG 622
                *      *   *** **   * ** *    * *** *    *** * *** * *   * *

FJ586244        AGACTCAGGTTGCGGAATGGATGAAAGCCAAATGGCCTATCAAAACTATTGGACCTACTT 706
FJ586245        AGTTCAGAGCTAAAGAAGCTTTG----TCAAACAGCTTAT--GCACTCCCGGTCTCGCAA 676
                **      * *   ***    **     ****  ** ***    ***   ** *   *

FJ586244        CATTACTGGAGAAA--CATAAGAAACTTGGCAA-TGACAAGAATCAAATAATTAGCCTGT 763
FJ586245        CGCCGCCGGTTTACGTCATCGGACCTTTGGTTGCTGAAACGAACAGGAAAAATGGCGGCG 736
                *    * **   *   ***  **   ****    *** * ***    * ** * **

FJ586244        TCGAACAAAATCATAAAGCATGTATAGATCAATGGCTAGACTCAA-TGGAAACATGCTCT 822
FJ586245        AGGAGCATGAATGCCTGAAATGGCTCGATTCTCAGCCAATTAAAAGCGTGATTTTCCTCT 796
                  ** **  *         ***  * ***     ** *     **  *  *   * ****

FJ586244        GTTGTCTACG-TGTCGTTGGGAAGTATTGCCTCGATTGGGAAAGAAGAGA--------TG 873
FJ586245        GTTTTGGTCGACGTGGCCTATTTTCAGCCGCCCAATTGAAAGAAATGGCGATCGGGCTTG 856
                *** *   **  ** *         *    * * ****  * * * *           **
```

```
FJ586244        GAAGAACTGGCTTGTGGTTT-----ATTGATGAGCAACTGCTACTTCTTGTGGGTAGTAA 928
FJ586245        AAAACAGTGGCCACCGGTTTCTCTGGTCTGTCCGCAGCCCGCCCGGCCCGGCGGCGGCGA 916
                 **  * ****    *****      *   *  *** *     *  *  *  **  *  *

FJ586244        GAGCTTCAGAACAGGACAA---GCTTCCTTCAGATTTCATGTCTTTGGCTTCTGAAAAAG 985
FJ586245        AAGATCCCGACTTGGACGCTTTGCTTCCGGAAGGTTTTATGGAGAGGACGAAAGATAGGG 976
                 ** * * **   ****     ******   ** *** ***     * *    ** *  *

FJ586244        GTCTAATCGTCAATTGGTG---TTGTCAAACTGAAGTCTTGGCTCATCCTGCCGTGGCAT 1042
FJ586245        GATTTGTTATAAAGACGTGGGCCCCGCAGAAGGAGGTGCTCAGCCACGAGGCCGTCGGCG 1036
                *  *  *  * **   ***       ** *  ** **  *    **    ***** *

FJ586244        GTTTTATGACGCACTGTGGATGGAATTCAACGCTCGAAGCAATTAGTTGTGGGGTCCCAC 1102
FJ586245        GGTTCGTGACTCACTGCGGCCGGAGCTCTGTTTTGGAAGCGGTGTCGTTTGGGGTTCCGA 1096
                * **  **** ***** **  ***  **     * *****  *    * ****** **

FJ586244        TTGTCACTATGGCCCA-GTGGGTGGACCAGCAGCCAAATGCCAAGTGTGTTGAAGATTTG 1161
FJ586245        TGATTGG-GTGGCCGATGTACGCGGAGCAGAGG----ATGCAGCGTGTTTTCATGGTGGA 1151
                *  *     ***** * **  * *** ***  *    ****   **** ** * * *

FJ586244        TGGAAAGTAGGCG----TTCGGATTAAAGGTCCTGAGAATGGGACTTTTGAGAGAGAAGA 1217
FJ586245        GGAAATGAAGGTGGCGCTGCCGCTTGCGGAGGAGGCGGACGGG-TTCGTGACGGCGGGGG 1210
                 * ** * *** *    * * * **   *     * * * ***  *  ***  * *  *

FJ586244        AATTGCTAGATGCATACAACA-AGTCATCGGAGGAGATAATGCAGATGAGCTTAGAGCAA 1276
FJ586245        AGCTGG-AGAAGCGAGTGAGAGAGTTGATGGGGTTGCCGGCGGGAAAGGCCGTGACGCAG 1269
                *  **  *** **     * * ***    ** *  *     *   * *  * *   ***

FJ586244        -ATGCTTGGAAATGGAAAAAATTGGCTCAAGATGCTATGGAGGAAAATGGAAATTCTACT 1335
FJ586245        CGCGTGGCGGAATTGAGGACGGCGGCCGAGGCTGCGGTGCGGAAGGGTGGATCGTCTGTT 1329
                   *    * *** **  *    ***  * * ***  **  * *   ****   ***  *

FJ586244        AAAAATATTGAGGATTTTGTTGTG-CAGTTTTTTAACATGTCCCTGCTCCTAACTTGA 1392
FJ586245        GTTGCTTTAGGAAAGTTCATTGAGACGGTGACCCGACGCTAA--------------- 1371
                     * * *   * **  *** * * **      **
```

3.3.2.5.4 CLUSTAL W (1.8) multiple sequence alignment of amino acid sequences of BMGT1 (FJ586244) and BMGT2 (FJ586245)

```
FJ586244        MESKGTGKEAHILVFPYPAQGHINPVLPFSKFLASKG--LKVTIIVTPSVKKLVNFPPNS 58
FJ586245        ---------MEDAIVLYSSAEHLNSMLVPAKFISKHHPSISVIIISTAAESAAASVASVP 51
                         .  :. *.:  *:*.:*  :**::.:   :.* ** *.: .  ..... .

FJ586244        SISIERISDGSED-VKETEDIEAYFNRFRREASQNLAKFIDEKKGWGAKVIVYDSTMPWV 117
FJ586245        SITYHRLPSAPLPPDLTTSIIELFFEIPRFHNPFLHEALLEISQKSNLRAFLIDFFCNST 111
                **: .*:....      *. ** :*:  * . .     ::: .:  . :.:: *     .

FJ586244        LDIAHERGLLGASFFTQSCFVSAVYCHLHQGTLKYPYEEEEKSTLLSLHPLLPTLQINDL 177
FJ586245        FEVSTSLNIP--TYFYLSGGACGLCALLYFPTIDEAVSPRDIGELNDFLEIPGCPPVHSL 169
                :::: . .:   ::*  *  ...: . *:  *:. . . .: . * .:  :     ::.*

FJ586244        PCFSKFDDPKHLVSKHLTDQFINLDKVDWILFNTFYDLETQVAEWMKAKWPIKTIGPTSL 237
FJ586245        DFPKAMWFRRSNTYKHFLDTAGNMRRASGIVTNSFDAIEFRAKEALSNSLCTPGLATPPV 229
                   . :   :  . **: *   *: :.. *: *:*  :* :. * :. .     :....:

FJ586244        LEKHKKLGNDKNQIISLFEQNHKACIDQWLDSMETCSVVYVSLGSIASIGKEEMEELACG 297
FJ586245        YVIGPLVAETNR-----KNGGEEHECLKWLDSQPIKSVIFLCFGRRGLFSAAQLKEMAIG 284
                      :.: :.      : ..:    :****    **::::.*  . :.  :::*:* *

FJ586244        LLMSNCYFLWVVR-----ASEQDK-----LPSDFMSLASEKGLIVN-WCCQTEVLAHPAV 346
FJ586245        LENSGHRFLWSVRSPPGPAAAKDPDLDALLPEGFMERTKDRGFVIKTWAPQKEVLSHEAV 344
                *  *.  *** **     *: :*      **..**. :.::*:::: *. *.***:* **

FJ586244        ACFMTHCGWNSTLEAISCGVPLVTMAQWVDQQPNAKCVEDLWKVGVRIKGPENGTFEREE 406
FJ586245        GGFVTHCGRSSVLEAVSFGVPMIGWPMYAEQRMQRVFMVEEMKVALPLAEEADGFVTAGE 404
                . *:**** .*.***:* ***::  . :.:*: :   : :  **.: :    :* .   *
```

```
FJ586244        IARCIQQVIGGDNADELRANAWKWKKLAQDAMEENGNSTKNIEDFVVQFFNMSLLLT 463
FJ586245        LEKRVRELMGLPAGKAVTQRVAELRTAAEAAVRKGGSSVVALGKFIETVTRR----- 456
                : : :::::*   .. :  .. : :. *: *:.:.*.*.  : .*:  . .
```

Both glycosyltransferase genes show very less homology at nucleotide as well as amino acid level.

3.3.2.5.5 Amino acid composition, theoretical pI and molecular weight of BMGT1 and BMGT2 genes

The deduced amino acid sequences of BMGT1 gene show the coding region consists of 463 amino acids, with theoretical molecular weight of 52347.1 Daltons. The theoretical pI value of BMGT1 gene is 5.41 (http://www.expasy.ch/cgibin/protparam). The amino acid composition of BMGT1 gene is given in table 3.1. In BMGT1 gene the total number of negatively charged residues (Asp + Glu) equals 60 and positively charged residues (Arg + Lys) equal 46. The empirical formula of BMGT1 gene is $C_{2358}H_{3653}N_{607}O_{690}S_{25.}$

Table 3.1 Amino acid composition of BMGT1 (FJ586244)

Amino acid	Total number	Percentage	Amino acid	Total number	Percentage
Ala (A)	29	6.3%	Leu (L)	45	9.7%
Arg (R)	10	2.2%	Lys (K)	36	7.8%
Asn (N)	22	4.8%	Met (M)	11	2.4%
Asp (D)	24	5.2%	Phe (F)	23	5.0%
Cys (C)	14	3.0%	Pro (P)	20	4.3%
Gln (Q)	19	4.1%	Ser (S)	34	7.3%
Glu (E)	36	7.8%	Thr (T)	22	4.8%
Gly (G)	24	5.2%	Trp (W)	13	2.8%
His (H)	12	2.6%	Tyr (Y)	9	1.9%
Ile (I)	28	6.0%	Val (V)	32	6.9%

Similarly deduced amino acid composition of BMGT2 gene is 456 amino acids with theoretical molecular weight of 50047 Daltons. The theoretical pI value of BMGT2 gene is 6.21. The amino acid composition of BMGT2 gene is given in table 3.2. In BMGT2 gene the total number of negatively charged residues (Asp + Glu) equals 50 and positively charged residues (Arg + Lys) equal 45. The empirical formula of BMGT2 gene is $C_{2256}H_{3530}N_{602}O_{645}S_{20.}$

Table 3.2 Amino acid composition of BMGT2 (FJ586245)

Amino acid	Total number	Percentage	Amino acid	Total number	Percentage
Ala (A)	46	10.1%	Leu (L)	44	9.6%
Arg (R)	26	5.7%	Lys (K)	19	4.2%
Asn (N)	13	2.9%	Met (M)	12	2.6%
Asp (D)	15	3.3%	Phe (F)	28	6.1%
Cys (C)	8	1.8%	Pro (P)	31	6.8%
Gln (Q)	7	1.5%	Ser (S)	38	8.3%
Glu (E)	35	7.7%	Thr (T)	22	4.8%
Gly (G)	30	6.6%	Trp (W)	5	1.1%
His (H)	12	2.6%	Tyr (Y)	8	1.8%
Ile (I)	25	5.5%	Val (V)	32	7.0%

3.3.2.5.6 Hydropathy index of both the glycosyltransferase 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 native 3-dimensional structure) while hydrophilic amino acids are more commonly found towards the protein surface. The amino acid sequences of BMGT1 (FJ586244) and BMGT2 (FJ586245) were analyzed using Kyte-Doolittle Hydropathy plot at http://gcat.davidson.edu/rakarnik/kd.cgi (Fig 3.23 A & Fig 3.23 B respectively) window size 9. When the window size is 9, strong negative peaks indicate possible surface regions of globular proteins.

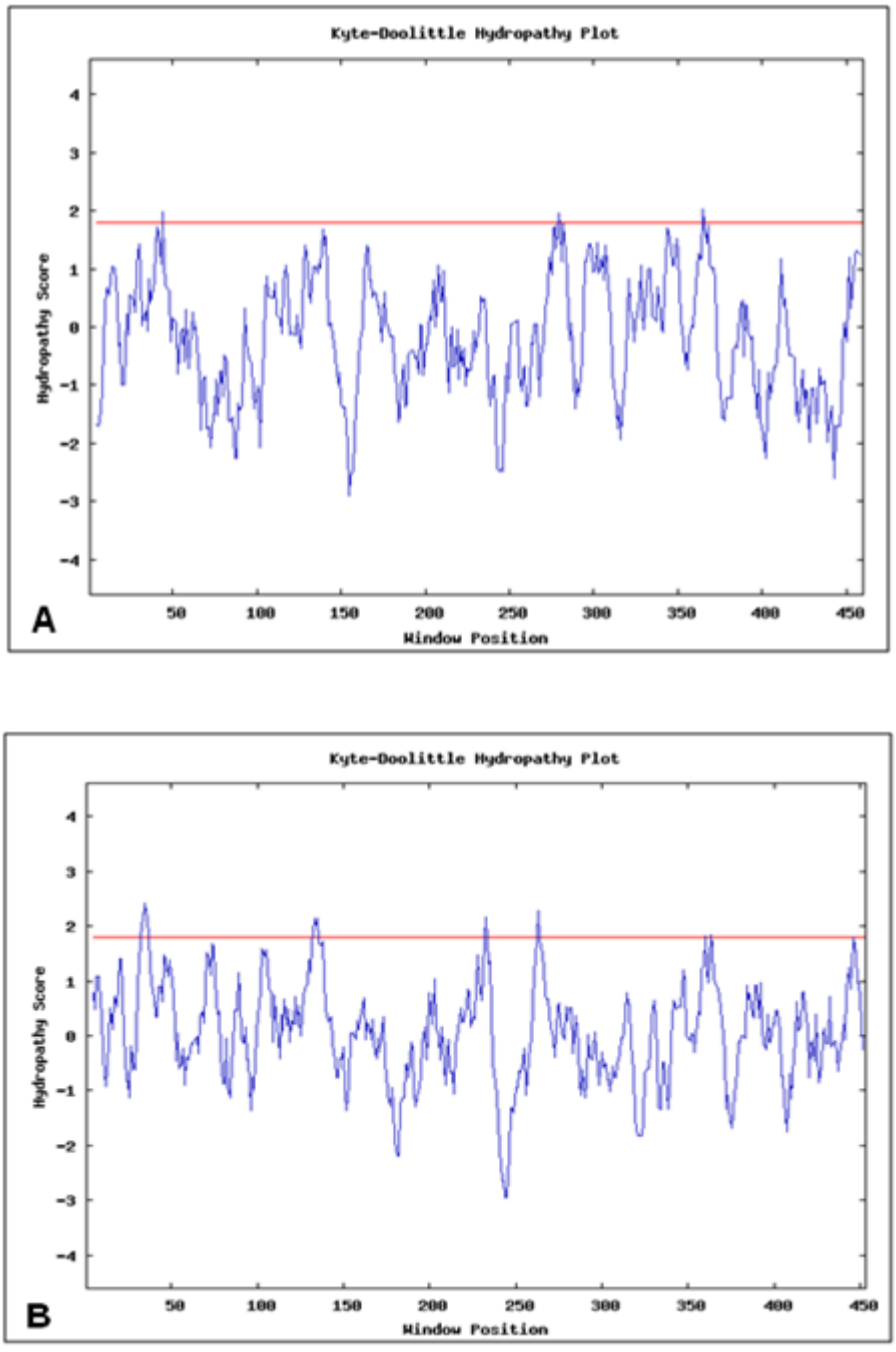


Fig. 3.23: Hydropathic plot. A Kyte-Doolittle Hydropathy for BMGT1 (FJ586244) and **B** KyteDoolittle Hydropathy for BMGT2 (FJ586245). Window size of 9 suggested a good value for finding putative surface-exposed regions. When the window size is **9**, strong negative peaks indicate possible surface regions of globular proteins. Both the *GTs* showed almost similar pattern in (Kyte-Doolittle plot).

3.3.2.5.7 Codon usage of both the BMGT1 and BMGT2

The GC% and codon usage of both, the cDNA clones BMGT1 (FJ586244) and BMGT2 (FJ586245) was calculated using online software (www.justbio.com). The GC content of the cDNA clone, BMGT1 (FJ586244) is 43.14% and for BMGT2 (FJ586245) is 55.26%. The codon usage for *GT* cDNA clone, BMGT1 (FJ586244) is given in Table 3.3 and codon usage of BMGT2 (FJ586245) is given in table 3.4. It is expressed as % of total codons. Standard genetic codes were used for determining codon usage.

Table 3.3 Codon usage of BMGT1 (FJ586244)

codon	mean	codon	mean	codon	mean	codon	Mean
UUU ( )	4.70	UCU ( )	1.34	UAU ( )	2.01	UGU ( )	1.57
UUC ( )	2.24	UCC ( )	1.57	UAC ( )	1.34	UGC ( )	2.01
UUA ( )	0.00	UCA ( )	2.46	UAA ( )	2.24	UGA ( )	2.91
UUG ( )	2.01	UCG ( )	0.22	UAG ( )	1.57	UGG ( )	2.24
CUU ( )	1.34	CCU ( )	1.12	CAU ( )	2.91	CGU ( )	2.24
CUC ( )	1.79	CCC ( )	1.12	CAC ( )	0.89	CGC ( )	0.22
CUA ( )	1.12	CCA ( )	1.34	CAA ( )	3.80	CGA ( )	2.01
CUG ( )	0.89	CCG ( )	0.22	CAG ( )	0.45	CGG ( )	0.22
AUU ( )	1.57	ACU ( )	1.12	AAU ( )	1.34	AGU ( )	1.34
AUC ( )	0.45	ACC ( )	0.45	AAC ( )	1.34	AGC ( )	2.24
AUA ( )	0.22	ACA ( )	2.01	AAA ( )	3.36	AGA ( )	4.03
AUG ( )	2.24	ACG ( )	0.00	AAG ( )	0.89	AGG ( )	1.79
GUU ( )	1.34	GCU ( )	2.24	GAU ( )	2.01	GGU ( )	1.79
GUC ( )	1.12	GCC ( )	1.12	GAC ( )	1.12	GGC ( )	1.79
GUA ( )	0.45	GCA ( )	0.89	GAA ( )	3.13	GGA ( )	4.03
GUG ( )	1.12	GCG ( )	0.00	GAG ( )	0.67	GGG ( )	0.67

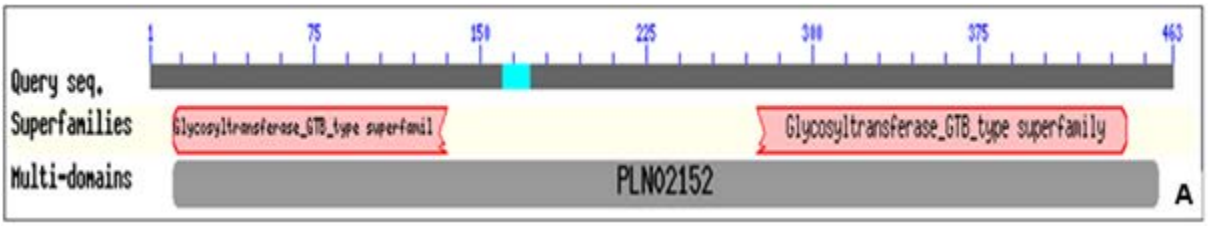
Table 3.4 Codon usage of BMGT2 (FJ586245)

codon	mean	codon	mean	codon	mean	codon	mean
UUU ( )	3.64	UCU ( )	2.05	UAU ( )	1.36	UGU ( )	1.82
UUC ( )	1.14	UCC ( )	2.73	UAC ( )	0.23	UGC ( )	0.91
UUA ( )	0.45	UCA ( )	1.14	UAA ( )	0.45	UGA ( )	1.59
UUG ( )	0.00	UCG ( )	0.68	UAG ( )	0.45	UGG ( )	1.82
CUU ( )	1.59	CCU ( )	1.82	CAU ( )	2.50	CGU ( )	2.50
CUC ( )	2.05	CCC ( )	2.05	CAC ( )	1.82	CGC ( )	3.64

CUA ( )	0.68	CCA ( )	1.36	CAA ( )	1.36	CGA ( )	3.64
CUG ( )	1.36	CCG ( )	1.59	CAG ( )	2.73	CGG ( )	3.18
AUU ( )	1.82	ACU ( )	0.23	AAU ( )	2.05	AGU ( )	0.45
AUC ( )	0.68	ACC ( )	0.68	AAC ( )	0.91	AGC ( )	2.05
AUA ( )	0.00	ACA ( )	0.45	AAA ( )	1.82	AGA ( )	1.82
AUG ( )	0.68	ACG ( )	0.45	AAG ( )	0.45	AGG ( )	0.45
GUU ( )	2.73	GCU ( )	2.05	GAU ( )	1.82	GGU ( )	2.27
GUC ( )	0.91	GCC ( )	1.59	GAC ( )	2.05	GGC ( )	3.41
GUA ( )	0.68	GCA ( )	1.14	GAA ( )	3.41	GGA ( )	3.64
GUG ( )	0.91	GCG ( )	1.14	GAG ( )	1.59	GGG ( )	1.36

3.3.2.5.8 Analysis of amino acid sequence of the *B. monniera* glycosyltransferase genes, BMGT1 (FJ586244) and BMGT2 (FJ586245) for their conserved domain

The amino acid sequences of BMGT1 (FJ586244) and BMGT2 (FJ586245) when analysed for their sequence similarity, it matched with glycosyltransferases which catalyze the transfer of sugar moieties from activated donor molecules to specific acceptor molecules, forming glycosidic bonds. The acceptor molecule can be a lipid, a protein, a heterocyclic compound, or another carbohydrate residue. The structures of the formed glycoconjugates are extremely diverse, reflecting a wide range of biological functions. The members of this family share a common GTB topology, one of the two protein topologies observed for nucleotide-sugar-dependent glycosyltransferases. GTB proteins have distinct N- and C- terminal domains each containing a typical Rossmann fold. The two domains have high structural homology despite minimal sequence homology. The large cleft that separates the two domains includes the catalytic center and permits a high degree of flexibility. The conserved domains of BMGT1 (Fig. 3.24 A) and BMGT2 (Fig. 3.24 B) are shown below.



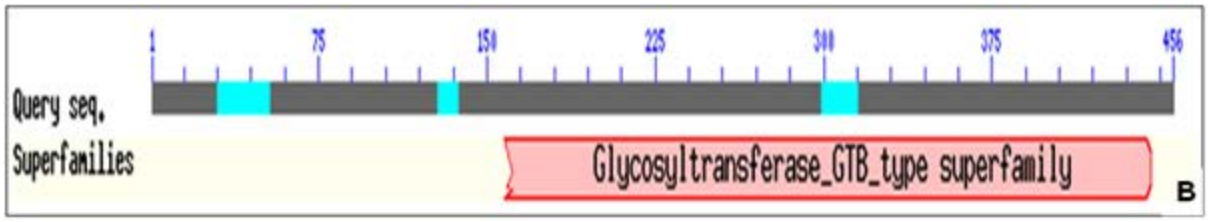


Fig. 3.24: Conserved domains of BMGT1 (A) and BMGT2 (B) genes.

3.3.2.5.9 Phylogenetic analysis of BMGT1 (FJ586244) and BMGT2 (FJ586245) genes with other reported Glycosyltransferases

Phylogenetic analysis was done using 66 reported protein sequences of GT, which were retrieved from the GenBank database. Both BMGT1 (FJ586244) and BMGT2 (FJ586245) genes were used for alignment. Multiple alignments of the predicted protein coding sequences were performed using Clustal X. The phylogenetic tree was constructed by neighbor joining methods using Mega 4.0 software program. It is clear from the phylogenetic tree that, BMGT1 is evolutionarily most similar with *Lycium barbarum*, *Vitis vinefera* and *Withania somnifera* but BMGT2 is most likely similar to *Antirrhinum majus*, *Sesamum indicum and Perilla frutescins* (Fig. 3.25). Another important point which could be predicted from the phylogenetic tree was that BMGT1 gene belonged to the UDP-Glucose:glycosyltransferase group whereas BMGT2 gene belonged to the UDP-Glucuronic acid:glycosyltransferase group which was also reflected in the amino acid sequence of the two genes respectively.

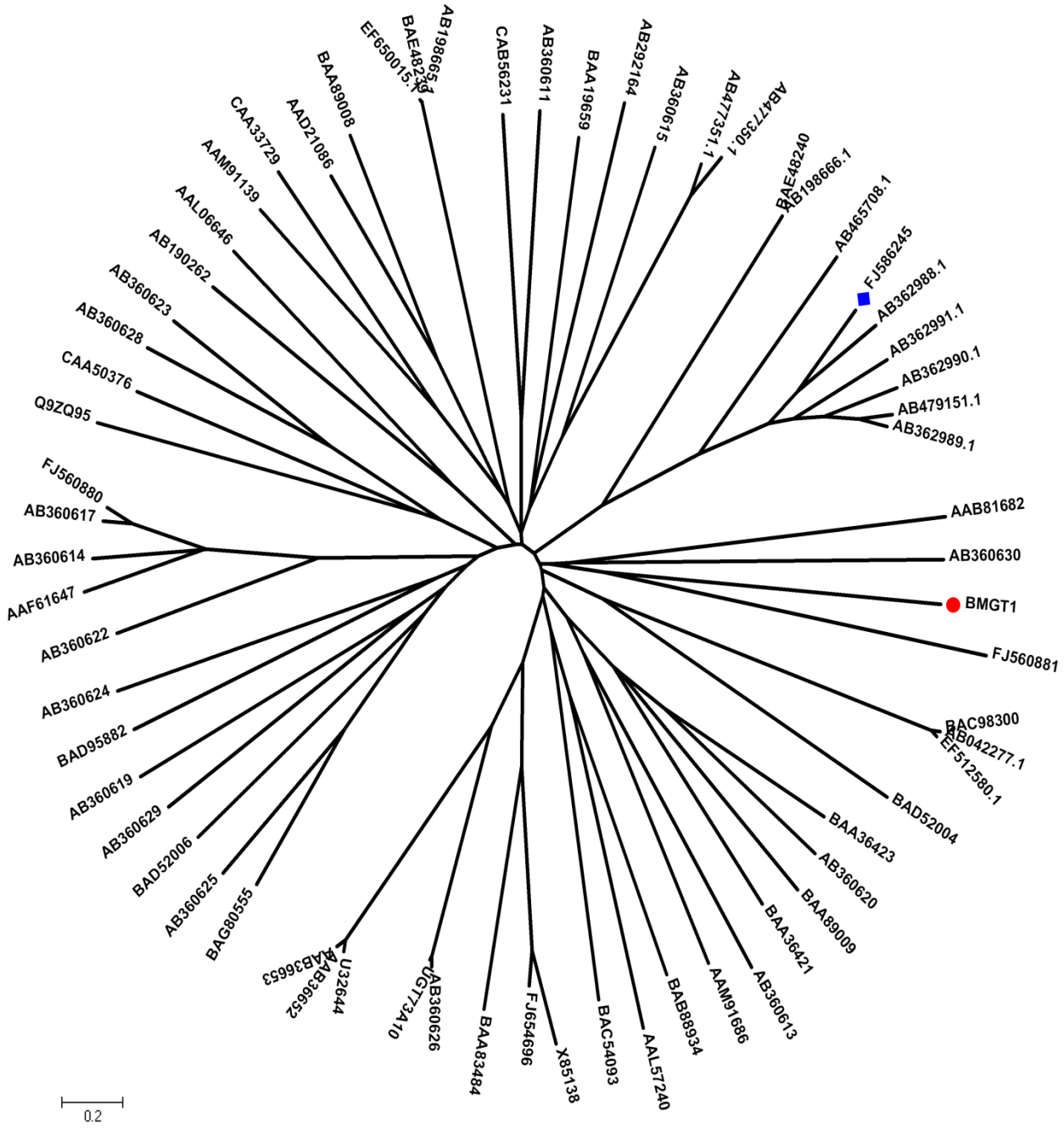


Fig. 3.25: Phylogenetic tree using 66 selected plant glycosyltransferase proteins, constructed by neighbor-joining methods using Mega 4.0 software program, after alignment of amino acids sequences by Clustal X. The tree shows the relatedness between BMGT1 (FJ586244) and BMGT2 (FJ586245), to the amino acid sequences of other plant gltcosyltransferases. The details of the accession numbers has been given in table 3.5.

Table 3.5: Details of the Accessions numbers:

S.No	Accession No.	Plant specie
1	AAB81682	*Vitis vinefera*
2	AB360630	*Lycium barbarum*
3	BMGT1 (FJ586244)	*Bacopa monniera*
4	FJ560881	*Withania Somnifera*
5	BAC98300	*Scutellaria baicalensis*
6	AB042277.1	*Scutellaria baicalensis*
7	EF512580.1	*Scutellaria baicalensis*
8	BAD52004	*Dianthus caryophyllus*
9	BAA36423	*Verbena x hybrid*
10	BAA36421	*Perilla frutescins*
11	AB360620	*Lycium barbarum*
12	BAA89009	*Petunia x hybrid*
13	AB360613	*Lycium barbarum*
14	AAM91686	*Arabidopsis thaliana*
15	BAB88934	*Nicotiana tabacum*
16	AAL57240	*Dorotheanthus bellidiformis*
17	BAC54093	*Torenia hybrid*
18	X85138	*Lycopersicum esculentum*
19	FJ654696	*Withania somnifera*
20	BAA43484	*Scutellaria baicalensis*
21	AB360626	*Lycium barbarum*
22	UGT73A10	*Lycium barbarum*
23	U32644	*Nicotiana tabacum*
24	AAB36652	*Nicotiana tabacum*
25	AAB36653	*Nicotiana tabacum*
26	BAG80555	*Lycium barbarum*
27	AB360625	*Lycium barbarum*
28	BAD52006	*Dorotheanthus bellidiformis*
29	AB360629	*Lycium barbarum*
30	AB360619	*Lycium barbarum*
31	BAD95882	*Ipomoea purpurea*
32	AB360624	*Lycium barbarum*
33	AB360622	*Lycium barbarum*
34	AAF61647	*Nicotiana tabacum*
35	AB360614	*Lycium barbarum*
36	AB360617	*Lycium barbarum*
37	FJ560880	*Withania somnifera*

38	Q9ZQ95	*Arabidopsis thaliana*
39	CAA50376	*Petunia x hybrid*
40	AB360628	*Lycium barbarum*
41	AB360623	*Lycium barbarum*
42	AB190262	*Bellis perennis*
43	AAL06646	*Citrus maxima*
44	AAM91139	*Arabidopsis thaliana*
45	CAA33729	*Hordeum vulgare*
46	AAD21086	*Forsythia x intermedia*
47	BAA89008	*Petunia x hybrida*
48	EF650015.1	*Antirrhinum majus*
49	BAE48239	*Antirrhinum majus*
50	AD198665.1	*Antirrhinum majus*
51	CAB56231	*Dorotheanthus bellidiformis*
52	AB360611	*Lycium barbarum*
53	BAA19659	*Perilla frutescins*
54	AB292164	*Glycine max*
55	AB360615	*Lycium barbarum*
56	AB477351.1	*Torenia hybrid cultivar*
57	AB477350.1	*Torenia hybrid cultivar*
58	BAE48240	*Linaria vulgaris*
59	AB198666.1	*Linaria vulgaris*
60	AB465708.1	*Veronica persica*
61	FJ586245 (BMGT2)	*Bacopa monniera*
62	AB362988.1	*Antirrhinum majus*
63	AB362989.1	*Scutellaria baicalensis*
64	AB362990.1	*Sesamum indicum*
65	AB362991.1	*Perilla frutescins*
66	AB479151.1	*Scutellaria baicalensis*

3.3.2.6 Isolation of full length genomic DNA GT genes:

For isolating gDNA clones of BMGT1 & BMGT2 genes, same forward primers from the initiation codon and reverse primers from the stop codon of the two genes which were used for obtaining full length cDNA clones of these genes were used and PCR was done using gDNA as template. In case of BMGT1 it resulted in amplification of different GT gene but in case of BMGT2 it resulted in an amplification of same ~1.4 kb fragment (Fig. 3.26 A) as was

obtained using cDNA as template. This fragment of BMGT2 was cloned in pGEMT- easy vector (Fig. 3.26 B) and confirmed by sequencing (Fig. 3.27)

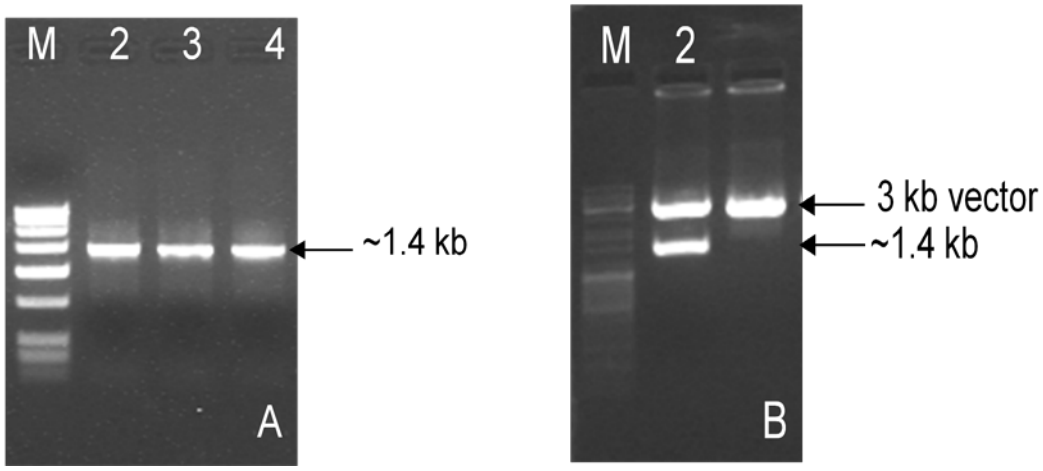


Fig. 3.26: A 1% Agarose gel showing full length BMGT2 gene PCR amplified from gDNA in lanes 2 to 4 (~1.4 kb) and lane M is marker. **B.** Restriction analysis of recombinant clone releasing ~1.4 kb insert (Lane 2) and lane M- marker.

Analysis of the sequence of genomic DNA clone of BMGT2 gene showed that it is almost similar to the cDNA sequence of the BMGT2 gene with few mismatches which may be due to sequencing error. The gDNA clone was found to be exactly of same size as cDNA clone without any introns present as is the case with several GTs isolated earlier from *Arabidopsis thaliana* (Ross *et al*., 2001).

3.3.2.6.1 Alignment of gDNA and cDNA clone of BMGT2 gene

```
gDNA            ATGGAAGACGCCATTGTTCTCTACTCCTCCGCCGAACACCTCAATTCCATGCTTGTCCTC 60
cDNA            ATGGAAGACGCCATTGTTCTCTACTCCTCCGCCGAACACCTCAATTCCATGCTTGTCCCC 60
                ************************************************************ *

gDNA            GCCAAATTCATCAGCAAACACCACCCCTCCATTTCCGTCATAATCATCAGCACCGCCGCC 120
cDNA            GCCAAATTCATCAGCAAACACCACCCCTCCATTTCCGTCATAATCATCAGCACCGCCGCC 120
                ************************************************************

gDNA            GAATCCGCCGCCGCCTCCGTCGCCTCTGTTCCTTCAATCACTTACCACCGCCTCCCCTCC 180
cDNA            GAATCCGCCGCCGCCTCCGTCGCCTCTGTTCCTTCAATCACTTACCACCGCCTCCCCTCC 180
                ************************************************************

gDNA            GCCCCTCTTCCTCCAGACTTAACCACCAATATCATCGAACTGTTCTTCGAAATTCCCCGC 240
cDNA            GCCCCTCTTCCTCCAGACTTAACCACCAGTATCATCGAACTGTTCTTCGAAATTCCCCGT 240
                *************************** ****************************
```

```
gDNA            TTCCACAACCCATTCCTTCACGAAGCCCTTCTCGAAATCTCTCAGAAATCAAATCTCAGA 300
cDNA            TTCCACAACCCATTCCTTCACGAAGCCCTTCTCGAAATCTCTCAGAAATCAAATCTCAGA 300
                ************************************************************


gDNA            GCATTTCTCATCGATTTCTTCTGCAATTCAACTTTTGAAGTCTCCACAAGCTTGAATATA 360
cDNA            GCATTTCTCATCGATTTCTTCTGCAATTCAACTTTTGAAGTCTCCACAAGCTTGAATATA 360
                ************************************************************

gDNA            CCCACCTACTTCTACCTCAGCGGCGGCGCGTGCGGCCTCTGCGCGCTTCTGTATTTCCCG 420
cDNA            CCCACCTACTTCTACCTCAGCGGCGGCGCGTGCGGCCTCTGCGCGCTTCTGTATTTCCCG 420
                ************************************************************

gDNA            ACCATCGACGAAGCCGTCAGTCCTCGAGATATCGGAGAATTGAACGATTTTCTTGAAATT 480
cDNA            ACCATCGACGAAGCCGTCAGTCCTCGAGATATCGGAGAATTGAACGATTTTCTTGAAATT 480
                ************************************************************

gDNA            CCGGGCTGTCCACCGATTCACTCGTTGGATTTTCCGAAAGCCATGTGGTTTCGCCGGAGC 540
cDNA            CCGGGCTGTCCACCGGTTCACTCGTTGGATTTTCCGAAAGCCATGTGGTTTCGCCGGAGC 540
                *************** ********************************************

gDNA            AATACGTACAAACATTTTCTTGACACCGCCGGAAATATGCGTAGAGCGAGTGGAATCGTC 600
cDNA            AATACGTACAAACATTTTCTTGACACCGCCGGAAATATGCGTAGAGCGAGTGGAATCGTC 600
                ************************************************************

gDNA            ACGAACTCGTTCGACGCGATCGAGTTCAGAGCTAAAGAAGCTTTGTCAAACAGCTTATGC 660
cDNA            ACGAACTCGTTCGACGCGATCGAGTTCAGAGCTAAAGAAGCTTTGTCAAACAGCTTATGC 660
                ************************************************************

gDNA            ACTCCCGGTCTCGCAACGCCGCCGGTTTACGTCATCGGACCTTTGGTTGCTGAAACGAAC 720
cDNA            ACTCCCGGTCTCGCAACGCCGCCGGTTTACGTCATCGGACCTTTGGTTGCTGAAACGAAC 720
                ************************************************************

gDNA            AGGAAAAATGGCGGCGAGGAGCATGAATGCCTGAAATGGCTCGATTCTCAGCCAATTAAA 780
cDNA            AGGAAAAATGGCGGCGAGGAGCATGAATGCCTGAAATGGCTCGATTCTCAGCCAATTAAA 780
                ************************************************************

gDNA            AGCGTGATTTTCCTCTGTTTTGGTCGACGTGGCCTATTTTCAGCCGCCCAATTGAAAGAA 840
cDNA            AGCGTGATTTTCCTCTGTTTTGGTCGACGTGGCCTATTTTCAGCCGCCCAATTGAAAGAA 840
                ************************************************************

gDNA            ATGGCGATCGGGCTTGAAAACAGTGGCCACCGGTTTCTCTGGTCTGTCCGCAGCCCGCCC 900
cDNA            ATGGCGATCGGGCTTGAAAACAGTGGCCACCGGTTTCTCTGGTCTGTCCGCAGCCCGCCC 900
                ************************************************************

gDNA            GGCCCGGCGGCGGCGAAAGATCCCGACTTGGACGCTTTGCTTCCGGAAGGTTTTATGGAG 960
cDNA            GGCCCGGCGGCGGCGAAAGATCCCGACTTGGACGCTTTGCTTCCGGAAGGTTTTATGGAG 960
                ************************************************************

gDNA            AGGACGAAAGATAGGGGATTTGTTATAAAGACGTGGGCCCCGCAGAAGGAGGTGCTCAGC 1020
cDNA            AGGACGAAAGATAGGGGATTTGTTATAAAGACGTGGGCCCCGCAGAAGGAGGTGCTCAGC 1020
                ************************************************************

gDNA            CACGAGGCCGTCGGCGGGTTCGTGACTCACTGCGGCCGGAGCTCTGTTTTGGAAGCGGTG 1080
cDNA            CACGAGGCCGTCGGCGGGTTCGTGACTCACTGCGGCCGGAGCTCTGTTTTGGAAGCGGTG 1080
                ************************************************************

gDNA            TCGTTTGGGGTTCCGATGATTGGGTGGCCGATGTACGCGGAACAGAGGATGCAGCGTGTT 1140
cDNA            TCGTTTGGGGTTCCGATGATTGGGTGGCCGATGTACGCGGAGCAGAGGATGCAGCGTGTT 1140
                ****************************************** ******************

gDNA            TTCATGGTGGAGGAAATGAAGGTGGCGCTGCCGCTTTCGGAGGAGGCGGACGGGTTCGTG 1200
cDNA            TTCATGGTGGAGGAAATGAAGGTGGCGCTGCCGCTTGCGGAGGAGGCGGACGGGTTCGTG 1200
                ************************************ ***********************

gDNA            ACGGCGGGGGAGCTGGAGAAGCGAGTGAGAGAGTTGATGGGATCGCCGGCGGGCAAGGCC 1260
cDNA            ACGGCGGGGGAGCTGGAGAAGCGAGTGAGAGAGTTGATGGGGTTGCCGGCGGGAAAGGCC 1260
                ****************************************** * ********* ******
```

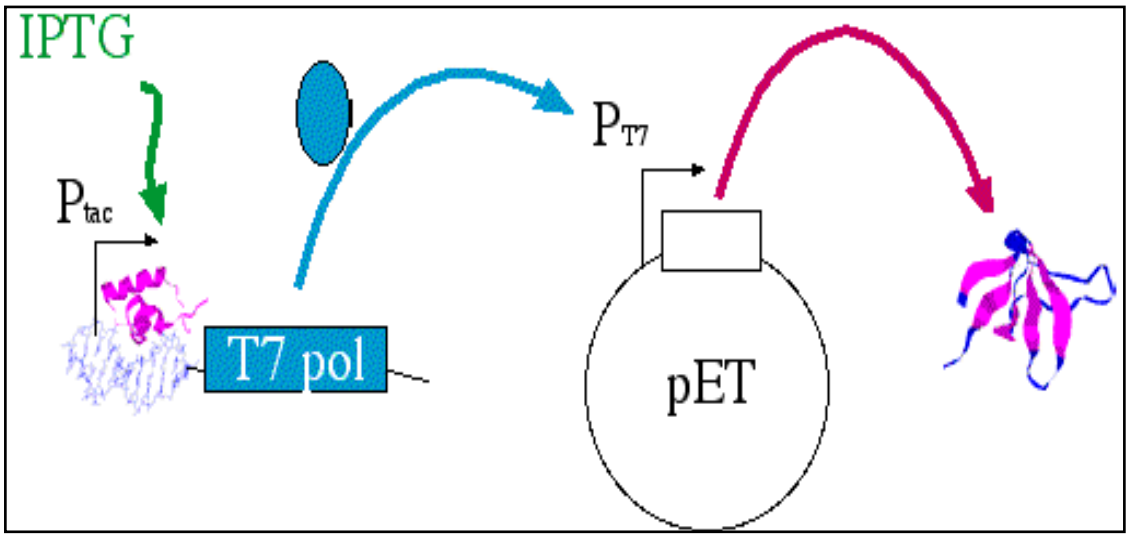
```
gDNA            GTGAGGCAGCGCGTGGCGGAATTGAGGACGGCGGCCGAGGCTGCGGTGCGGAAGGGTGGG 1320
cDNA            GTGACGCAGCGCGTGGCGGAATTGAGGACGGCGGCCGAGGCTGCGGTGCGGAAGGGTGGA 1320
                **** ************************************************** 

gDNA            TCGTCTGTTGTTGCTTTAGGAAAGTTCATTGAGACGGTGACCCGACGCTAA 1371
cDNA            TCGTCTGTTGTTGCTTTAGGAAAGTTCATTGAGACGGTGACCCGACGCTAA 1371
                ************************************************* 
```

3.4 Conclusions

- PCR based approach was used to fish out the GT gene. Two GT cDNA clones were isolated namely BMGT1 (Accession No. **FJ586244)** and BMGT2 (Accession No. **FJ586245).**
- BMGT1 clone showed 49% identity at amino acid level with the UDP-glucose:glycosyltransferase gene from *Lycium barbarum* and BMGT2 clone showed 72% identity at amino acid level with the UDP-glucose:glycosyltransferase gene from *Sesamum indicum.*
- RACE reaction yielded 5'UTR and 3'UTRs. The putative poly A tails were identified in the 3' UTRs of both the genes. The full length cDNA clone of both the glycosyltransferases were of 1392 bp and 1371 bp respectively.
- Analysis of deduced amino acid sequence of BMGT1 and BMGT2 showed the conserved domains for glycosyltransferase acceptor and donor binding sites. It also showed the most conserved 45 bp domain called as PSPG box which is found to be present in all the plant glycosyltransferase genes.
- Phylogenetic analysis of BMGT1 deduced amino acid sequence was done using MEGA 4 software program. The results show that BMGT1 is evolutionarily most similar to UDP-glucose:glycosyltransferases from *Lycium barbarum* of *Lamiaceae* family whereas BMGT2 is evolutionary similar to UDP-glucuronic acid:glycosyltransferases from *Sesamum indicum* of *Lamiaceae* family.
- The genomic clone of BMGT2 is of 1371 bp same as cDNA clone without any intron present.

Chapter: 4(A)

Heterologous Expression of *B. monniera* GT gene(s), its Purification & Characterization



4.1 Introduction

Glycosyltransferases (GTs) that use secondary metabolites as substrates are minor constituents in plant cells (Vogt & Jones, 2000). Although many of these enzymes have been isolated from several plant species and assayed *in vitro*, in many cases their roles in the secondary metabolism of these plants are still unknown. Is there one GT that glycosylates all aglycones or one GT for each and every substrate? The answer is most likely somewhere in between; although we cannot exclude that there is a great variation in the range of substrates that different GTs are able to accept. It has been suggested that GTs are highly regiospecific (or selective) rather than substrate specific (Vogt & Jones, 1997). In other words, acceptance is strictly based on a certain substructure, with little regards for the entire structure of the acceptor molecule. In this regard it is interesting to over express these recombinant GTs in suitable expression systems and carry out the structure function relationships studies.

Analysis of gene function is of central importance for the understanding of physiological processes. Expression of genes in heterologous organisms has allowed the isolation of many important genes (e.g. for nutrient uptake and transport) and has contributed a lot to the functional analysis of the gene products. Heterologous expression systems are powerful tools for isolating new genes and for characterizing proteins from all organisms. An efficient way to clone and simultaneously prove the function of a gene is possible with its functional expression in heterologous host cells. Heterologous expression of plant genes provides a new technique for determining gene-product function.

In chapter 3, two cDNAs encoding *GT* gene were characterized. *In silico* studies of these two sequences revealed that both the genes have different molecular weight, isoelectric point and hydropathicity. Therefore, both of the cDNA encoding *GT* were chosen for over-expression and further studies.

4.2 Materials and methods

4.2.1 Materials

Glassware and plasticware: As discussed in Chapter 2, Section 2.2 and Section 2.3

Chemicals: As discussed in Chapter 2, Section 2.4

***Escherichia coli* and plasmids used in the study:**

Escherichia coli XL-1 Blue (Stratagene, USA)

E. coli, BL 21 (DE3) (Invitrogen, USA)

pGEM-T Easy Cloning vector (Promega, USA)

pET30b (+) Expression vector (Novagen, USA).

Stock solutions: As discussed in Chapter 2, Section 2.6.3

4.2.2 Methods

4.2.2.1 Bacterial culture conditions: As discussed in Chapter 2, Section 2.8.1

4.2.2.2 Bacterial cells transformation: *E. coli* transformation and selection was done as discussed in Chapter 2, Section 2.8.2.3

4.2.2.3 Isolation of plasmid DNA from *E. coli*: As described in Chapter 2, Section 2.8.4.1

4.2.2.4 Restriction digestion of DNA: As described in Chapter 2, Section 2.8.4.3

4.2.2.5 Extraction and purification of DNA from agarose gel: As described in Chapter 2, Section 2.8.4.4

4.2.2.6 Polymerase Chain Reaction (PCR): As described in Chapter 2, Section 2.8.4.9

4.2.2.7 Colony PCR method: As described in Chapter 2, Section 2.8.3

4.2.2.8 Cloning of BMGT1 and BMGT2 genes in pET-30b (+)

GT genes were cloned in pGEM-T Easy vector by incorporating appropriate restriction sites in the primers. In case of BMGT1 gene, *Nde* I restriction site was added to the forward primer and *Xho* I restriction site was added to the reverse primer (namely BMGT1F and BMGT1R) whereas *Nde* I restriction site was added to the forward primer and *Bam* HI restriction site was added to the reverse primer of the BMGT2 gene (namely BMGT2F and BMGT2R). High fidelity *Taq pol* (*Pfx* Invitrogen) was used to amplify both the *GT* genes using the above set of primers from the cDNA clone, sequenced and maintained in pGEM-T Easy vector as the template. The above recombinant plasmid DNA was diluted 1:100 times and 1 µL was used as the template and PCR was performed (Chapter 2, Section 2.8.4.9). Two 1.4 kb bands were amplified (exactly 1392 bp coding region of BMGT1 gene + *Nde* I and *Xho* I sites added into primer sequences and 1371 bp coding region of BMGT2 gene + *Nde* I and *Bam* HI sites added into primer sequences). The bands were excised from gel, purified, ligated in pGEM-T Easy vector and transformed in *E.coli* XL1 MRF cells. Clones were screened by colony PCR and a few colonies were inoculated in 5 mL LB (Ampicilin 100 µg/mL) medium in tubes. Isolated individual plasmids of both genes were restriction digested with respective restriction enzymes to confirm the integration of the genes.

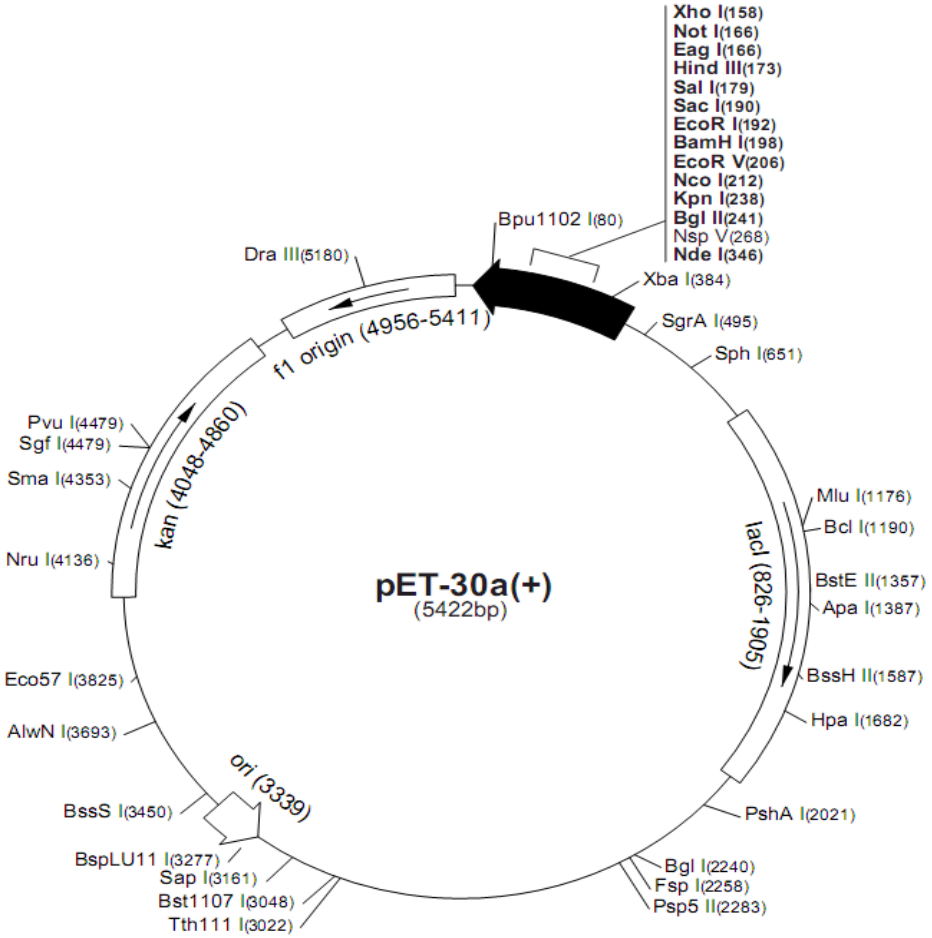
PCR cycling conditions for BMGT1 gene

Temperature	Time	Cycles
95 ºC	5 min	1 cycle
95 ºC	1 min	35 cycles
55 ºC	0.45 min	
72 ºC / 1 kb	1.30 min	
72 ºC / 1 kb	5 min	1 cycle

PCR cycling conditions for BMGT2 gene

Temperature	Time	Cycles
95 ºC	5 min	1 cycle
95 ºC	1 min	35 cycles
58 ºC	0.45 min	
72 ºC / 1 kb	1.30 min	
72 ºC / 1 kb	5 min	1 cycle

The recombinant *GT* clones with respective restriction sites were directionally cloned in pET-30b (+) vector (Fig 4A.1). Colony PCR (Chapter 2 Section 2.8.3) was done to screen the recombinant pET-30b (+) clones. Integration of *GT* gene in pET-30b (+) was confirmed by digestion with respective restriction enzymes (Fig. 4A.2).



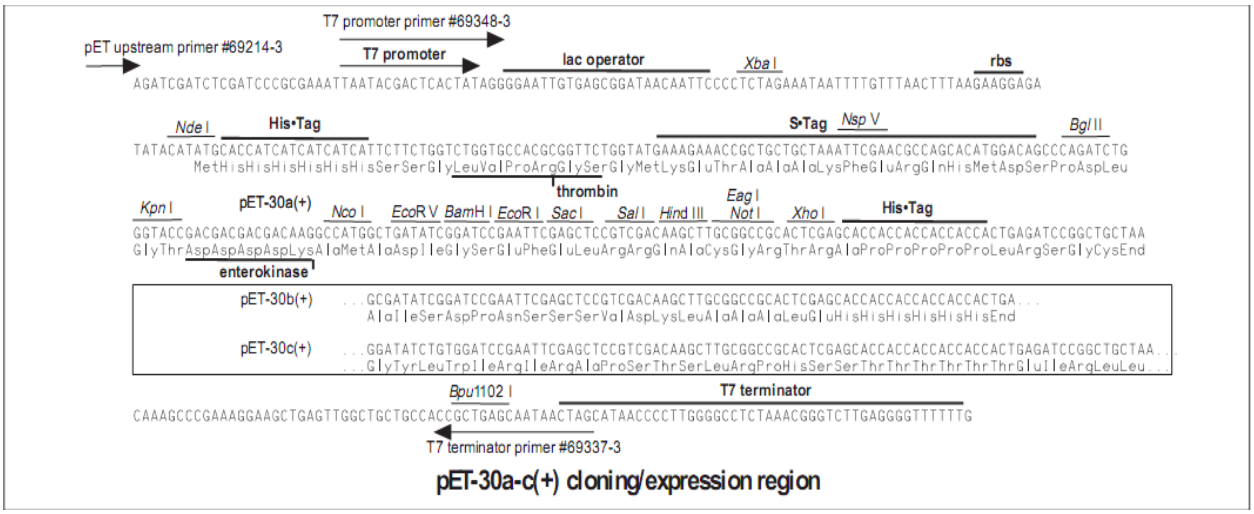


Fig. 4A.1: Vector map of pET-30 and pET 30b (+) is shown in box.

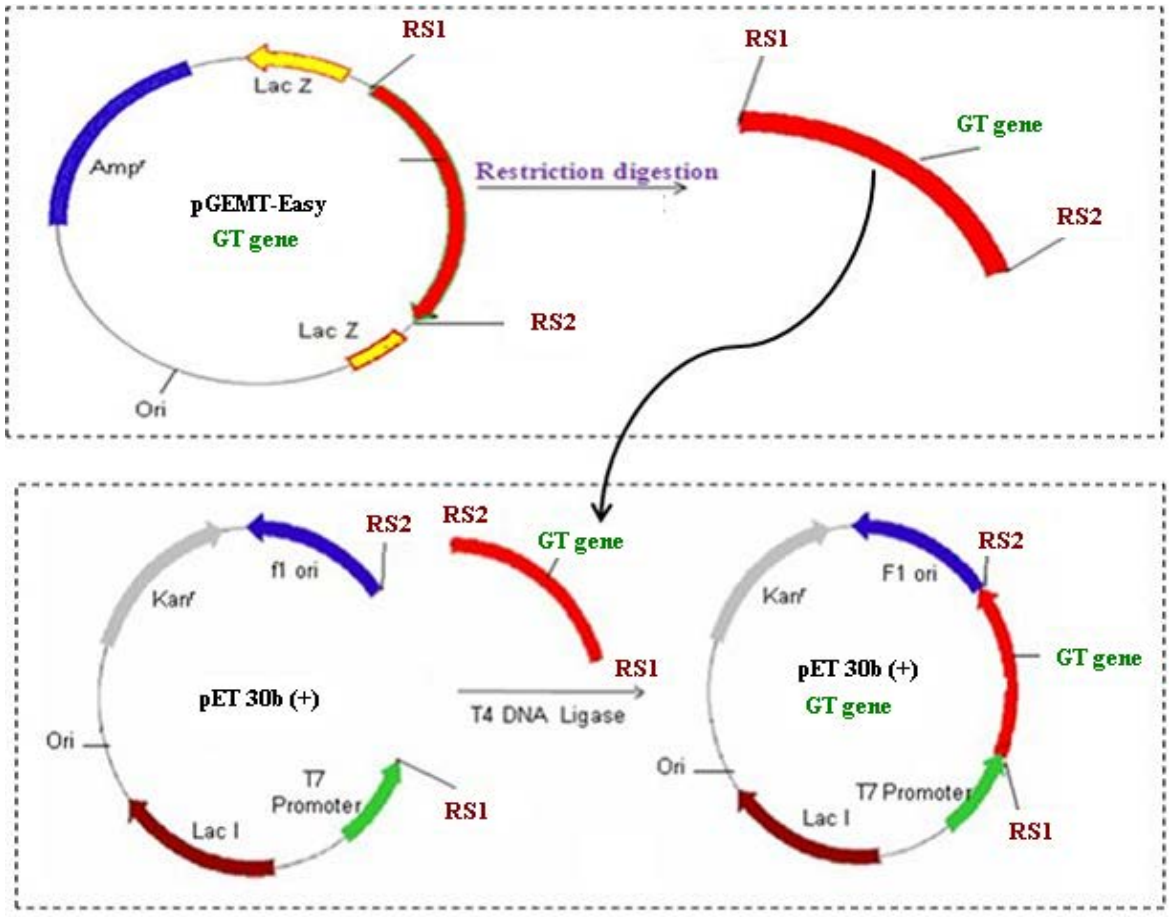


Fig. 4A.2: Strategy used for directional cloning of *GT* genes in pET-30b (+) vector. RS1 and RS2 indicate the respective restriction sites for BMGT1 and BMGT2 genes.

4.2.2.9 Recombinant GT protein expression and its purification from inclusion bodies

4.2.2.9.1 Recombinant protein expression in *E. coli* (BL21)

A single bacterial (BL21) colony carrying recombinant pET-30b (+) with *GT* gene, from freshly streaked plates (grown on LB agar medium containing 50 µg/mL kanamycin) was used to inoculate 5 mL liquid LB medium containing the same concentrations of the antibiotic. Culture was grown overnight with shaking at 225 rpm at 37 °C. One mL aliquot of over-night grown culture was used to inoculate 100 mL LB broth containing 50 µg/mL kanamycin. Once the cultures reached OD_{600} 0.4 to 0.5, recombinant protein expression was induced by the addition of isopropyl β-D thiogalactopyranoside (IPTG), and the culture was grown for 4 to 6 h at 37 °C with shaking at 150 rpm. Recombinant protein extraction was done according to the protocol described in Section 2.9. Likewise, four positive recombinant pET-30b (+) clones were screened for GT protein over-expression on 10% SDS PAGE (Chapter 2, Section 2.9.3). The clone showing maximum over-expression was chosen for further studies.

4.2.2.9.2 Purification of recombinant protein (GT)

His-tagged recombinant protein (GT) was 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-NTA (nickel-nitrilotriacetic 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. Purification of recombinant GT protein carrying a 6xHis-tag was conducted using Ni NTA Agarose beads (Qiagen) (Chapter 2, Section 2.9.2). Purity of protein was checked on 10% SDS-PAGE (Chapter 2, Section 2.9.3).

4.2.2.9.3 Raising polyclonal antibody against purified GT protein in rabbit

The purified GT protein was used for raising polyclonal antibodies in New Zealand White rabbit.

4.2.2.9.4 Pre-treatment of serum: As described in Chapter 2, Section 2.10.1

4.2.2.9.5 Determination of titre of antibodies

ELISA was performed to determine the titre of first, second and third bleed of rabbit serum (Chapter 2, Section 2.10.2). Once the antibody titre was determined then, a fixed dilution of antibody was used for rest of the experiments.

4.2.2.9.6 Standardization of time and temperature for protein expression in soluble form

The information obtained from above section 4.2.2.9 was utilized for this experiment. Several flasks with 50 mL LB broth (kanamycin 50 µg/mL) were inoculated with O/N grown culture of *E. coli* BL21 harboring recombinant pET-30b (+) plasmid carrying *GT* genes. The cells were grown till A_{600} reached 0.5-0.6 and induction was done with 0.01-1.0 mM IPTG. The cultures were grown at different temperatures (12-37 °C) and different durations (4 to 18 h) for optimization of maximum expression of recombinant GT proteins in soluble form. One flask as an uninduced control sample was used before inducing with IPTG. The soluble fraction of cell lysate was analysed on 10% SDS-PAGE.

4.2.2.9.7 Protein estimation

Protein estimation was done using Bradford assay (1976). A standard graph was made for BSA and concentration of unknown sample was determined by plotting standard graph (See appendix).

4.2.2.9.8 GT enzyme assay

The crude lysate of both the glycosyltransferase proteins were used for performing enzyme assay studies. The control reaction contained crude lysate of only pET 30b (+) vector transformed in *E. coli* (BL21). Assay reaction was carried out as described in Chapter 2, Section 2.12. See appendix, Table 2.9 substrates used for the study.

4.3 Results and discussions

4.3.1 Cloning of *B. monniera* BMGT1 and BMGT2 genes in pET-30b (+) vector

BMGT1 and BMGT2 genes were cloned in pET-30b (+) expression system to get the gene expressed in its active form and for its characterization.

4.3.1.1 Incorporation of restriction sites

The BMGT1 gene cloned in pGEM-T Easy vector was amplified using gene specific primers **BMGT1 Forward** and **BMGT1 Reverse** to incorporate the *Nde* I site at the 5' end and *Xho* I site at 3' end of BMGT1 gene.

BMGT1 Forward primer- 5' CAT ATG GAG AGC AAA GGA ACA GGG AAG GA 3'
BMGT1 Reverse primer- 5' CTC GAG AGT TAG GAG CAG GGA CAT GTT AAA 3'

Approximately, 1.4 kb (1392 bp BMGT1 + *Nde* I and *Xho* I restriction sites) band was amplified. The band was cut, purified, ligated in pGEM-T Easy vector and transformed in *E. coli* XL1 Blue cells. Clones with BMGT1 gene with *Nde* I and *Xho* I restriction sites were screened by inoculating a few colonies in 5 mL LB media (ampicilin 100 μg/mL) tubes. Isolated individual plasmids were restriction digested with *Nde* I and *Xho* I enzymes to confirm the integration of BMGT1 gene.

Similarly BMGT2 gene cloned in pGEM-T Easy vector was amplified with gene specific primers **BMGT2 Forward** and **BMGT2 Reverse** to incorporate *Nde* I site at the 5'end and *Bam* HI site at the 3'end of the BMGT2 gene.

BMGT2 Forward primer- 5' CAT ATG GAA GAC GCC ATT GTT CTC TAC TC 3'
BMGT2 Reverse primer- 5' GGA TCC ATG CGT CGG GTC ACC GTC TCA AT 3'

Approximately, 1.4 kb (1371 bp BMGT2 + *Nde* I and *Bam* HI restriction sites) band was amplified which was excised, gel purified, ligated in pGEM-T Easy vector and transformed in *E. coli* XL1 Blue cells. Clones with BMGT2 gene over hanged with *Nde* I and *Bam* HI restriction sites were screened by inoculating a few colonies in 5 mL LB (ampicilin 100 μg/mL) medium in tubes. Isolated individual plasmids were digested with restriction enzymes *Nde* I and *Bam* HI to confirm the integration of BMGT2 gene.

4.3.1.2 Directional cloning of *B. monniera* BMGT1 and BMGT2 genes in pET-30b (+)

The above clone of BMGT1 gene in pGEM-T Easy vector was restriction digested with *Nde* I and *Xho* I restriction enzymes and gene of interest was purified. pET-30b (+) vector DNA was also digested with same restriction enzymes and purified. BMGT1 gene was directionally cloned in purified restriction digested pET-30b (+) vector. Ligation mixture was transformed into *E. coli* XL1 competent cells and plated on LB-agar plate (kanamycin 50 μg/mL). *E. coli* colonies having recombinant plasmids were screened by colony PCR (Fig. 4A.3), with reaction cycles as shown in table below. Plasmids were isolated from PCR positive clones and were digested with *Nde* I and *Xho* I to confirm the integration of BMGT1 gene fragment in pET-30b (+) vector (Fig. 4A.4).

PCR cycling conditions for BMGT1 gene

Temperature	Time	Cycles
95 ºC	5 min	1 cycle
95 ºC	1 min	

55 ºC	0.45 min	35 cycles
72 ºC / 1 kb	1.30 min	
72 ºC / 1 kb	5 min	1 cycle

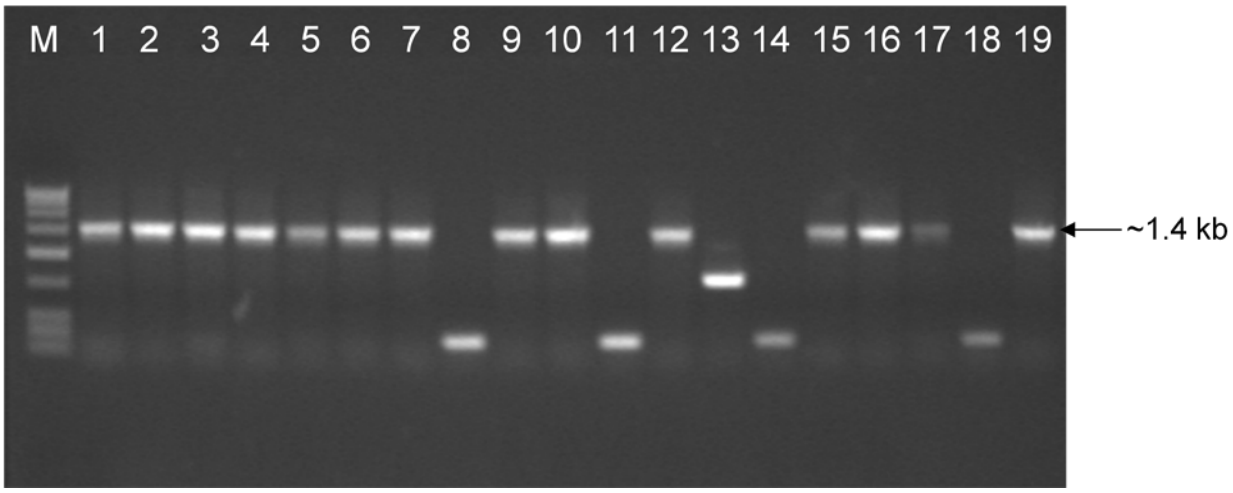


Fig. 4A.3: Colony PCR showing ~1.4 kb BMGT1 gene containing recombinant colonies; Lane M- marker, lane 1-7, 9, 10, 12, 15, 16, 17 & 19 are positive clones.

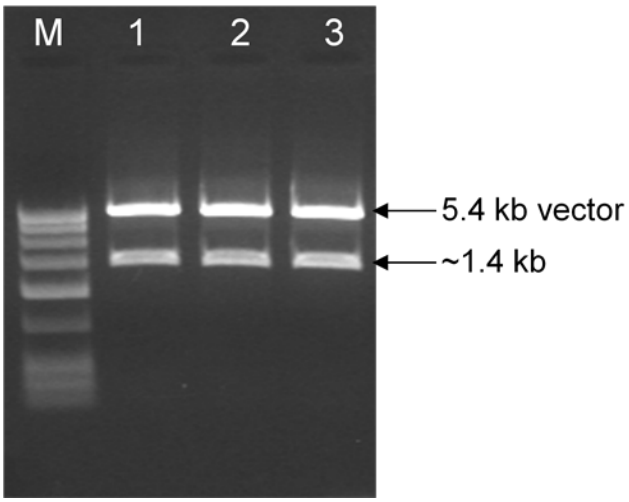


Fig. 4A.4 Restriction analysis of recombinant pET-30b (+): lane 1-3 recombinant vector releasing ~1.4 kb BMGT1 gene and 5.4 kb vector backbone, lane M- marker.

Integration of the BMGT1 gene along with restriction sites was further confirmed by sequencing.

The sequence was translated using proteomic tools available at www.expasy.ch and was checked for in frame translation up to HIS tag.

Similarly BMGT2 gene in pGEM-T Easy vector was restriction digested with *Nde* I and *Bam* HI restriction enzymes and digested product was purified. pET-30b (+) vector DNA was also digested with same restriction enzymes and purified. BMGT2 gene was directionally cloned in purified restriction digested pET-30b (+) vector. Ligation mixture was transformed into *E. coli* XL1 competent cells and plated on LB-agar plate (kanamycin 50 μg/mL). *E.coli* colonies harbouring recombinant plasmids were screened by colony PCR (Fig 4A.5), with reaction cycles as shown in table below. Plasmids were isolated from PCR positive clones and were confirmed by restriction digestion (Fig. 4A.6) and sequence validation.

PCR cycling conditions for BMGT2 gene

Temperature	Time	Cycles
95 ºC	5 min	1 cycle
95 ºC	1 min	35 cycles
58 ºC	0.45 min	
72 ºC / 1 kb	1.30 min	
72 ºC / 1 kb	5 min	1 cycle

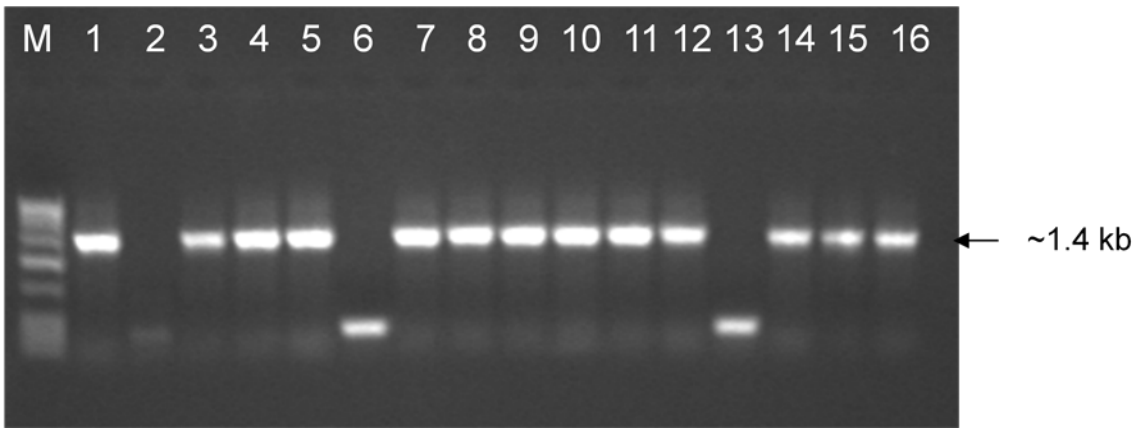


Fig. 4A.5 Colony PCR showing ~1.4 kb BMGT2 gene containing recombinant colones; Lane M- marker, lane 1, 3-5, 7-12 and 14-16 are positive clones.

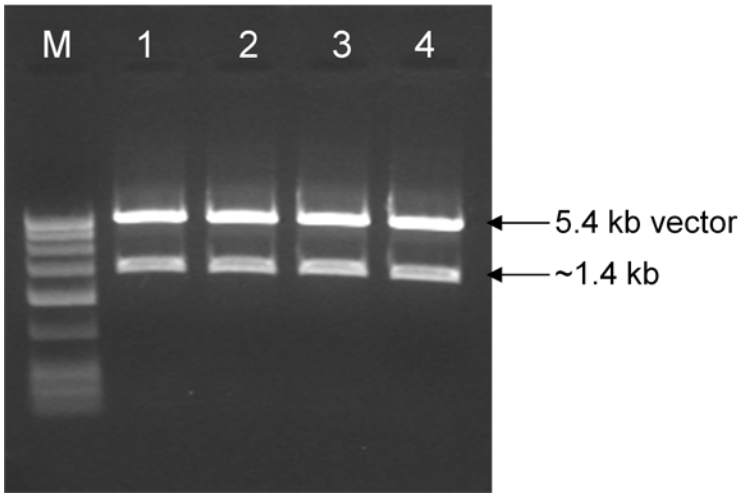


Fig. 4A.6 Restriction analysis of recombinant pET-30b(+): lane 1-4 recombinant vector releasing ~1.4 kb BMGT2 gene and 5.4 kb vector backbone, lane M- marker.

4.3.2 Recombinant BMGT1 and BMGT2 protein expression and their purification from inclusion bodies

4.3.2.1 Recombinant BMGT1 protein expression and purification

E. coli BL 21 (DE3) cells transformed with recombinant pET-30b (+) plasmids having BMGT1 gene were screened for over-expression of protein. A few positive recombinant clones were screened for recombinant BMGT1 protein over-expression. An approximately 52 kDa protein was expressed in all clones analysed on 10% SDS-PAGE with varying levels of expression. Among those recombinant clones one of the clone showing maximum expression was chosen and large scale protein purification was done using Ni-NTA Agarose beads (Fig. 4A.7).

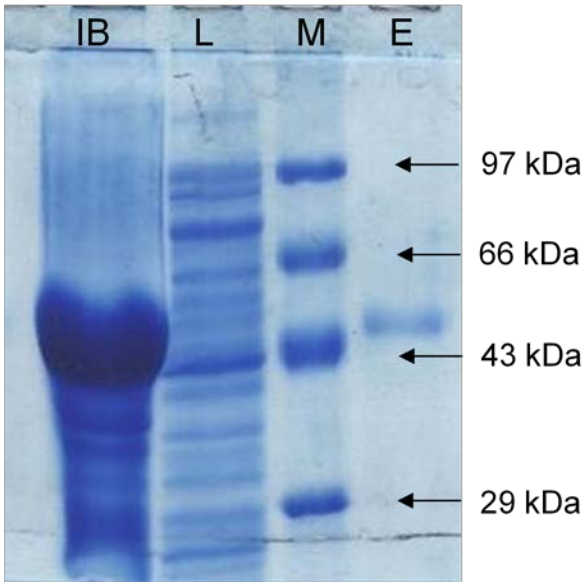


Fig. 4A.7: 10% SDS PAGE; Commassie Blue staining: Lane IB- Inclusion bodies from BMGT1 protein, Lane L- Lysate of BMGT1 Protein, Lane M- Protein molecular weight marker. Lane E- purified BMGT1 using Ni-NTA Agarose beads.

Purified BMGT1 protein was dialyzed against 1X PBS buffer supplemented with 1 mM DTT and 0.1 mM PMSF overnight with two changes of fresh buffer. Dialysed BMGT1 protein was

concentrated, quantified by Bradford assay (Bradford reagent, Promega, USA) and given for raising antibodies in New Zealand rabbit.

4.3.2.2 Recombinant BMGT2 protein expression and purification

Plasmids harboring BMGT2 gene in pET-30b (+) vector which was transformed in *E. coli* BL 21 (DE3) cells were screened for over-expression. A few positive recombinant clones were screened for recombinant BMGT2 protein over-expression. An approximately 50 kDa protein was found to express in all clones analysed on 10% SDS-PAGE with varying levels of expression. Among those recombinant clones one of the clone showing maximum expression was chosen and protein was purified using Ni-NTA Agarose beads (Fig. 4A.8).

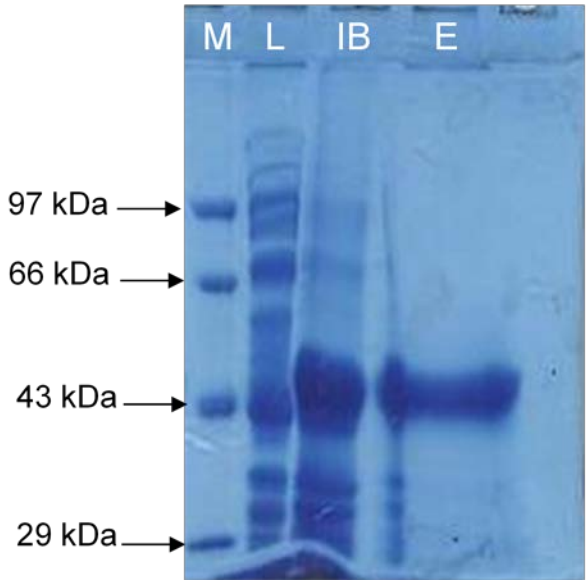


Fig. 4A.8: 10% SDS PAGE; Commassie Blue staining : Lane M- Protein molecular weight marker, Lane L- Lysate of BMGT2 Protein, Lane IB- Inclusion bodies from BMGT2 protein and Lane E- Purified BMGT2 using Ni-NTA Agarose beads.

4.3.3 Raising antibodies in rabbit against BMGT1 protein

300 µg of purified protein was used for first injection in New Zealand rabbit to raise antibodies. Same amount of protein was used for booster doses. Antibody titer of first, second and third bleed was determined by ELISA. Titer of third bleed is ~1:150000 (Fig. 4A.9). 3rd bleed serum dilution of 1:10000 was used for further experiments.

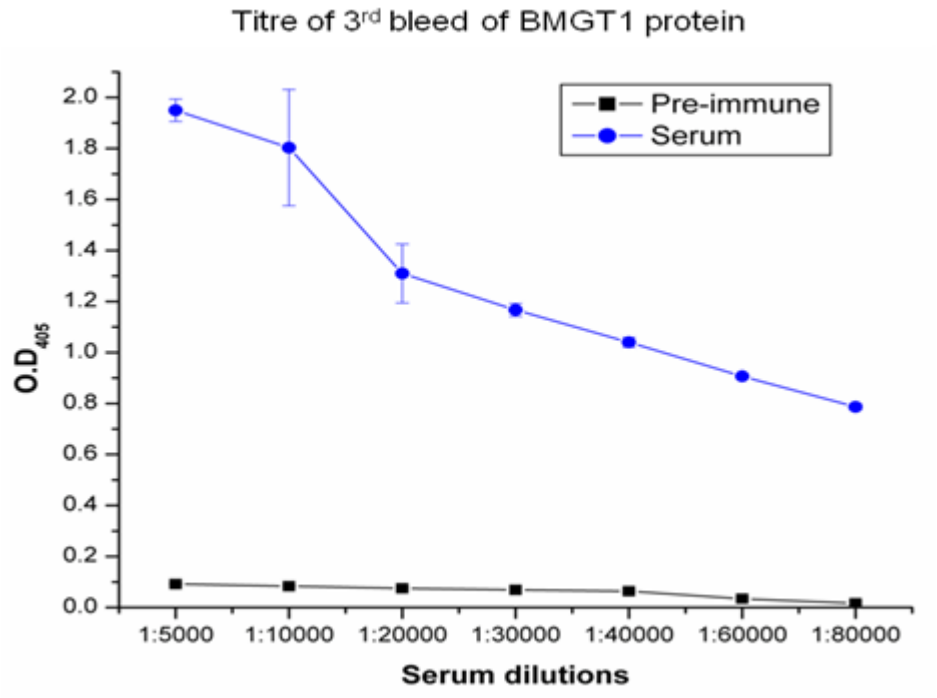
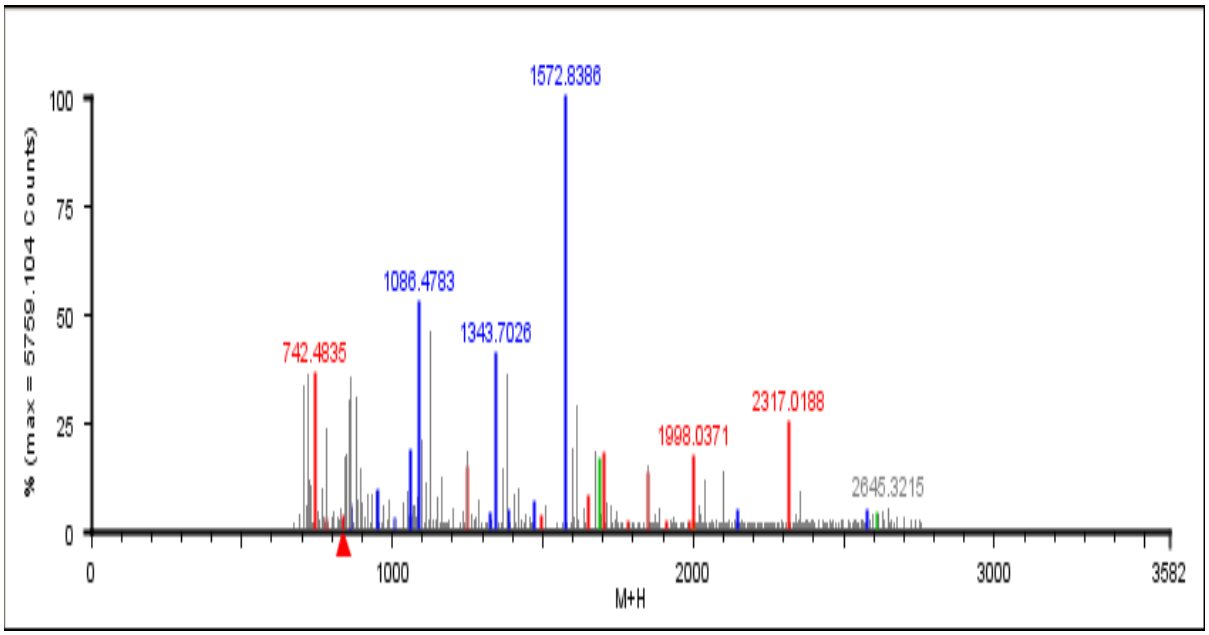


Fig. 4A.9: Determination of titre of 3rd bleed serum; Black square (■) represents OD_{405} of Pre-immune serum and blue circles (●) represents OD_{405} of serum dilution.

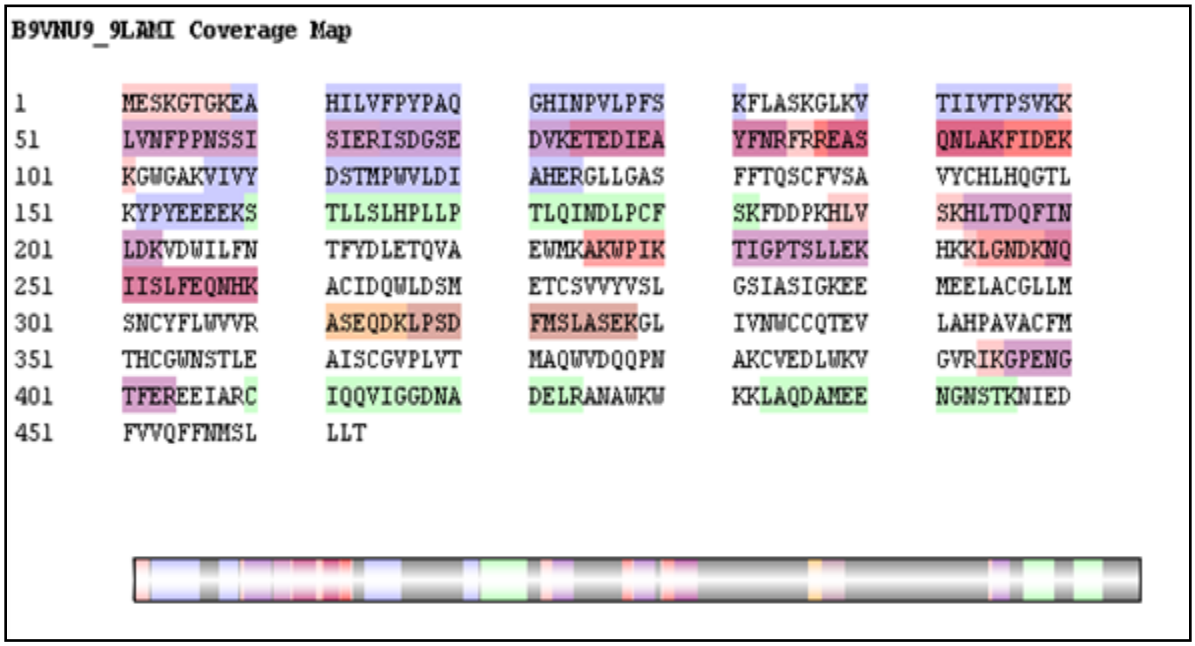
4.3.4 MALDI MS/MS analysis of BMGT1 and BMGT2 proteins

MALDI MS/MS was done as described in Chapter 2, Section 2.11. MS/MS analysis was done for confirmation of recombinant BMGT1 and BMGT2 proteins with the sequences already submitted to the Uniprot Database. The MALDI MS/MS spectra and coverage map of BMGT1 and BMGT2 proteins are shown in Fig 4A.10 & Fig 4A.11 respectively. Ionization spectra shows mass of different peptides of the query protein and coverage map shows

number of peptides of query protein which shows exact match with the template (BMGT1 and BMGT2) present in the Uniprot database.

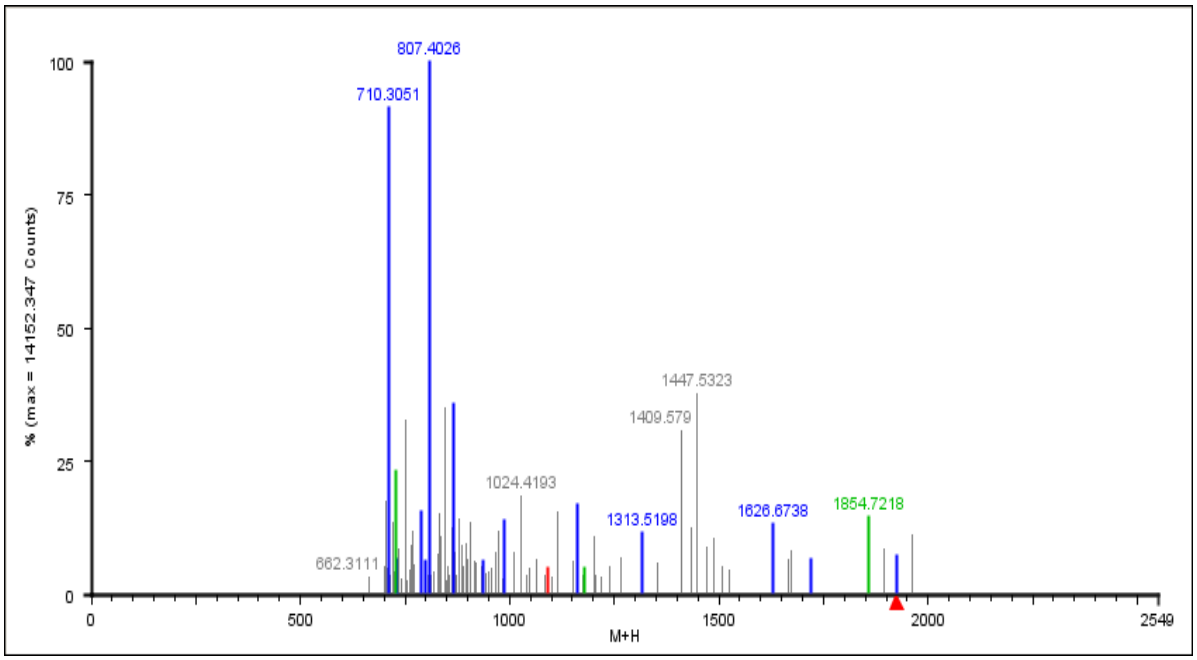


A

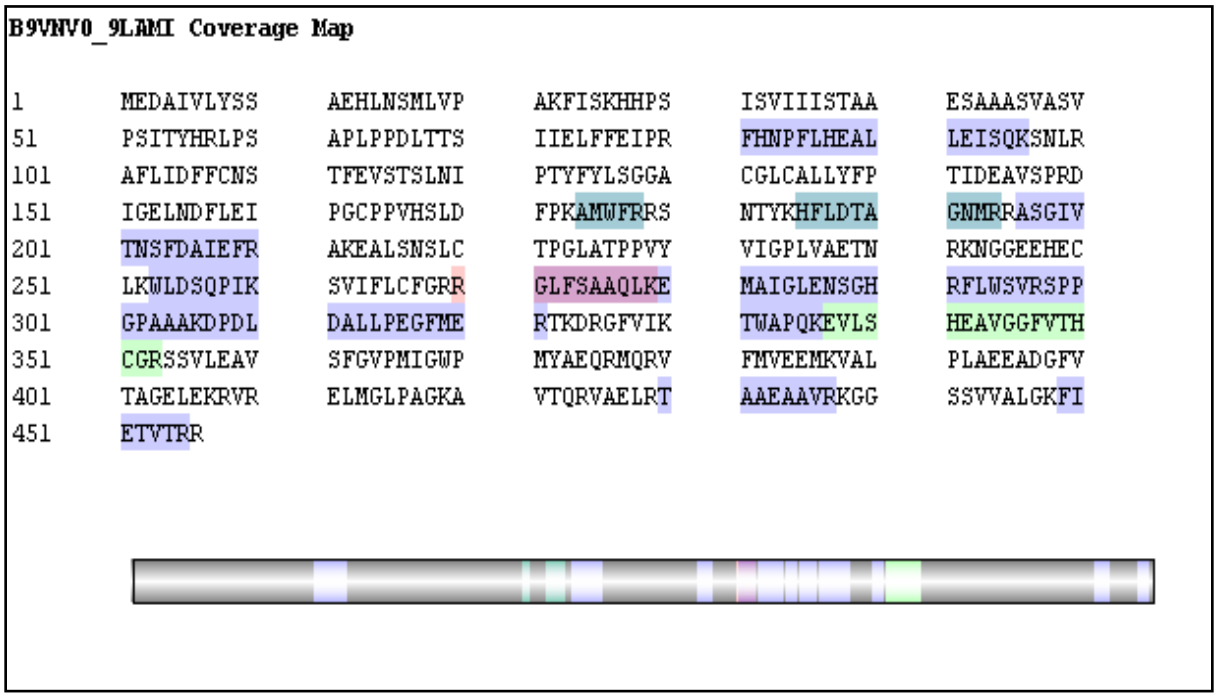


B

Fig. 4A.10: MALDI MS/MS analysis of recombinant BMGT1 protein **(A)** ionization spectra **(B)** coverage map of recombinant BMGT1 protein.



A



B

Fig. 4A.11: MALDI MS/MS analysis of recombinant BMGT2 protein **(A)** ionization spectra **(B)** coverage map of recombinant BMGT2 protein.

4.3.5 Extraction of recombinant BMGT1 and BMGT2 proteins in soluble form

BMGT1 clone which showed maximum expression as described above in section 4.3.2.1 was used for extraction of recombinant protein in soluble form to test the enzyme activity. Temperature and time required after induction was standardized for maximum BMGT1 protein expression in soluble form that is in cell lysate. The culture was grown at different temperatures ranging from 12-37 °C and at different time durations (4 to 18 h) for optimization of maximum expression of recombinant BMGT1 protein in soluble form. IPTG concentration was also checked ranging from 0.03-1 mM. The optimum parameters were as follows; after initial growth at 37 ^{o}C till A_{600} reached 0.6 cells were induced with 0.1 mM IPTG and grown 12-18 hours at 15 °C as shown in Fig. 4A.12.

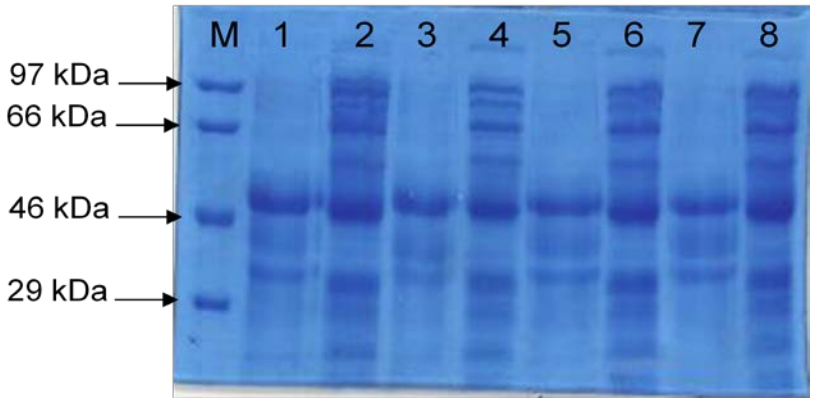


Fig. 4A.12: 10% SDS PAGE analysis of BMGT1 protein; Commassie Blue staining: Lane M- Protein molecular weight marker, Lane 1,3,5 and 7- Inclusion bodies from BMGT1 grown at 18, 16, 14 and 12 hours respectively, Lane 2,4,6,8- Lysate of BMGT1 protein grown at 18, 16, 14 and 12 hours respectively.

Since the expression seen was almost same in 12-18 hours lysate. The optimum temperature chosen for further study was 18 hours.

Similarly BMGT2 protein was grown at different temperatures (12-28 °C), for different time intervals (4-18 h) and at different IPTG concentrations (0.1- 1 mM) for maximum expression in soluble form. The optimum temperature, time and IPTG

concentration for BMGT2 protein was found to be 21 °C, 18 hours and 0.3 mM respectively (Fig. 4A.13).

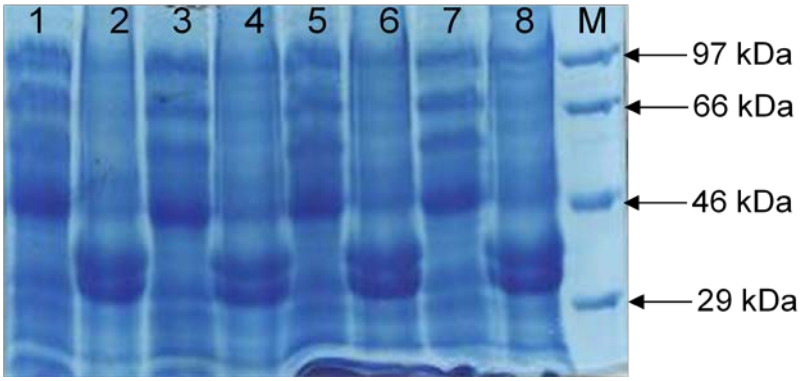


Fig 4A.13: 10% SDS PAGE analysis of BMGT2 protein; Commassie Blue staining : Lane 1,3,5 and 7- Lysate of BMGT2 protein grown at 12, 14, 16 and 18 hours respectively, Lane 2,4,6,8- Inclusion bodies from BMGT2 protein grown at 12, 14, 16 and 18 hours respectively, Lane M- Protein molecular weight marker,

4.3.6 Glycosyltransferase enzyme assay

4.3.6.1 BMGT1 enzyme assay

The activity of BMGT1 protein was checked by glycosylation reaction of various aglycone moieties by different sugar donor molecules. The assay reaction (0.5 mL) contained 20 mM Tris-HCl buffer pH 7.0, 5 mM glycosyl donor (UDP-glucose and UDP-galactose), 600 µM glycosyl acceptor and BMGT1 enzyme. After a 10 minutes preincubation of the mixture without the enzyme at 30 ^{o}C, the reaction was initiated by addition of the BMGT1 enzyme. After incubation at 30 ^{o}C for 3 hours, the reaction was extracted thrice by the addition of equal volume of ethyl acetate. The extracted reaction was then completely vacuum dried and dissolved in appropriate volume of methanol. The substrates and the glycosylated products were then analysed by LC-MS and HPLC. Same reaction was run for control (i.e. lysate of pET 30b (+) vector transformed in *E. coli* BL21) with all the substrates also in order to check the background activity if any.

4.3.6.1a LC-MS

LC-MS was done as described in Chapter 2, Section 2.12.1. Different substrates (Table 2.9) were used in order to check the specificity of the BMGT1 enzyme using different sugar donors like UDP-glucose and UDP-galactose. Almost 22 different substrates belonging to different classes like flavonols, flavones, benzoic acids, isoflavones, flavanones and phenolic compounds were checked using LC-MS but BMGT1 enzyme showed expected mass only with three substrates i.e. genistein, naringenin and kaempferol and only UDP-glucose was able to support the glycosylation reaction. No activity was observed with UDP-galactose. LC-MS profile of genistein control reaction (i.e. lysate of pET-30b (+) vector transformed in *E. coli* BL21 with genistein as substrate and UDP-glucose as donor) Fig. 4A.14, BMGT1 enzyme with genistein as substrate and UDP-glucose as donor (Fig. 4A.15), Naringenin control reaction (i.e. lysate of pET-30b (+) vector transformed in *E. coli* BL21 with naringenin as substrate and UDP-glucose as donor molecule) Fig. 4A.16, BMGT1 enzyme with naringenin as substrate and UDP-glucose as sugar donor (Fig. 4A.17) and BMGT1 enzyme with kaempferol as substrate and UDP-glucose as donor molecule (Fig. 4A.18) are shown below.

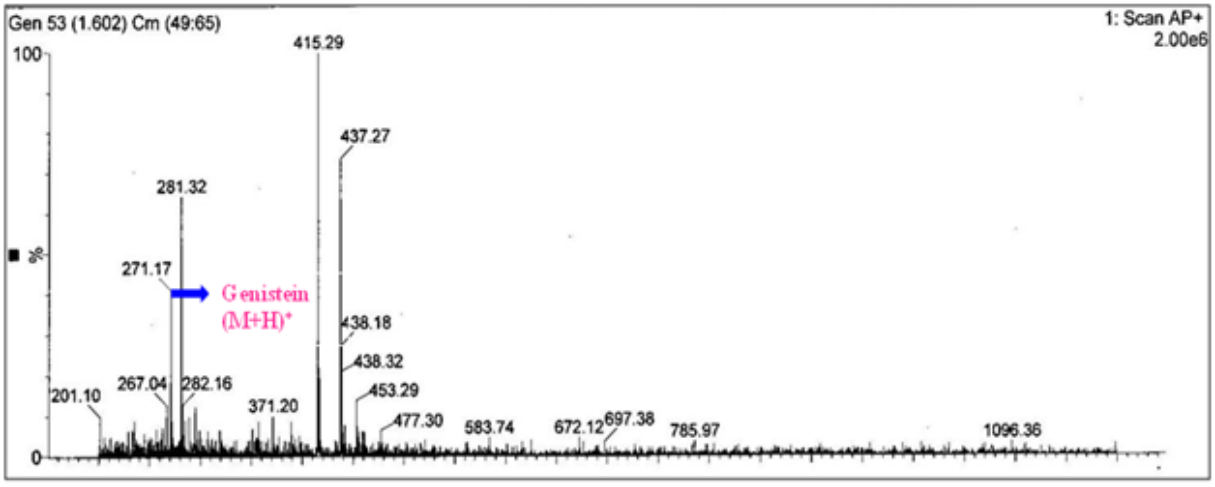


Fig. 4A.14: LC-MS profile of control i.e. lysate of pET 30b (+) with genistein as substrate and UDP-glucose as donor molecule. Peak which is marked with blue arrow indicates the substrate peak i.e. genistein.

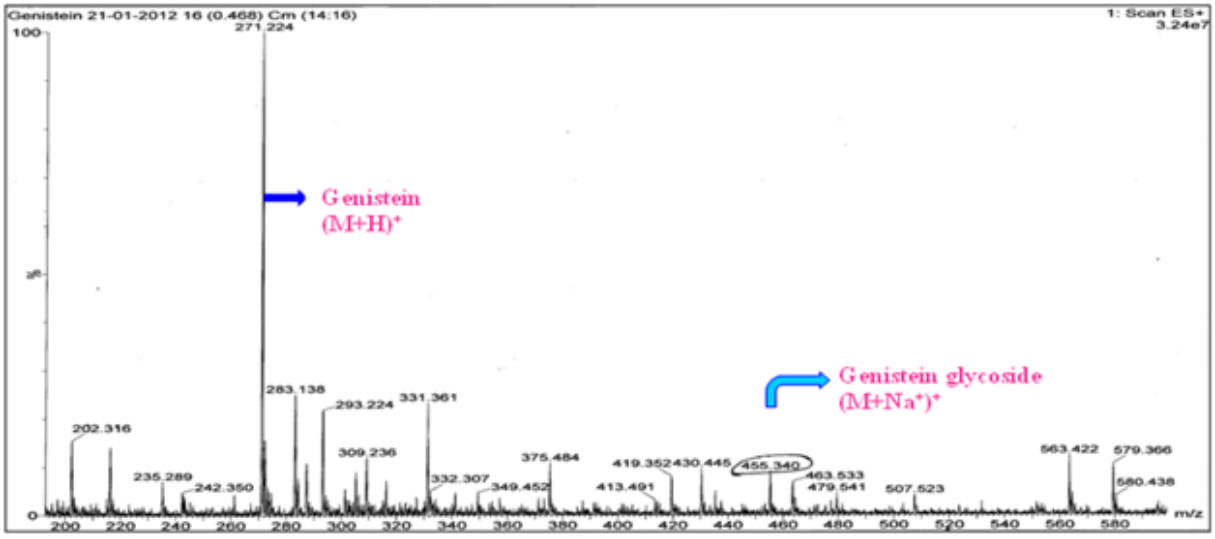


Fig. 4A.15: LC-MS profile of BMGT1 enzyme with genistein as substrate and UDP-glucose as donor molecule. Peak which is marked with blue arrow indicates the substrate peak i.e. genistein and peak which is marked with turquoise is the product i.e. genistein glycoside.

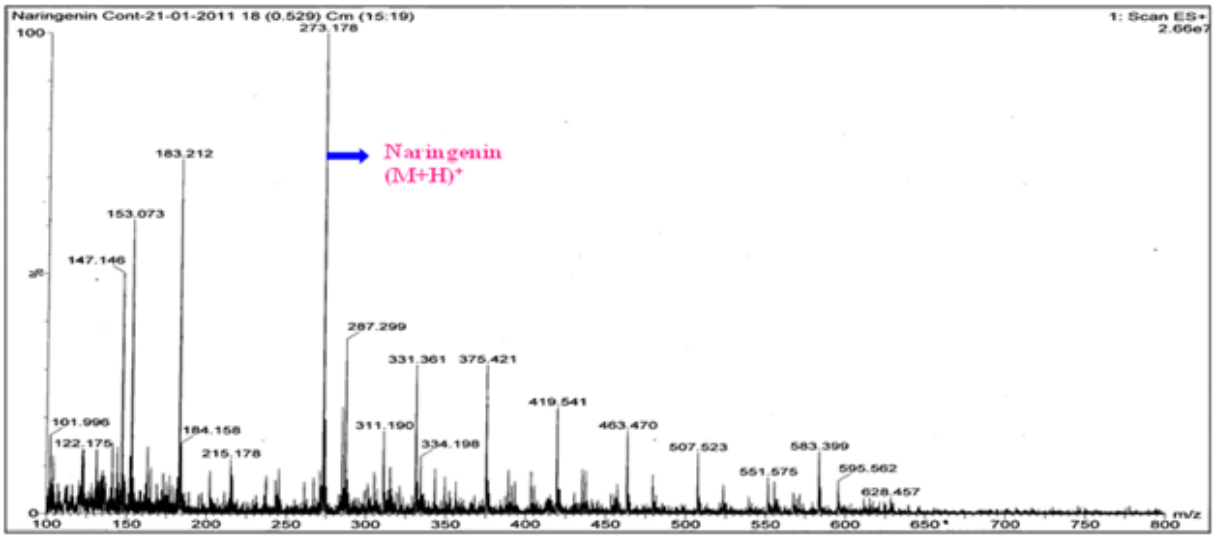


Fig. 4A.16: LC-MS profile of control i.e. lysate of pET 30b (+) with naringenin as substrate and UDP-glucose as donor molecule. Peak which is marked with blue arrow indicates the substrate peak i.e. naringenin.

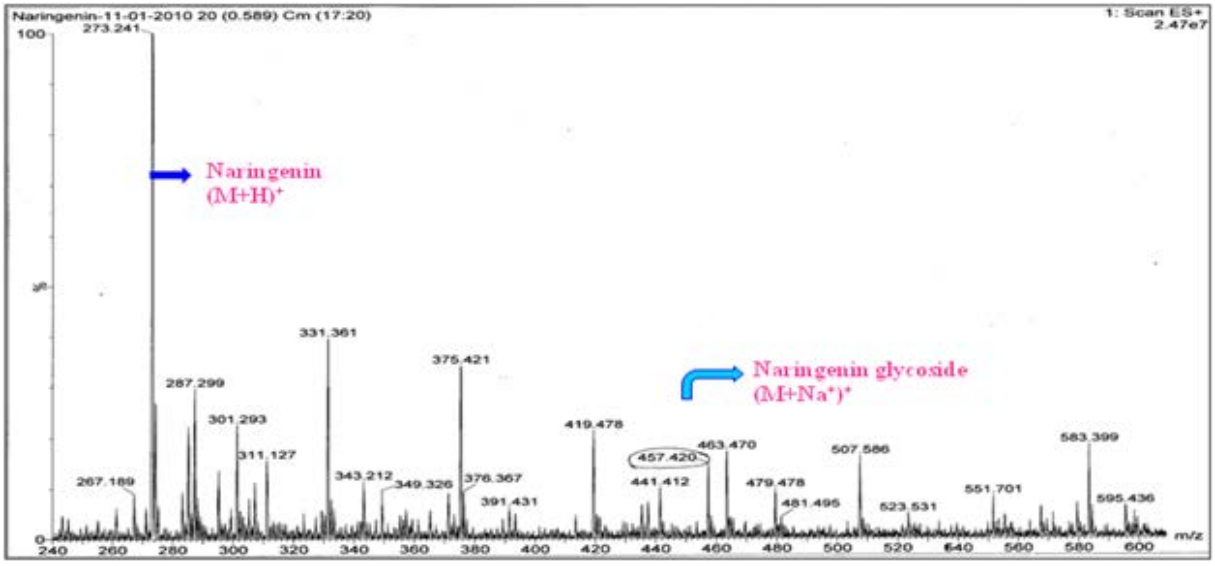


Fig. 4A.17: LC-MS profile of BMGT1 enzyme with naringenin as substrate and UDP-glucose as donor molecule. Peak which is marked with blue arrow indicates the substrate peak i.e. naringenin and peak which is marked with turquoise is the product i.e. naringenin glycoside.

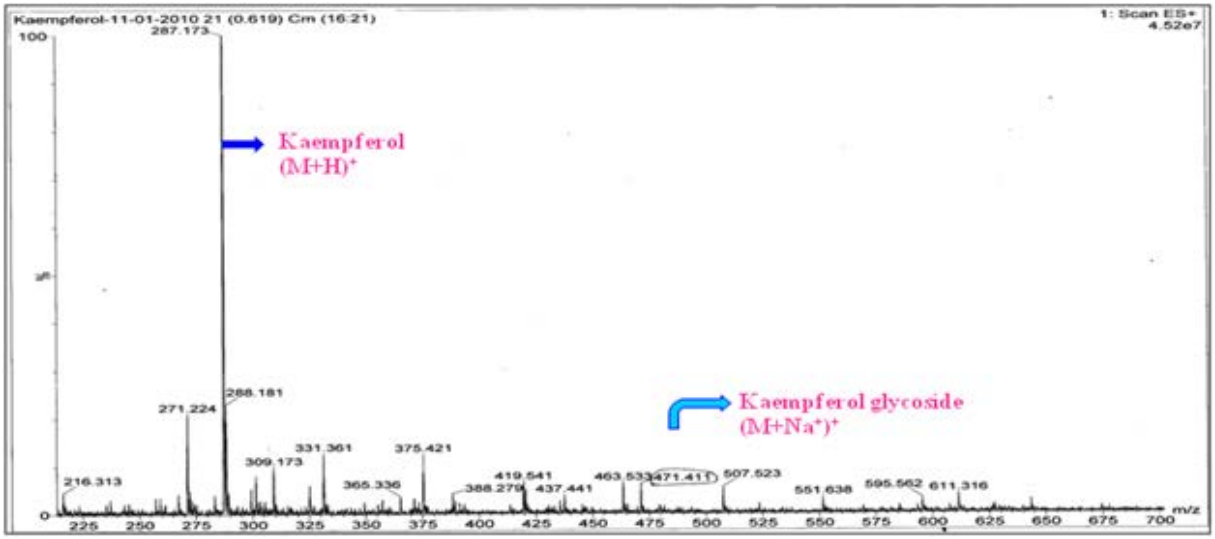


Fig. 4A.18: LC-MS profile of BMGT1 enzyme with kaempferol as substrate and UDP-glucose as donor molecule. Peak which is marked with blue arrow indicates the substrate peak i.e. kaempferol and peak which is marked with turquoise is the product i.e. kaempferol glycoside.

The LC-MS results showed that BMGT1 enzyme can glycosylate only three substrates namely genistein, naringenin and kaempferol out of 22 substrates tested (Table: 4A.1) as indicated by their expected mass in LC-MS profile and it accepts only UDP-glucose as sugar donor but the control reactions fail to produce respective glycosides which proved that *E.coli* did not show any background activity of its own. These results were further confirmed by HPLC.

Table 4A.1: List of substrates tested for LC-MS

Substrate	LC-MS result	Substrate	LC-MS result
Apigenin	_	Diosgenin	_
Daidzein	_	Capsaicin	_
Naringenin	+	Vannilic acid	_
Catechin	_	Salicylic acid	_
Myricetin	_	Resorcinol	_
Isorhamnetin	_	Thymol	_
Genistein	+	Euginol	_
Luteolin	_	Benzoic acid	_
Hesperetin	_	4-nitro phenol	_
Curcumin	_	Methyl vanilate	_
Kaempferol	+	3,4dimethoxyphenol	_

4.3.6.1b HPLC

On the basis of LC-MS results genistein, naringenin and kaempferol were used as substrates and only UDP-glucose was used as donor molecule for HPLC analysis for further confirmation. Analysis of the above mentioned substrates and their glycosylated products was done by using reversed-phase HPLC (Perkin Elmer series 200) on a supelco C18 column (5µm, 25cm x 4.6 mm). For genistein mobile phase consisted of 50 mM Tris buffer (pH 8.0). 20 µL of sample dissolved in methanol was injected in HPLC column and was programmed as follows: 10% acetonitrile for 0.5 min, 40% acetonitrile for 10 min and 60% acetonitrile for

5 min at a flow rate of 1.0 mL/min. A dual wavelength recorder set at 270 nm and 340 nm was used to detect the compounds eluting from the column.

For naringenin mobile phase consisted of milliQ water. 20 µL of sample dissolved in methanol was injected in HPLC column and was programmed as follows: 20% acetonitrile for 0.5 min, 40% acetonitrile for 5 min, 70% acetonitrile for 5 min, 90% acetonitrile for 5 mins and 20% acetonitrile for 5 mins at a flow rate of 1.0 mL/min. A dual wavelength recorder set at 270 nm and 340 nm was used to detect the compounds eluting from the column.

For kaempferol mobile phase consisted of milliQ water. 20 µL of sample dissolved in methanol was injected in HPLC column and was programmed as follows: 10% acetonitrile for 0.5 min, 10% acetonitrile for 5 mins, 30% acetonitrile for 5 min, 75% acetonitrile for 5 min and 95% acetonitrile for 5 mins at a flow rate of 1.0 mL/min. Compounds eluted from the column were detected at 420 nm.

Genistein gets glycosylated mainly at two positions and can form two products i.e. genistein 7-glycoside and genistein 4-glycoside depending on the specificity of the glycosyltransferase enzyme. Whereas, naringenin and kaempferol gets glycosylated at 3 positions i.e. 3,7 & 5-OH positions giving rise to respective glycosides. In that case HPLC profiles need to be generated of substrate as well as all the products it produces in order to find out the specificity of the glycosyltransferase enzyme (BMGT1).

HPLC profile of genistein (substrate) Fig. 4A.19, genistein 7-glycoside (product) Fig. 4A.20, genistein 4-glycoside (product) Fig. 4A.21, Assay reaction i.e. BMGT1 enzyme with genistein as substrate and UDP-glucose as sugar donor molecule Fig. 4A22 and control reaction (i.e. lysate of pET 30b (+) transformed in *E. coli* BL21 cell line with genistein as substrate and UDP-glucose as donor molecule) Fig. 4A.34. HPLC profile of naringenin (substrate) Fig. 4A.24, naringenin 7-glycoside (product) Fig. 4A.25, Assay reaction i.e. BMGT1 enzyme with naringenin as substrate and UDP-glucose as sugar donor molecule Fig. 4A.26.

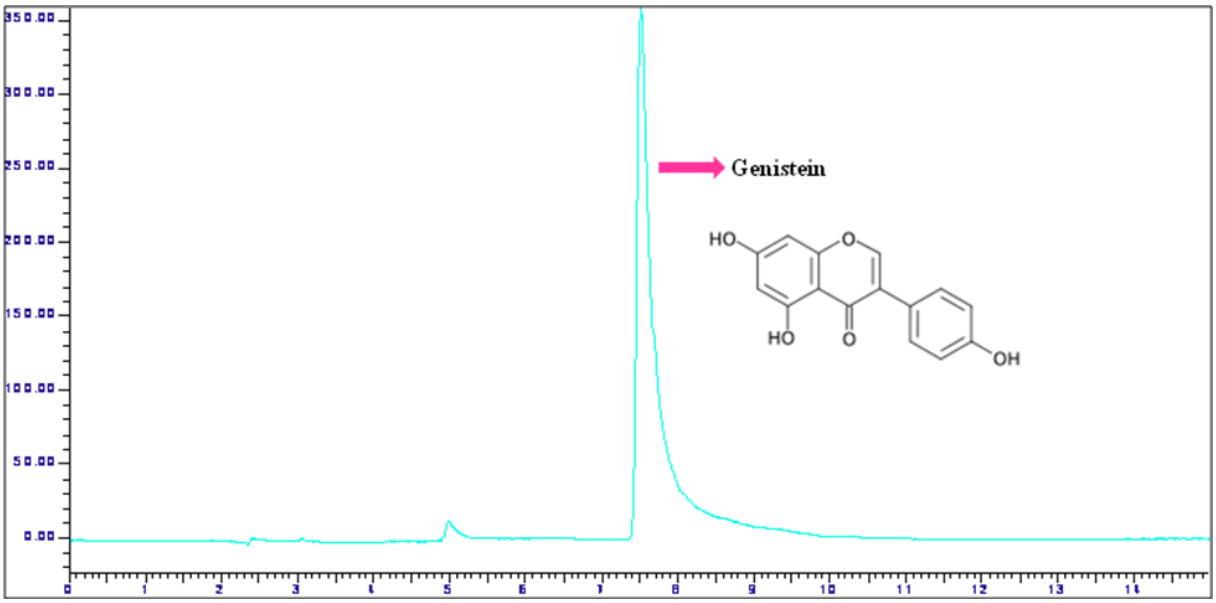


Fig. 4A.19: HPLC profile of standard substrate (genistein). Peak which is marked with pink arrow indicates the substrate peak i.e. genistein with Retention time (t_R) of 7.6.

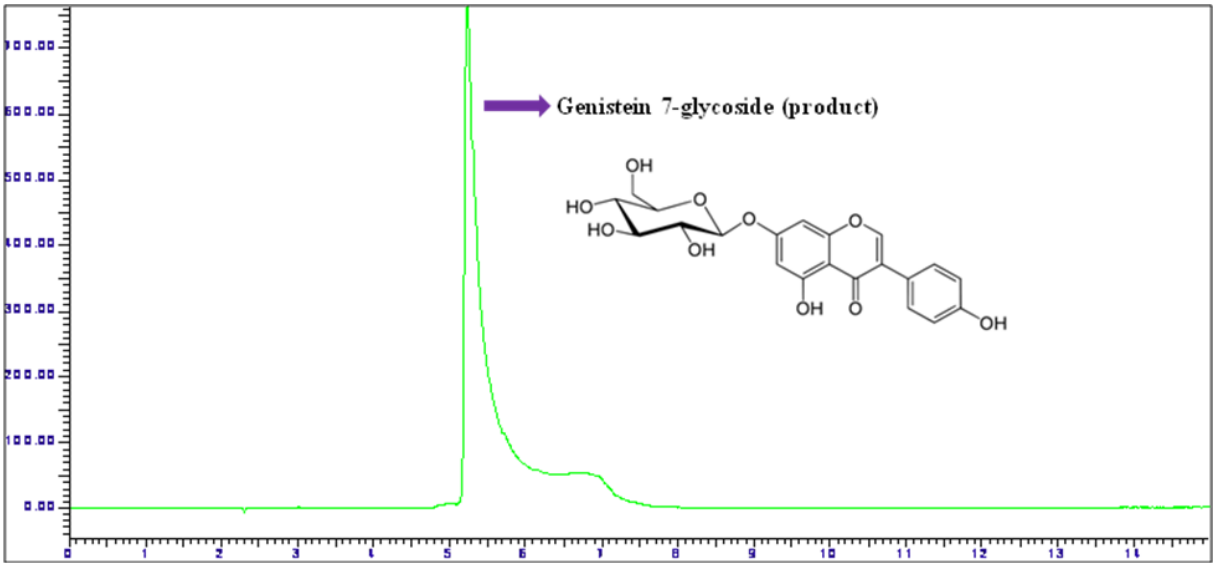


Fig. 4A.20: HPLC profile of standard product (genistein 7-glycoside). Peak which is marked with purple arrow indicates the product peak i.e. genistein 7-glycoside (t_R -5.2).

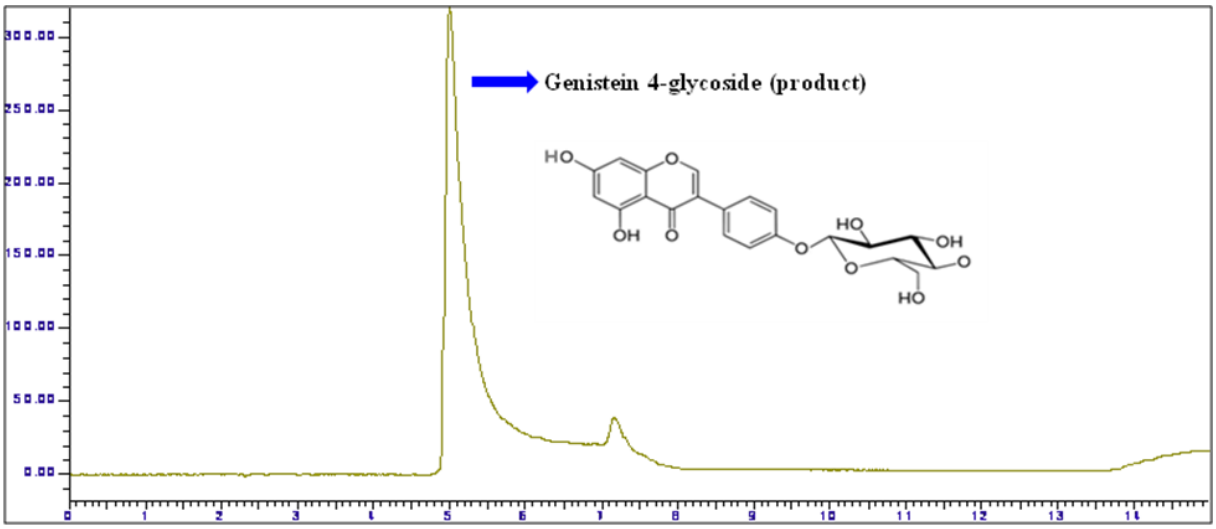


Fig. 4A.21: HPLC profile of standard product (genistein 4-glycoside). Peak which is marked with blue arrow indicates the product peak i.e. genistein 4-glycoside (t_R -4.9).

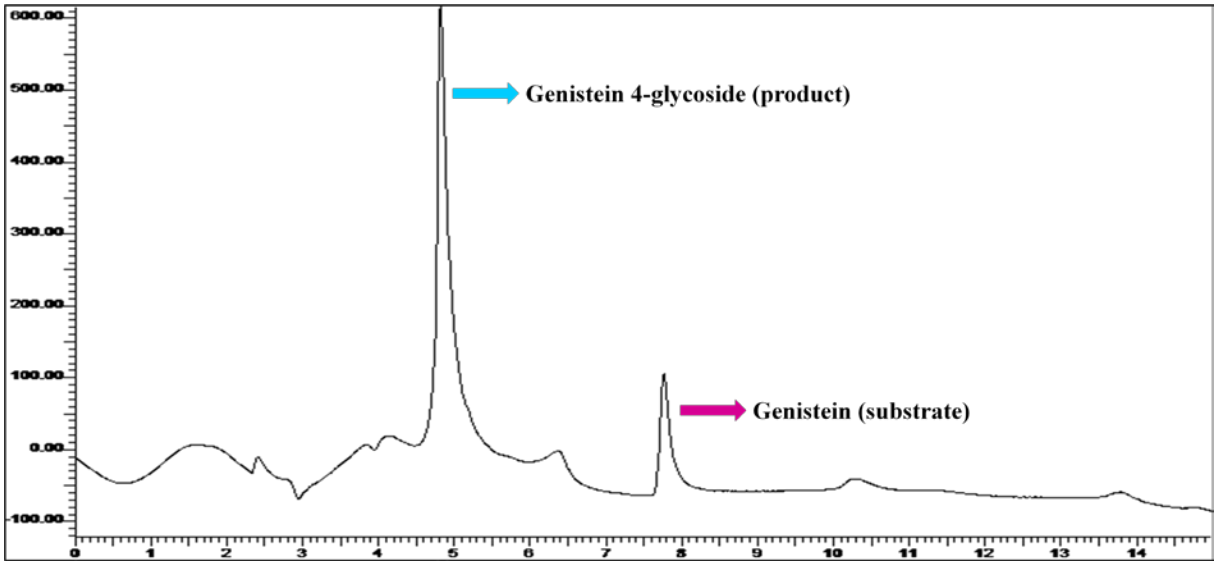


Fig. 4A.22: HPLC profile of assay reaction i.e. BMGT1 enzyme with genistein as substrate and UDP-glucose as donor molecule. Peak which is marked with turquoise arrow indicates the product peak i.e. genistein 4-glycoside and the peak which is marked by pink arrow indicates unused substrated genistein.

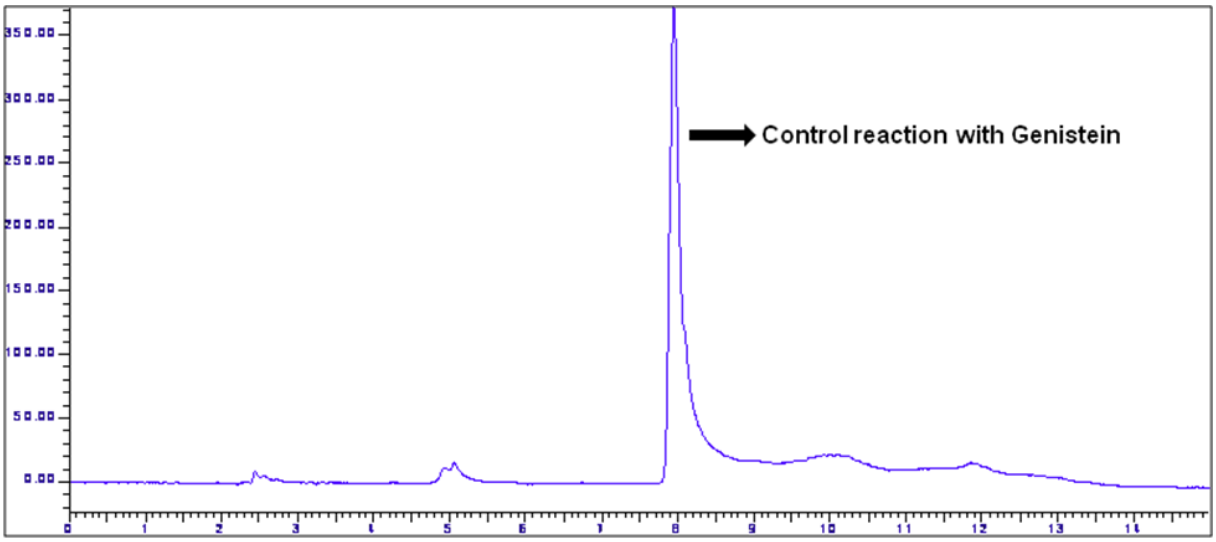


Fig. 4A.23: HPLC profile of control reaction (i.e. lysate of pET 30b (+) vector transformed in *E. coli* BL21 cells with genistein as substrate and UDP-glucose as donor molecule). Peak which is marked with black arrow indicates the substrate peak i.e. genistein.

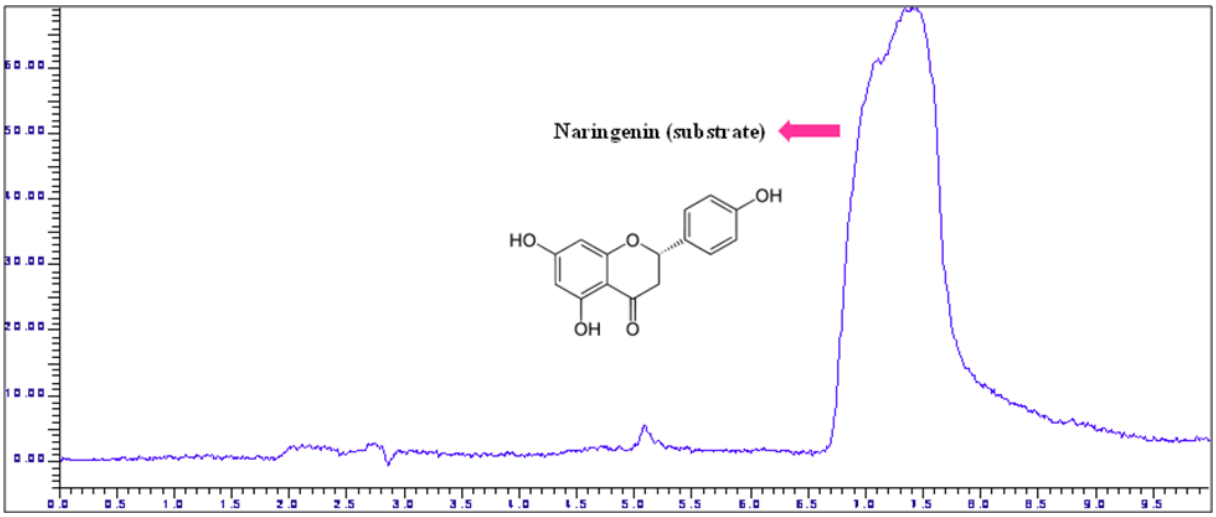


Fig. 4A.24: HPLC profile of standard substrate (naringenin). Peak which is marked with pink arrow indicates the substrate peak i.e. naringenin (t_R -7.4).

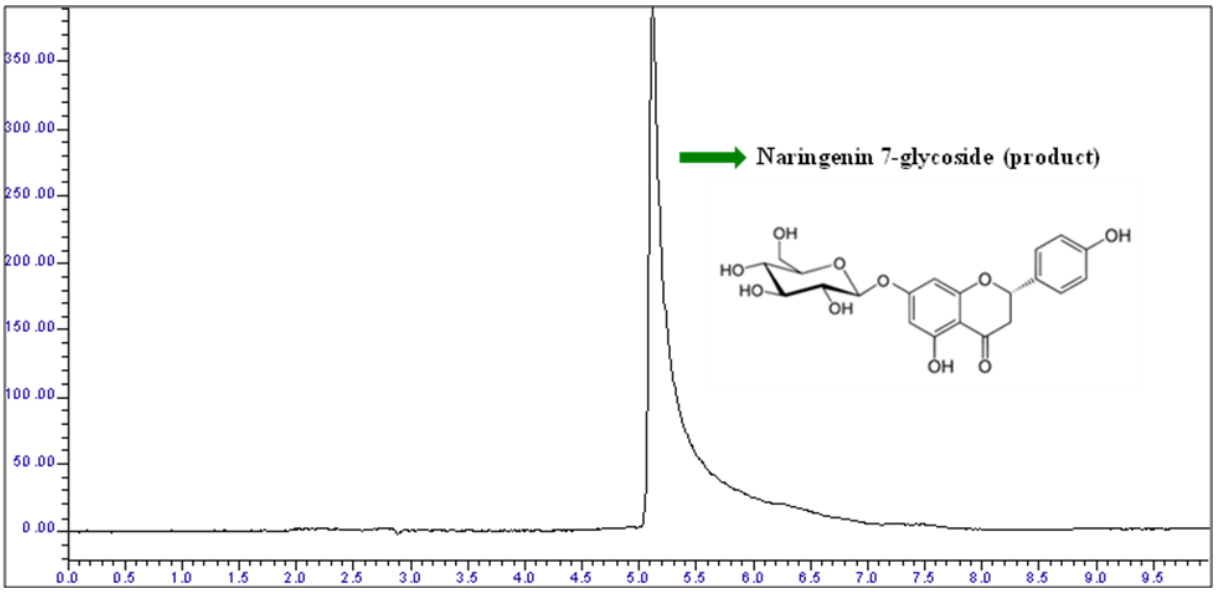


Fig. 4A.25: HPLC profile of standard product (naringenin 7-glycoside). Peak which is marked with dark green arrow indicates the product peak i.e. naringenin 7-glycoside (t_R -5.1).

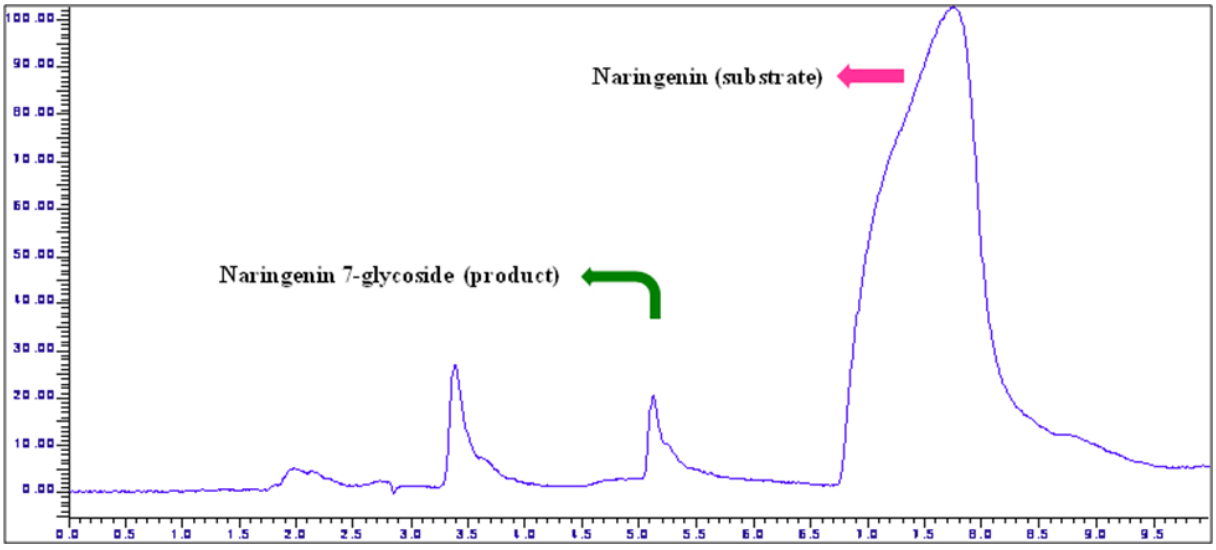


Fig. 4A.26: HPLC profile of assay reaction i.e. BMGT1 enzyme with naringenin as substrate and UDP-glucose as donor molecule. Peak which is marked with dark green arrow indicates the product peak i.e. naringenin 7-glycoside and the peak which is marked by pink arrow indicates unused substrated naringenin.

HPLC was run for all the three substrates i.e. genistein, naringenin and kaempferol. Genistein could provide the expected results whereas in case of naringenin and kaempferol, relative specificity of BMGT1 was low towards the respective substrates. HPLC results also showed that BMGT1 enzyme is specific to genistein only and out of two products it could produce only one products i.e. genistein 4-glycoside (also referred to as sophoricoside) which is evident from their retention time (retention time of genistein 4-glycoside is 4.9 min and retention time of genistein 7-glycoside is 5.2 min). Complementing LC-MS result, here also control reaction could not generate any product which indicates that *E. coli* does not produce any background activity of its own.

4.3.6.2 BMGT2 enzyme assay

The activity of BMGT2 protein was also checked by glycosylation reaction as done with BMGT1 protein. The assay reaction was run as described in section 4.3.6.1. All the reactions were analysed by LC-MS.

LC-MS results did not show expected mass in any of the reactions where UDP-glucose and UDP-galactose were used as sugar donors. This result was quite expected as the gene sequence of BMGT2 itself showed that it has conserved region which is specific to UDP-glucuronic acid as donor. The assay with UDP-glucuronic acid could not be performed since the substrates which accept glucuronic acid as donor molecule were not available so only way specificity of BMGT2 enzyme could be predicted is through *in silico* studies i.e. homology modelling and docking studies. Homology and docking studies have been discussed in detail in chapter 5 of this thesis.

4.3.7 Protein extraction from plant tissue

Plant tissue was harvested from *B. monniera* and was crushed to a fine powder in liquid nitrogen. Ground tissue was homogenized in buffer (100 mM Tris-HCl pH.7.5, 2% PVPP, 2% PEG 4000, DTT 5 mM and PMSF 1 mM). Homogenized tissue was centrifuged at 12000 rpm at 4 °C and supernatant was collected in fresh microfuge tube. Supernatant was quantified using Bradford assay (Bradford reagent, Sigma).

4.3.7.1 Western blot analysis

The recombinant protein expressed in Chapter 4, Section 4.3.2 was used to raise polyclonal antibody against BMGT1 protein in rabbit, Chapter 4, Section 4.3.3. The above raised anti-BMGT1 polyclonal antibody obtained from 3rd bleed serum was diluted 1:10000 times and was used for western blot experiments. Secondary antibody i.e. goat anti-rabbit IgG conjugated with alkaline phosphatase was purchased from Bangalore Genei, Bangalore, India. iBlot Gel Transfer System was used to transfer the protein from SDS-PAGE gels or from slot blot membranes by strictly following manufacturer's (Invitrogen, USA) instructions. The PVDF membrane containing transferred protein samples were processed as per standard procedure, blocking, treatment with primary antibody (anti-BMGT1 antibody Chapter 4, Section 4.3.3), washing followed by treatment with secondary antibody which is ALP conjugated, followed by washing and colour development using BCIP/NBT as substrate. The samples subjected for western analysis includes purified recombinant BMGT1 protein expressed in pET 30b (+) expression system and crude protein from *B. monniera* plant tissue. Total crude protein used for western blot experiment was 200 µg. Purified protein of BMGT1 was also used as positive control. The plant extract was run on 10% SDS-PAGE and blotted on to PVDF membrane (Fig. 4A.27).

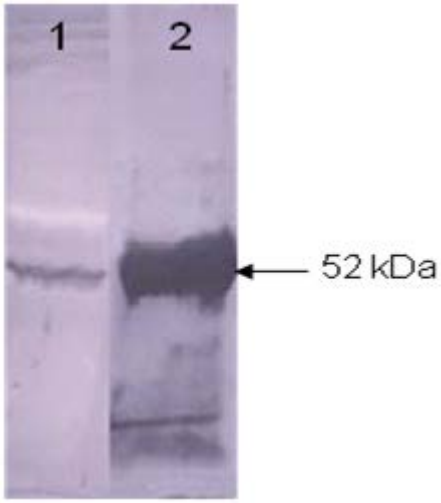


Fig. 4A.27: Western blot analysis: Lane 1-*B. monniera* plant extract and Lane 2- Purified recombinant BMGT1 protein expressed in pET 30b (+) expression vector.

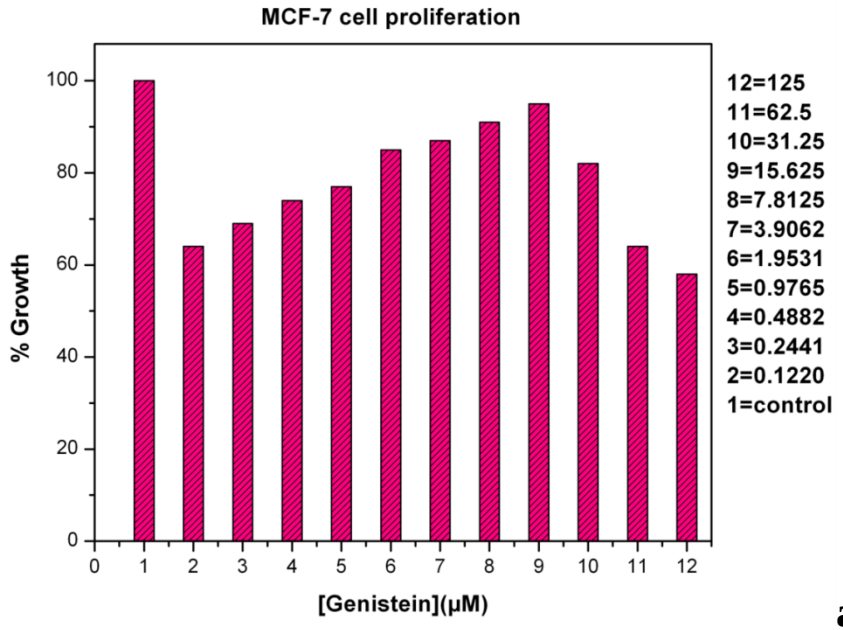
In Western blot, in spite of using polyclonal antibody raised against BMGT1 protein, only one protein band was detected in the crude plant extract of *B. monniera.* This band corresponds to the same molecular weight (~52 kDa) as that of BMGT1 protein which may be due to very less cross reactivity of the antibody against other GTs present in the *B. monniera* plant.

4.3.8 Cytotoxicity test of genistein and genistein 4-glycoside on different cancer cell lines

With reference to the anti-cancerous properties of the aglycone moiety (i.e. genistein), comparative study of aglycone part and its glycosylated product has been done by cell proliferation MTT assay using MCF-7 and THP-1 cell lines. As mentioned in earlier reports the activity of aglycone and its glycosylated product may or may not be same (One et al., 2010). A lot of work has been done on genistein and its glycosylated product genistein 7-glycoside but not much has been done on its other product i.e. genistein 4-glycoside. So in order to check whether genistein and genistein 4-glycoside exhibit same activity or not on cancer cell lines their cytotoxicity test was performed. MTT assay was done as described in chapter 2 section 2.11. Two cancer cell lines were used MCF-7 and Thp-1 for the assay.

4.3.8.1 Effect of genistein and sophoricoside (genistein 4-glycoside) on MCF-7 and Thp-1 cell proliferation

To determine the minimum concentration of genistein and sophoricoside that will stimulate or inhibit MCF-7 and Thp-1, we conducted a cell proliferation dose-response study. MCF-7 and Thp-1 cells were monitored in response to various concentrations of genistein and sophoricoside ranging from 0.12-125 µM (Fig. 4A.28a & 28b for MCF-7 and Fig. 4A.29a & 29b for Thp-1 cell lines). Data are expressed as a percentage of the control cell cultures. These levels were chosen because the genistein blood levels reported in animals and humans consuming diets high in genistein (such as soy-containing diets) have blood concentrations ranging from 0.1-6 µM.



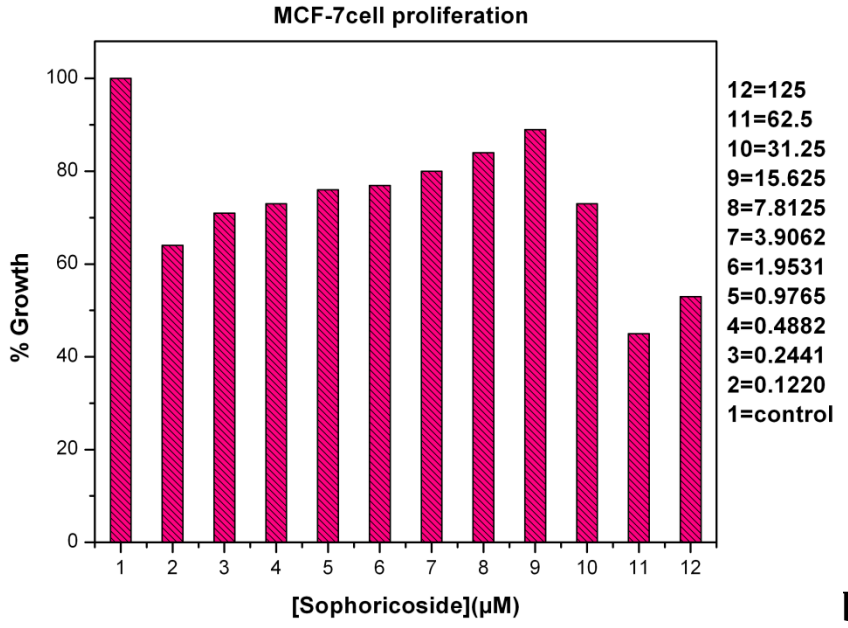
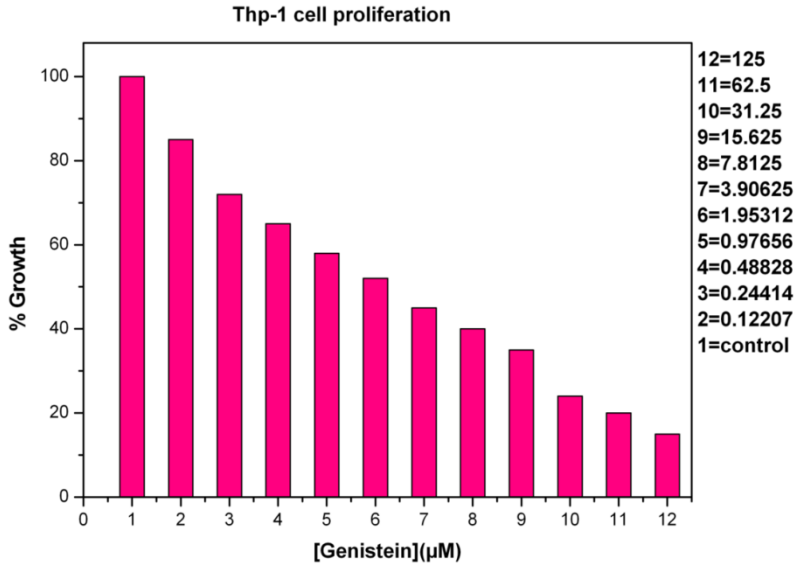
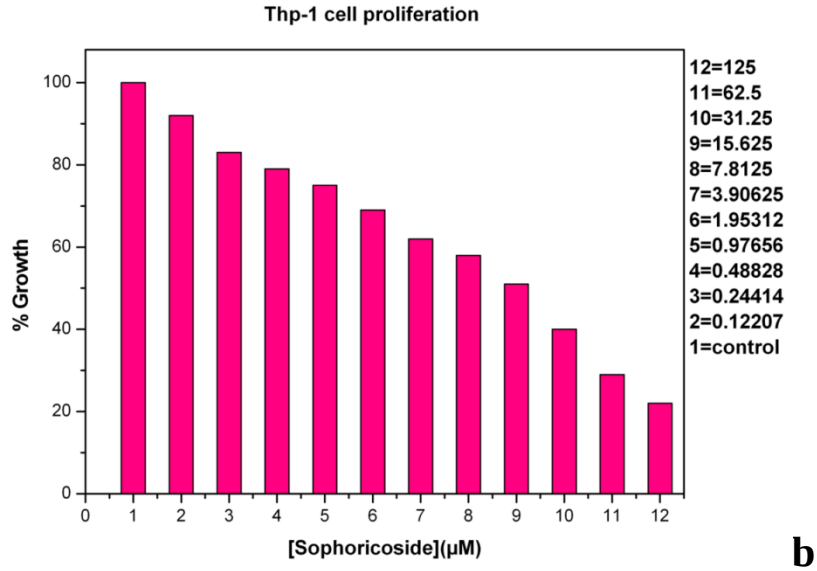


Fig. 4A.28: Effects of genistein **(a)** and sophoricoside **(b)** on the cell growth of MCF-7 cells. MCF-7 cells were cultured in the presence of various concentrations of genistein and sophoricoside (0.12 μM to 125 μM) for 120 h in MEM containing 10% FBS at 37 °C in a 95% humidified atmosphere of 5% CO, in air. Proliferation was assessed by MTT dye reduction analysis. MTT assay was measured at 490nm.

Genistein and sophoricoside increased cell growth in a dose-dependent manner in the range of 0.01-15 µM. Maximal growth stimulation was observed at 15 µM. In contrast, higher concentrations (30-125 µM) of genistein and sophoricoside produced a dose-dependent decrease in cell growth when compared with that of untreated controls. Here substrate shows same level of inhibition as product.



a



b

Fig. 4A.29: Effects of genistein **(a)** and sophoricoside **(b)** on the cell growth of Thp-1 cells.

Genistein and sophoricoside gradually decreases the cell proliferation when compared with that of untreated controls. It follows proper sigmoidal curve. In Thp-1 cells substrate shows higher inhibition rates than product.

4.4 Conclusions

- The BMGT1 and BMGT2 genes isolated from *B. monniera* were directionally cloned in pET 30b (+) expression system.

- Recombinant BMGT1 and BMGT2 proteins were standardized for over-expression and purified from inclusion bodies. A 52 kDa purified protein of BMGT1, from inclusion bodies, was used to raise polyclonal antibodies in New Zealand rabbit.

- Optimization of expression of native protein was done for both BMGT1 and BMGT2 proteins. The best expression was achieved when bacterial cultures were grown in LB media for 15 ºC for 18 h after induction with 0.1 mM IPTG whereas in case of BMGT2 protein best expression was obtained at 21 ºC for 18 h after induction with 0.3 mM IPTG.

- BMGT1 enzyme activity for substrates like genistein, naringenin and kaempferol was determined using crude lysate of BMGT1 protein. Data showed that recombinant BMGT1 enzyme showed maximum activity with genistein producing only one product i.e. genistein 4-glycoside rather than two and accepts only UDP-glucose as donor molecule.

- BMGT2 enzyme assay could not be performed because of the non availability of substrates so only way to know its specificity is by *in silico* method i.e. homology modeling and docking studies explained in chapter 5 of this thesis.

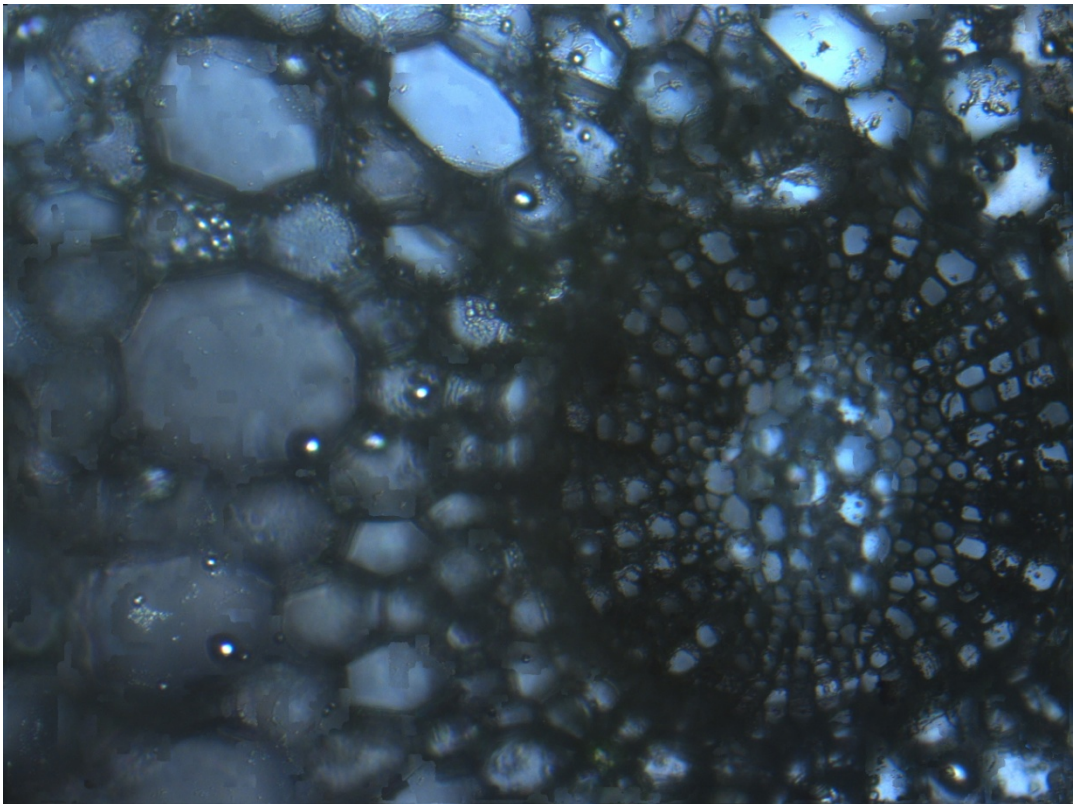
- The candidate protein, glycosyltransferase was detected in *B. monniera* by Western blotting.

- The functional expression of recombinant BMGT1 and BMGT2 proteins confirms that the cDNA isolated from *B. monniera* encodes for glycosyltransferase protein.

- Cytotoxicity test of genistein and genistein 4-glycoside was performed in order to check their effects on cancer cell lines. In MCF-7 cell line the both substrate and product are showing same inhibition effects on cell proliferation whereas in case of Thp-1 cell line substrate shows higher inhibition rates than product.

Chapter 4B

Tissue Specific Expression & Immuno-cytolocalization of *B. monniera* GT gene(s)



4.5 Introduction

Although the role of different glycosyltransferase genes from different plants have been elucidated, our knowledge is still scarce on how these GTs work to glycosylate different secondary metabolites and how plant metabolism responds to changes in the expression of glycosyltransferase genes. With the advent of genomic tools that enable unbiased transcriptome and metabolome wide analyses, such interactions can now be elucidated.

In the present study, we used reverse transcription followed by polymerase chain reaction (real-time RT-PCR), since the method is very sensitive and specific when differentiation between highly similar transcripts is needed (Huggett *et al*., 2005).

In the present study, expression of the identified glycosyltransferase genes was performed in more detail in *B. monniera* to verify the expression patterns and to explore their roles under different stress conditions. Studying different aspects of GTs like transcript analysis using RT-PCR and immuno-cytolocalization can provide insight in our understanding.

4.6 Materials and methods

4.6.1 Plant material

B. monniera shoots were treated as described in Chapter 2: section 2.1.1. The treated explants were transferred to MS basal medium for about 20 days till proliferating multiple adventitious shoots were separated and re-inoculated on proliferation medium for further growth. Proliferated shoots were then used for stress treatments. Root, shoot and leaves were harvested separately after 2 months and were used for Real-Time PCR analysis.

4.6.2 Primary antibodies for anti BMGT1 protein

Purified BMGT1 protein was used to raise antibodies in rabbit as discussed in Chapter two, Section 2.9 and 2.10.

4.6.3 Secondary antibody

Alkaline phosphatase tagged goat anti-rabbit IgG antibodies were purchased from Bangalore, genei.

4.7 Methods

4.7.1 Total RNA extraction and cDNA synthesis

Total RNA was isolated from experimental *in-vitro* shoots according to the RNA purification kit (Sigma). First strand cDNA synthesis, primed with an oligo $(dT)_{15}$ primer, was performed with Avian Myeloblastosis Virus Reverse Transcriptase (AMV-RT) according to the manufacturer's protocol (Promega Corp., Madison, USA).

4.7.2 Quantitative Real-Time PCR (QRT-PCR)

The cDNA first strands were used as template for semi and absolute quantification of GT gene transcripts. Brilliant II™ SYBR® Green Q-PCR master mix (Stratagene, Agilent Technology, TX, USA) was used for real-time PCR. The reaction mixture was prepared as described in chapter 2; section 2.8.5.2.

4.7.3 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 Ct method is also known as the $2^{-\Delta\Delta Ct}$ method, where
$\Delta\Delta C_t = \Delta$ Ct, sample - Δ Ct, reference.

Here, Δ Ct, $_{sample}$ is the Ct value for any sample normalized to the endogenous housekeeping gene and Δ Ct, $_{reference}$ is the Ct value for the calibrator also normalized to the endogenous housekeeping gene.

Means,

Δ Ct, $_{sample}$ = Ct, $_{sample}$ – Ct, $_{endogenous\ control\ (18S\ rRNA\ gene\ or\ any\ other\ housekeeping\ gene)}$

Δ Ct, $_{reference}$ = Ct, $_{reference/\ normal/\ untreated\ sample}$ – Ct, $_{endogenous\ control.}$

For the $\Delta\Delta$Ct 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 Δ Ct varies with template dilution. If the plot of cDNA dilution versus Δ 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.

4.7.3.1 QRT-PCR considerations: See chapter 2: section 2.8.5.1

4.7.3.2 Preparing the QRT-PCR reactions: See chapter 2: section 2.8.5.2

4.7.3.3 Real-time quantitative PCR for *GT* genes

Total RNA was extracted individually from roots, stem and leaves from *in-vitro* shoots of *B. monniera* and also from the plants under stress. One µg of total RNA was used for making cDNA using ImProm cDNA synthesis kit (Promega, Madison, USA). Brilliant SYBRGreen Q-PCR kit (Stratagene, USA) and Stratagene Mx3000P real time machine were used for all reactions. The primer sequences that were designed for *B. monniera GT* genes and 18S rRNA 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. QRT-PCR reactions were performed following conditions mentioned in Chapter 2: section 2.8.5.3 with annealing temperature of 55 ^{o}C. The reaction was run in triplicates and repeated twice. It was ensured

that equal quantity of RNA template was used for each reaction (Pfaffl, 2001; Freeman *et al*., 1999; Edyta Zdunek-Zastocka, 2008).

Table 4B.1 Primer sequences designated for *B. monniera GT* genes and 18S rRNA

Primers	**Sequence 5'- 3'**	**Tm in ºC**
RBMGT1F	AAA CAG CTC CAT TAG CAT AGA AAG C	49.3
RBMGT1R	GCC TCT CTC CTA AAG CGG TTG	51.2
RBMGT2F	TCC CGA CTT GGA CGC TTT G	60.0
RBMGT2R	CCG AGC ACC TCC TTC TGA G	62.0
Ctrl 18S F	GCA CGC GCG CTA CAA TGA AAG TAT	57.0
Ctrl 18S R	TGT GTA CAA AGG GCA GGG ACG TAA	57.1

4.8 Immuno-cytolocalization of GT protein in *B. monniera*

Solutions: See Appendix: Table 2.10

Protocol: See Chapter 2: Section 2.10.3

4.9 Results and discussion

Expression studies were done using *in vitro* seedlings of *B. monniera.* Stratagene Mx 3000P real time machine was used for expression studies.

4.9.1 Spatio-Temporal expression pattern of *GT* genes (BMGT1 and BMGT2), Real-time PCR

Temporal expression of BMGT1 and BMGT2 genes was carried out by isolating total RNA from leaves, root and stem of developing plants of *B. monniera* and plants which were subjected to different stresses like 50 mM NaCl, 20 µM methyl jasmonate, 30 µM salicylic acid and 2% mannitol, heat (37 °C) and cold (-20 °C) shock treatments. Control plant was

without any stress treatment. Since salicylic acid was dissolved in ethanol so a separate control plant was grown for salicylic acid treated with ethanol.

4.9.1a QRT-PCR results for BMGT1 gene

Real-time Q-PCR was performed using RBMGT1F and RBMGT1R as primers and cDNA isolated from untreated and treated plants. Ct value is inversely proportional to the level of expression. The relative expression of the BMGT1 gene in terms of mean fold expression ($2^{-\Delta\Delta Ct}$) is shown in Fig. 4B.1 (mannose stress with respect to control i.e. untreated plant), Fig. 4B.2 (methyl jasmonate stress with respect to control plant), Fig. 4B.3 (salt stress with respect to control plant), Fig. 4B.4 (cold shock with respect to control plant), Fig. 4B.5 (heat shock treatment with respect to control plant), Fig. 4B.6 (salicylic acid stress with respect to salicylic acid control plant).

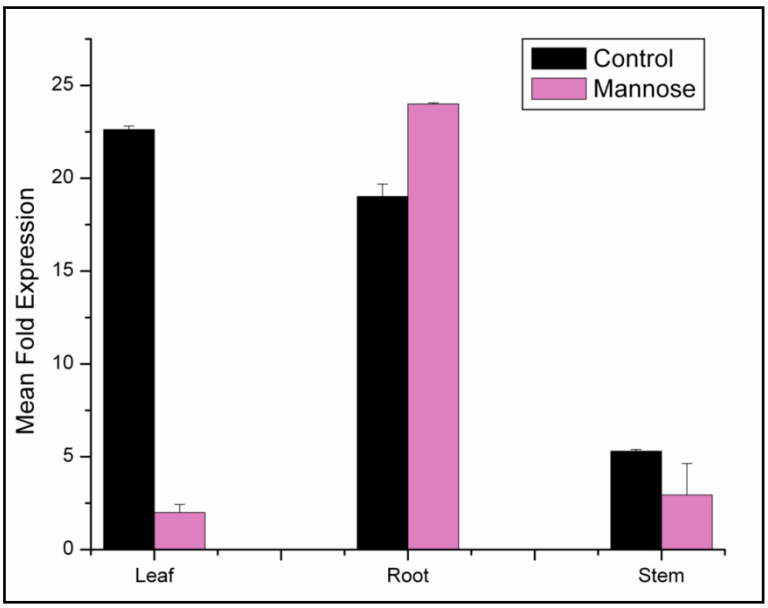


Fig. 4B.1: Relative expression in terms of mean fold expression of BMGT1 gene under mannose stress with respect to control plant. All values are plotted with standard deviation taken into account.

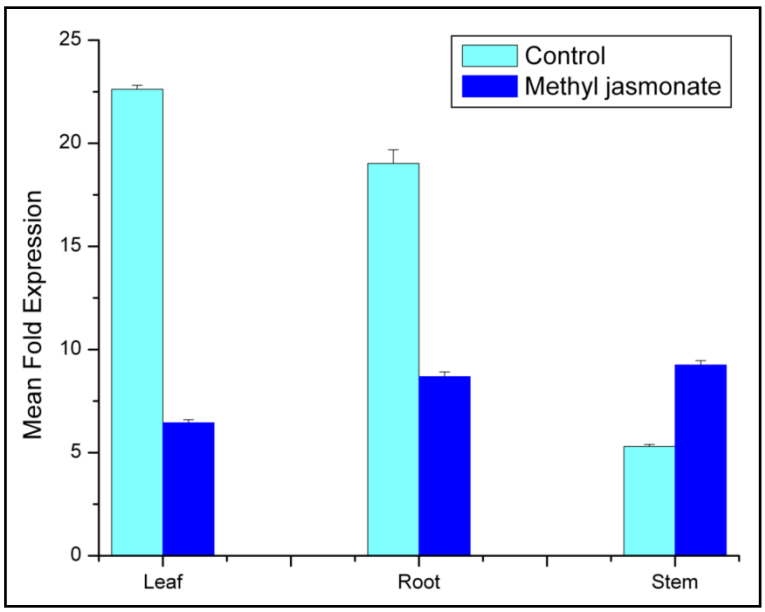


Fig. 4B.2: Relative expression in terms of mean fold expression of BMGT1 gene under methyl jasmonate stress with respect to control plant. All values are plotted with standard deviation taken into account.

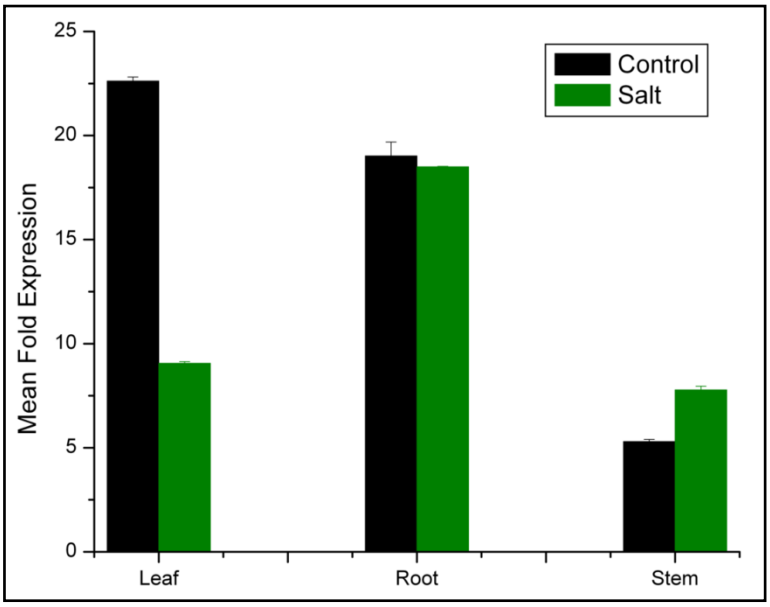


Fig. 4B.3: Relative expression in terms of mean fold expression of BMGT1 gene under salt stress with respect to control plant. All values are plotted with standard deviation taken into account.

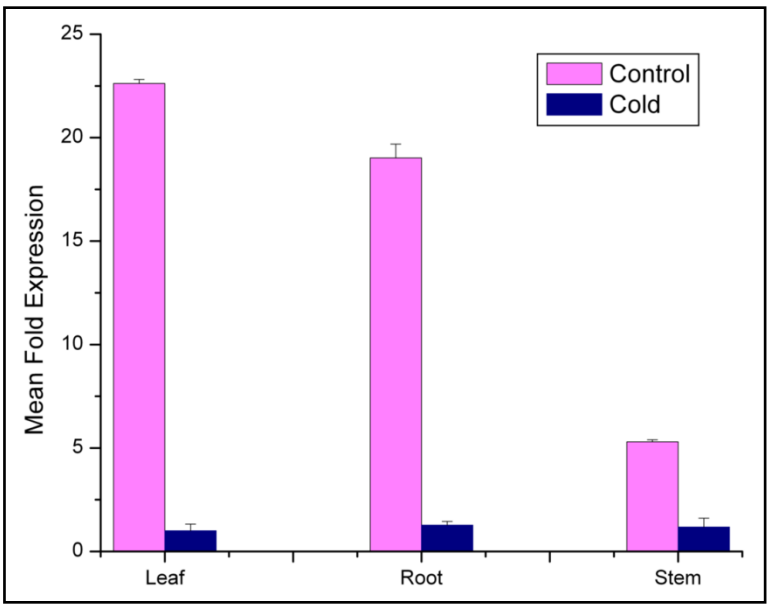


Fig. 4B.4: Relative expression in terms of mean fold expression of BMGT1 gene under cold shock with respect to control plant. All values are plotted with standard deviation taken into account.

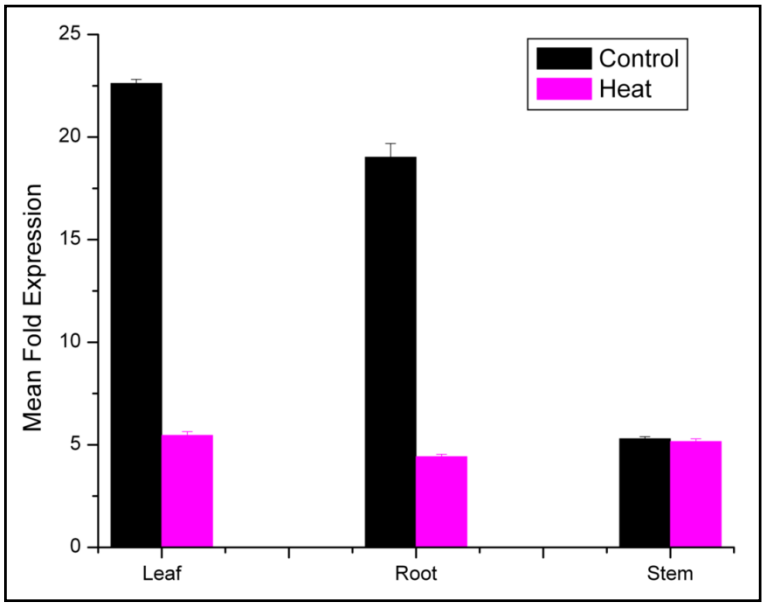


Fig. 4B.5: Relative expression in terms of mean fold expression of BMGT1 gene under heat shock treatment with respect to control plant. All values are plotted with standard deviation taken into account.

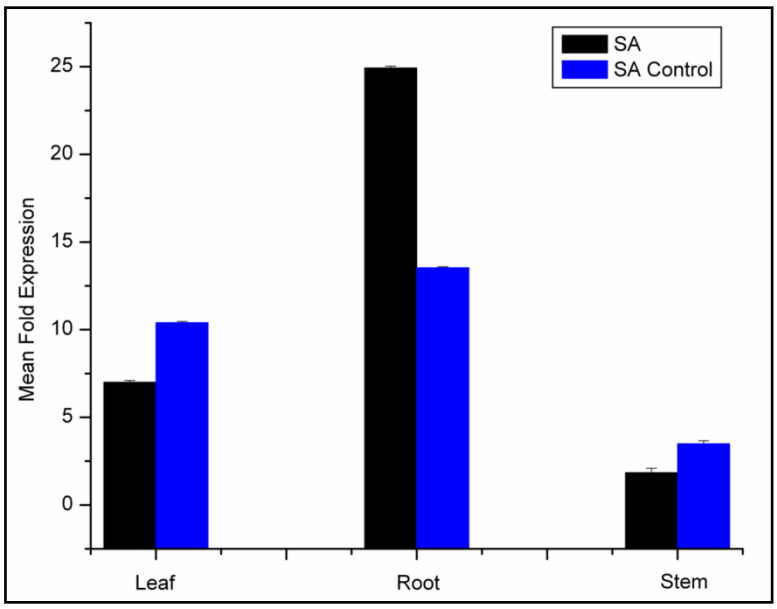


Fig. 4B.6: Relative expression in terms of mean fold expression of BMGT1 gene under salicylic acid (SA) stress with respect to salicylic acid control plant. All values are plotted with standard deviation taken into account.

Relative expression in terms of mean fold expression in control plant i.e. untreated plant using BMGT1 gene primers is highest in leaf and lowest in stem but under mannose stress conditions expression has increased almost 5-fold in root and huge dip in fold expression has been seen in leaf. In case of methyl jasmonate and salt stress expression has increased in stem and decreased in leaf. In cold and heat shock treatments expression has totally decreased. Finally in case of salicylic acid stress fold expression has increased almost 10 times in root and has decreased in stem and leaves. Numerical data of relative expression of BMGT1 gene is shown in Table 4B.2.

Table 4B.2 Relative expression of BMGT1 gene.

Sample name	Fold Expression	Sample name	Fold Expression
Control Leaf	22.62	Salt Leaf	9.06
Control Root	10.02	Salt Root	18.5
Control Stem	5.3	Salt Stem	7.78

Mannose Leaf	2	Heat Leaf	5.46
Mannose Root	24	Heat Root	4.43
Mannose Stem	2.94	Heat Stem	5.16
Methyl jasmonate Leaf	6.45	Salicylic acid Leaf	7.01
Methyl jasmonate Root	8.96	Salicylic acid Root	24.93
Methyl jasmonate Stem	9.25	Salicylic acid Stem	1.85
Cold Leaf	1	Salicylic acid Control Leaf	10.41
Cold Root	1.27	Salicylic acid Control Root	13.64
Cold Stem	1.18	Salicylic acid Control Stem	3.5

4.9.1b QRT-PCR results for BMGT2 gene

Real-time Q-PCR was performed using RBMGT2F and RBMGT2R as primers and cDNA isolated from untreated and treated plants. The relative expression of the BMGT2 gene in terms of mean fold expression is shown in Fig. 4B.7 (mannose stress with respect to control i.e. untreated plant), Fig. 4B.8 (methyl jasmonate stress with respect to control plant), Fig. 4B.9 (salt stress with respect to control plant), Fig. 4B.10 (cold shock with respect to control plant), Fig. 4B.11 (heat shock treatment with respect to control plant), Fig. 4B.12 (salicylic acid stress with respect to salicylic acid control plant).

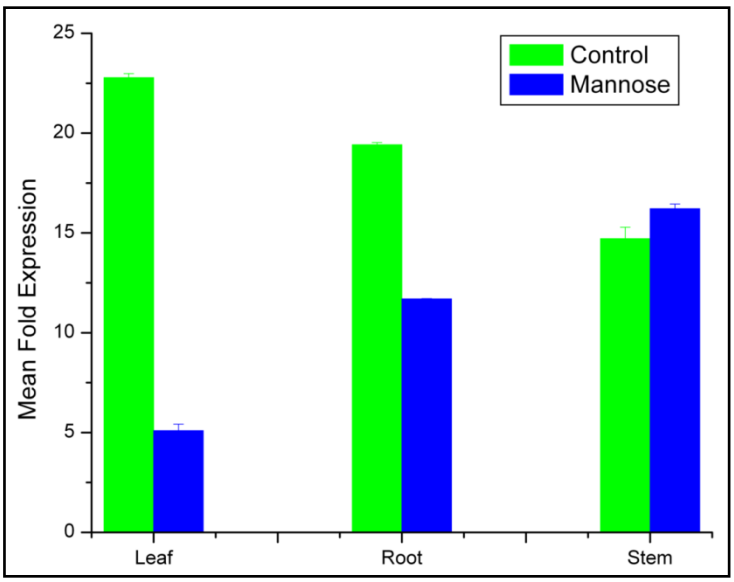


Fig. 4B.7: Relative expression in terms of mean fold expression of BMGT2 gene under mannose stress with respect to control plant.

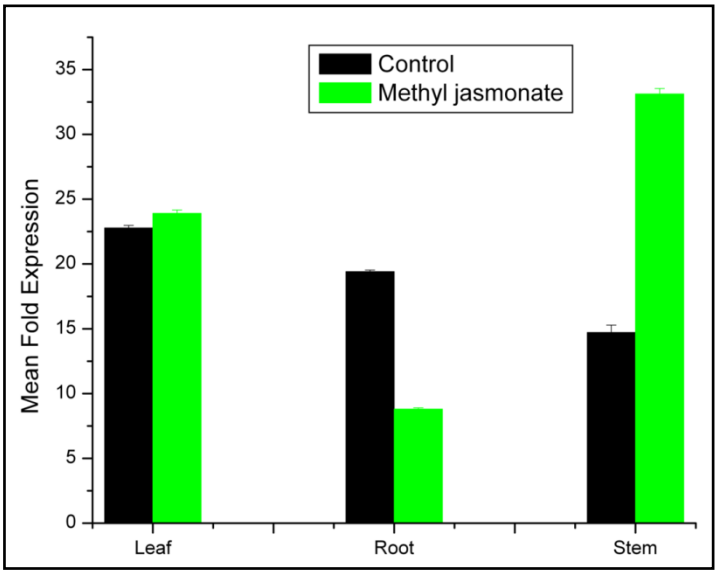


Fig. 4B.8: Relative expression in terms of mean fold expression of BMGT2 gene under methyl jasmonate stress with respect to control plant.

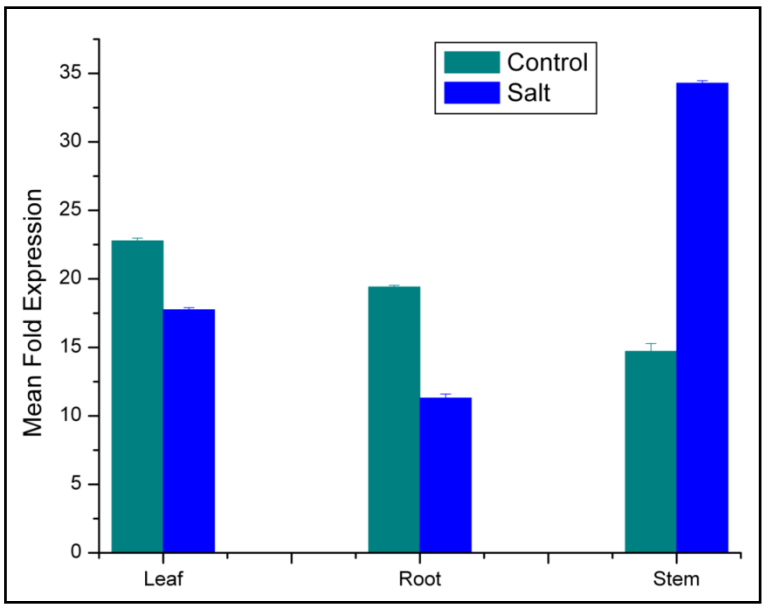


Fig. 4B.9: Relative expression in terms of mean fold expression of BMGT2 gene under salt stress with respect to control plant.

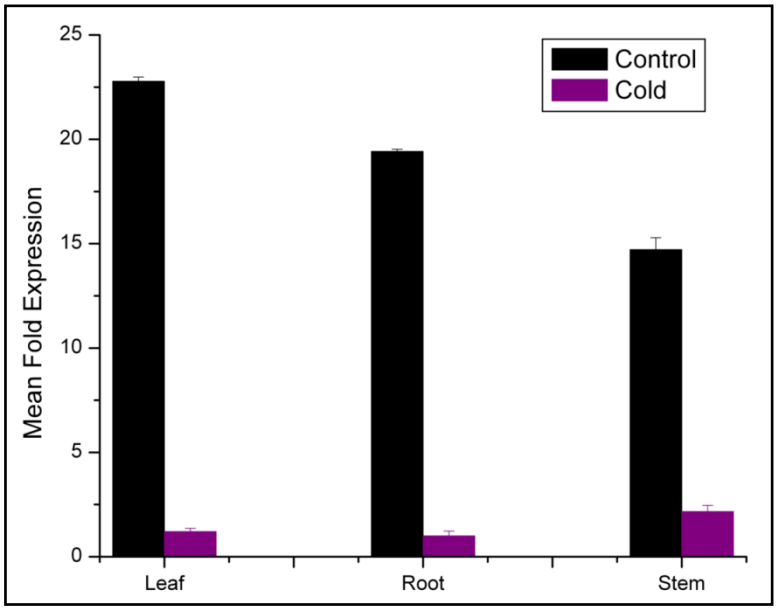


Fig. 4B.10: Relative expression in terms of mean fold expression of BMGT2 gene under cold stress with respect to control plant.

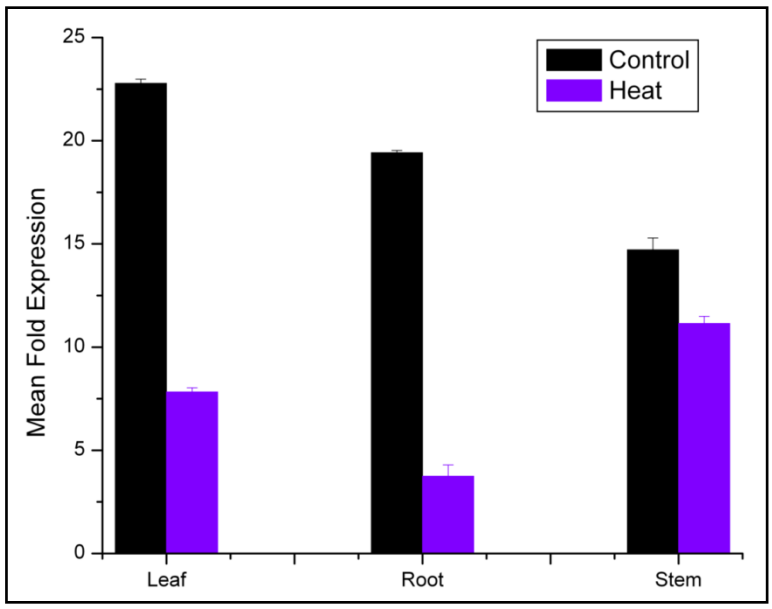


Fig. 4B.11: Relative expression in terms of mean fold expression of BMGT2 gene under heat stress with respect to control plant.

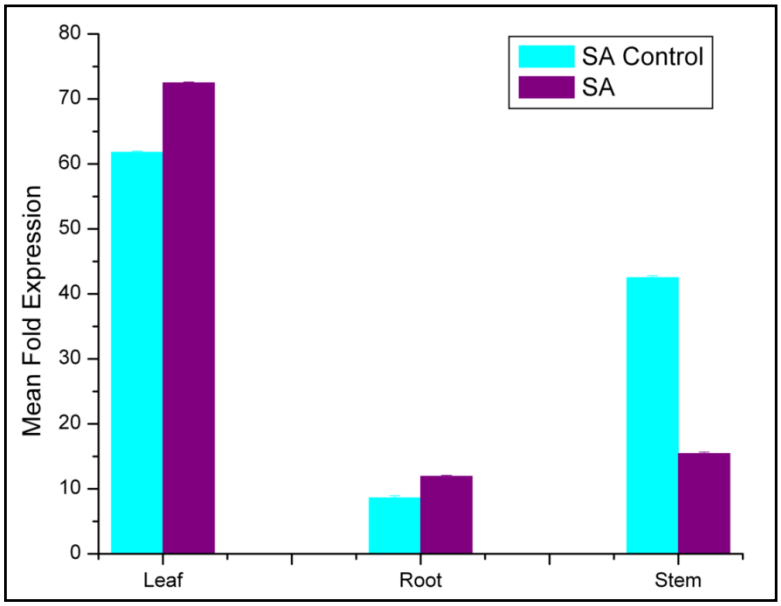


Fig. 4B.12: Relative expression in terms of mean fold expression of BMGT2 gene under salicylic acid stress with respect to salicylic acid control plant.

Relative expression in terms of mean fold expression in control plant i.e. untreated plant using BMGT2 gene primers is highest in leaf and lowest in stem but under mannose (2%) stress conditions expression has increased in stem and huge dip in fold expression has been seen in leaf. In case of methyl jasmonate (20 mM) and salt (50 mM) stress fold expression has increased almost 20 times in stem and shows almost 10 folds decrease in root. In cold (-20 °C) and heat (37 °C) shock treatments expression has decreased in all three i.e. leaf, root and stem. Finally in case of salicylic acid (30 mM) stress fold expression has increased almost 10 times in leaf and has decreased 25 folds in stem. Numerical data of relative expression of BMGT2 gene is shown in Table 4B.3.

Table 4B.3 Relative expression of BMGT2 gene.

Sample name	Fold Expression	Sample name	Fold Expression
Control Leaf	22.78	Salt Leaf	17.75
Control Root	19.42	Salt Root	11.31
Control Stem	14.72	Salt Stem	34.29
Mannose Leaf	5.1	Heat Leaf	7.83
Mannose Root	11.7	Heat Root	3.75
Mannose Stem	16.22	Heat Stem	11.15
Methyl jasmonate Leaf	23.91	Salicylic acid Leaf	72.50
Methyl jasmonate Root	8.81	Salicylic acid Root	11.95
Methyl jasmonate Stem	33.12	Salicylic acid Stem	15.45
Cold Leaf	1.21	Salicylic acid Control Leaf	61.81
Cold Root	1	Salicylic acid Control Root	8.63
Cold Stem	2.17	Salicylic acid Control Stem	42.51

4.9.2 Immuno-cytolocalization of GT in *Bacopa* plant

GT protein was immuno-cytolocalized in tissues of *B. monniera*. Stem, root and leaf tissues section of *Bacopa* plant were treated with antibody raised against BMGT1 protein. The deposition of blue-black to dark brownish colour after incubating with BCIP/NBT mix confirms the presence of GT protein near vascular bundle. Since both the isolated GTs from *B. monniera* are flavonoid glycosyltransferases and flavonoids being polyphenols present mostly in the vascular bundles which might be the possible reason why *B. monniera* GTs got localized in the vascular bundle region. In transverse sections of stem, leaf and root vascular bundles shows presence of GT protein which is clearly absent in pith tissue and cortex. Fig 4B.13 A, B, C and D shows the control transverse section of stem at 10x magnification, control stem section at 20x magnification, stem section treated with BMGT1 antibody at 10x magnification and stem section treated with BMGT1 antibody at 20x magnification respectively. Fig 4B.14 A & B Transverse section of leaf of control and leaf section treated with BMGT1 antibody respectively. Fig 4B.15 A, B, C and D transverse sections of root of control and root section treated with BMGT1 antibody at different magnifications respectively.

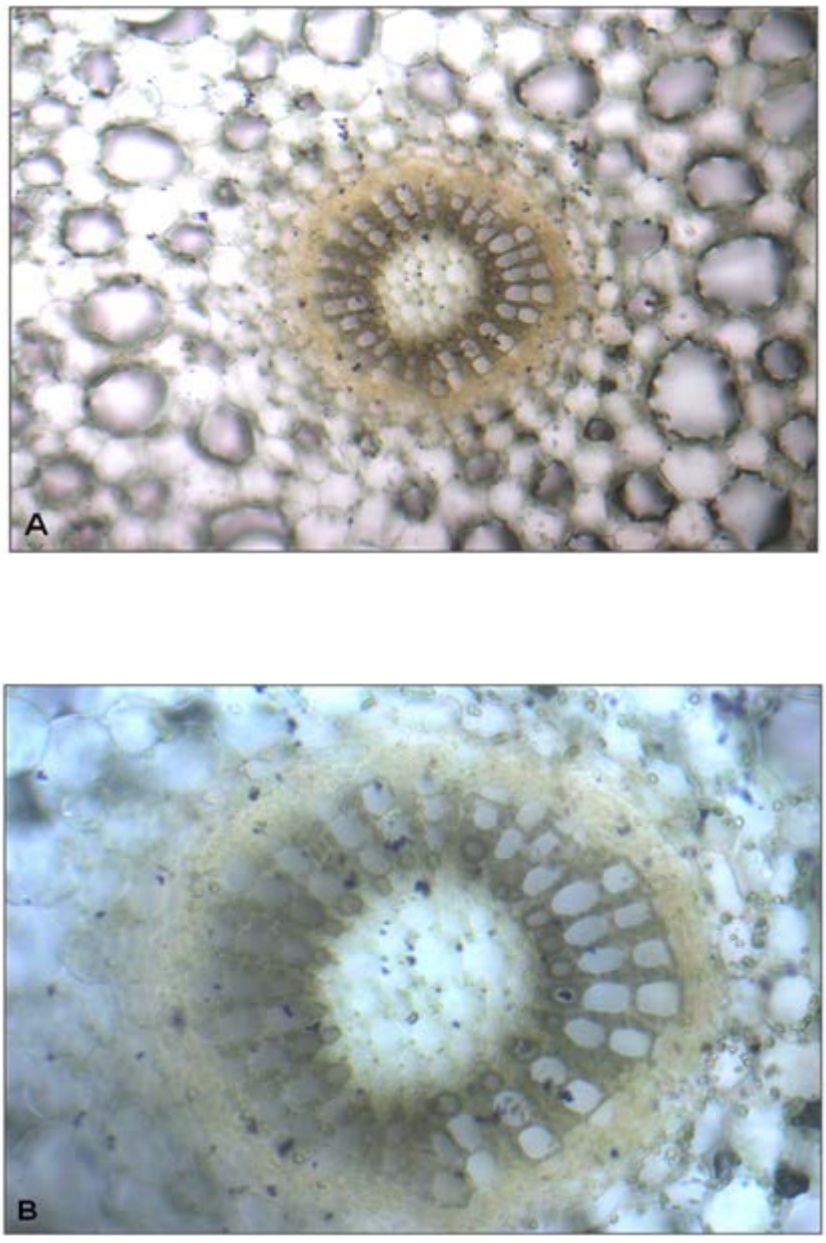


Fig. 4B.13: Immunocytolocalization of GT in *Bacopa monniera* stem as indicated by black arrow **(A)** Control transverse section of stem at 10x magnification, **(B)** control stem section at 20x magnification.

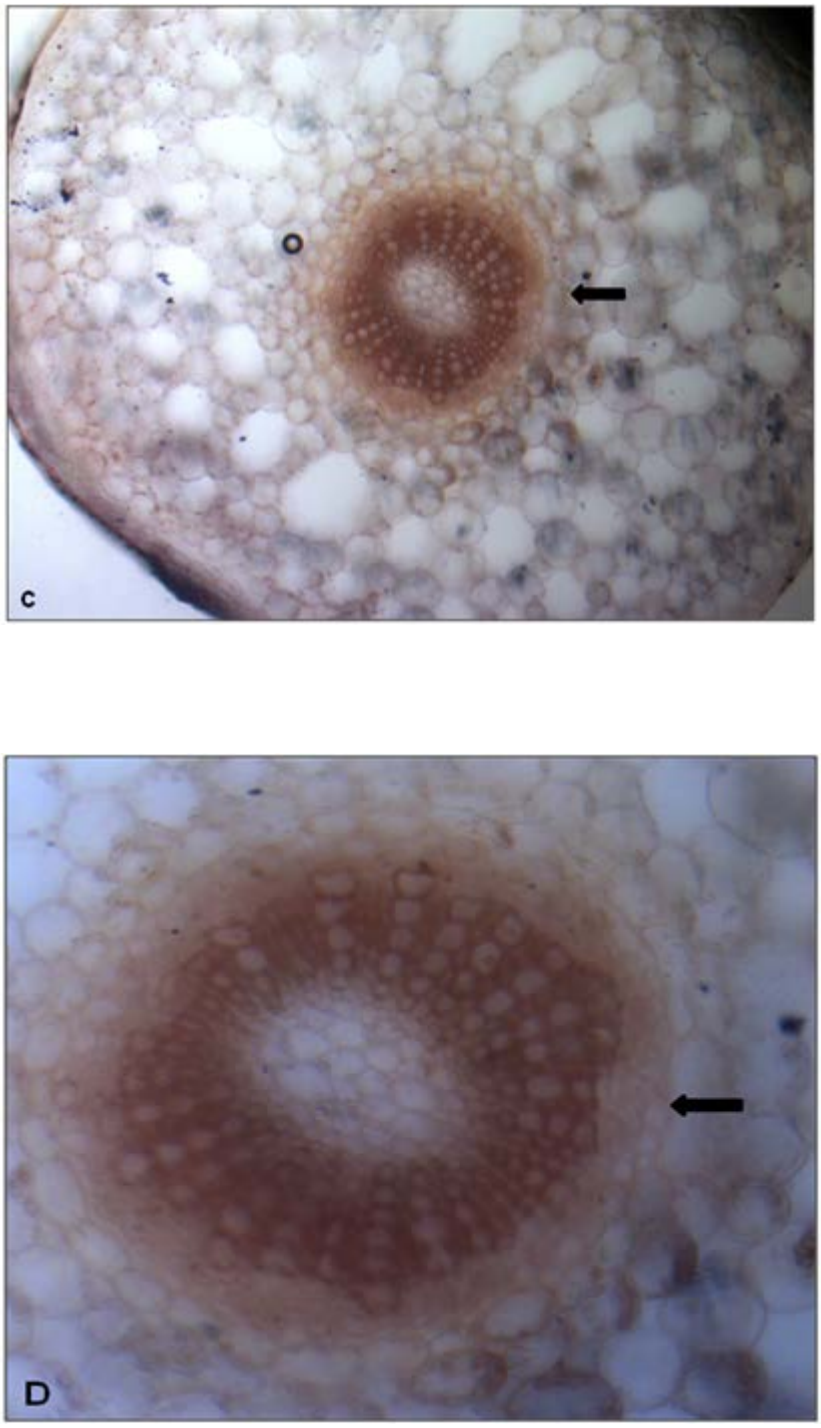


Fig. 4B.13: Immunocytolocalization of GT in *Bacopa monniera* stem as indicated by black arrow **(C)** Transverse section of stem treated with BMGT1 antibody at 10x magnification and (**D)** stem section treated with BMGT1 antibody at 20x magnification respectively.

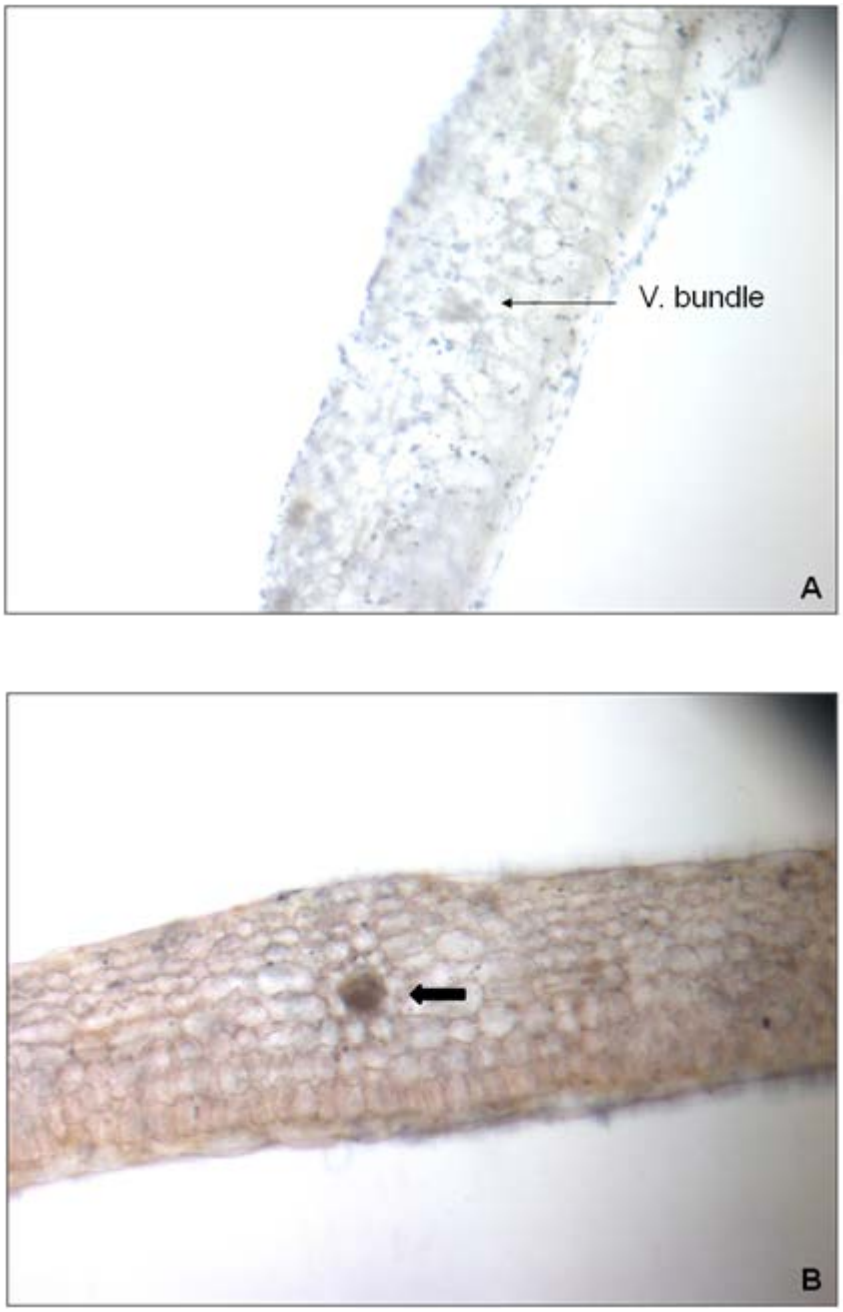


Fig. 4B.14: Immuno-cytolocalization of GT in *Bacopa monniera* leaf as indicated by black arrow **(A)** Control transverse section of leaf at 10x magnification, **(B)** Transverse section of leaf treated with BMGT1 antibody at 10x magnification respectively.

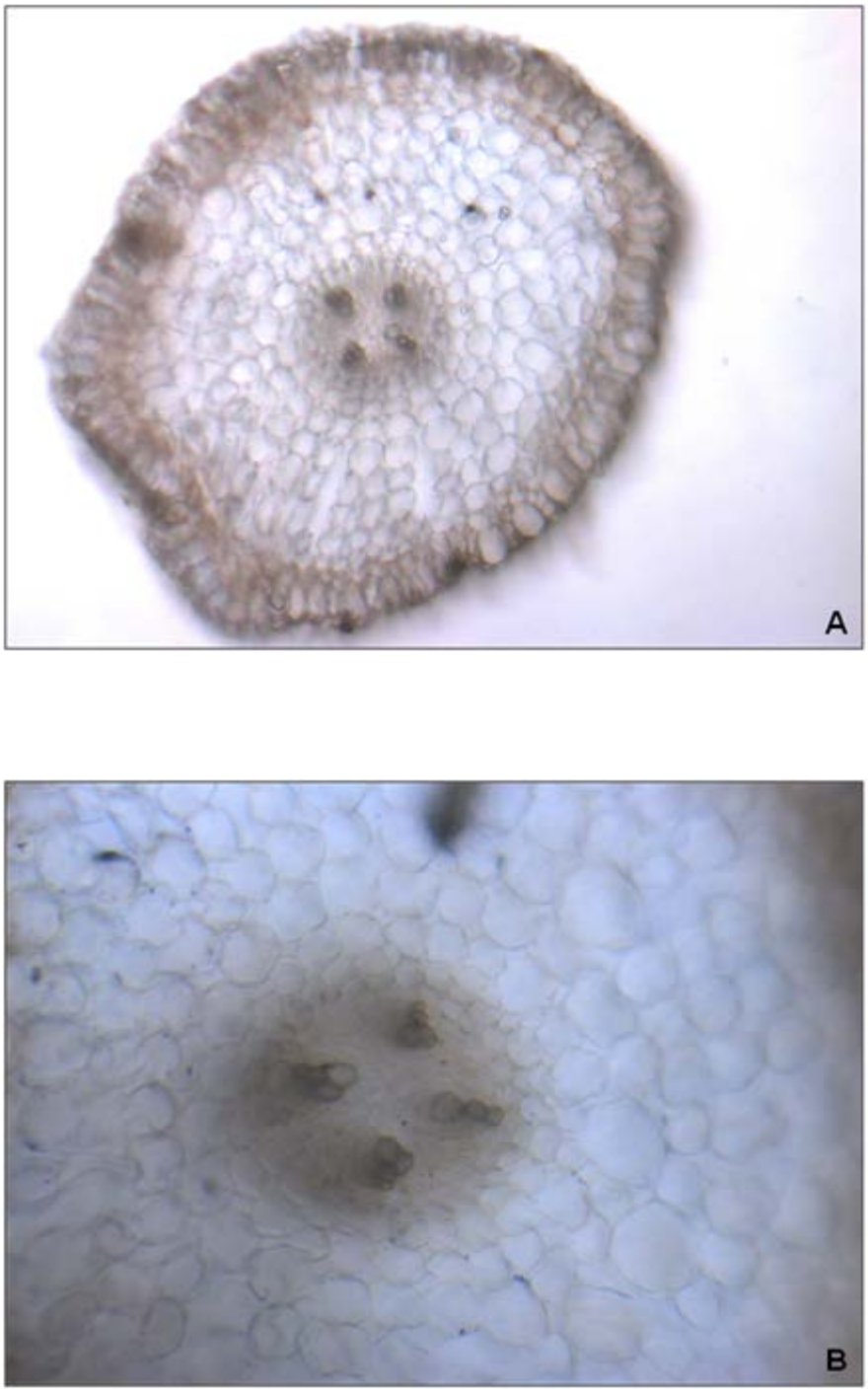


Fig. 4B.15: Immuno-cytolocalization of GT in *Bacopa monniera* root as indicated by black arrow **(A)** Control transverse section of root at 10x magnification, **(B)** control root section at 20x magnification.

Fig. 4B.15: Immuno-cytolocalization of GT in *Bacopa monniera* root as indicated by black arrow **(C)** Transverse section of root treated with BMGT1 antibody at 10x magnification and **(D)** Root section treated with BMGT1 antibody at 20x magnification respectively.

4.10 Conclusions

- Root, shoot and leaf tissue of *B. monniera* control plant as well as the plants which were grown under different stresses like mannose, salt, heat, cold, methyl jasmonate, salicylic acid etc. were harvested and used for real-time PCR analysis.

- Real-time PCR analysis showed differential expression at transcript level with maximum expression in leaf tissue in control plant for both BMGT1 AND BMGT2 genes.

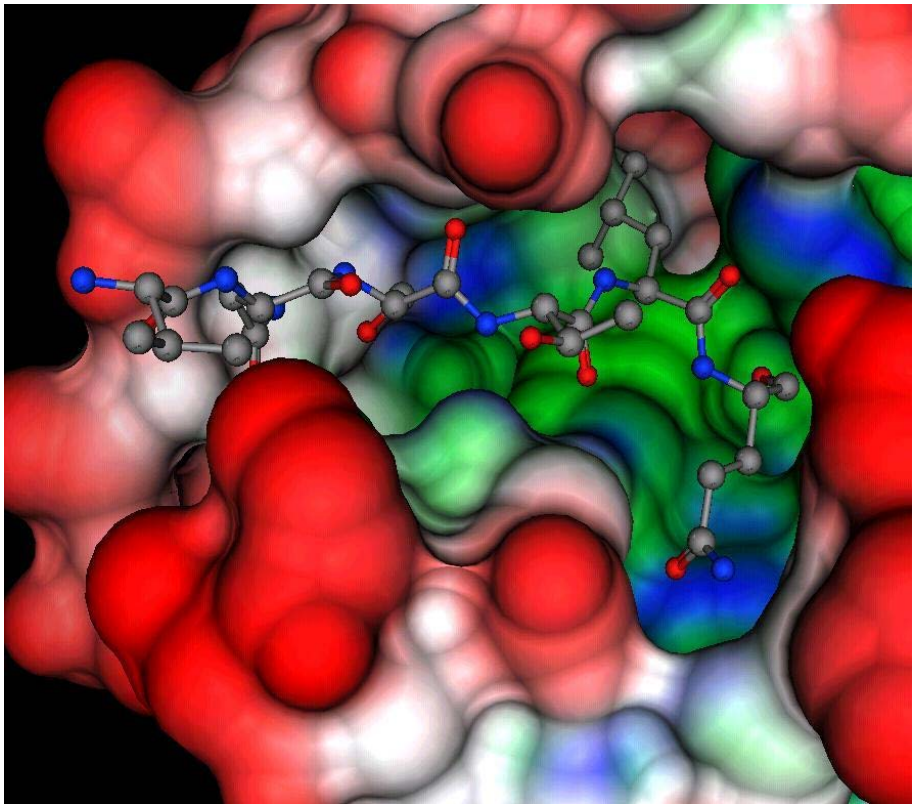
- In case of mannose stress, expression has increased 5 folds in roots whereas in case of salicylic acid stress expression has increased almost 20 folds in root for BMGT1 gene.

- In BMGT2 the relative expression has increased 20 folds in stem incase of methyl jasmonate and salt stress whereas 10 folds increase in expression in leaf was seen in case of salicylic acid stress.

- Transverse sections of immuno-cytolocalized root, leaf and stem showed that glycosyltransferases were localized in the vascular bundle.

Chapter: 5

Homology Modeling and Docking Studies of the B. monniera GT gene(s)



5.1 Introduction

Glycosyltransferases (GTs) catalyze the transfer of a sugar residue of an activated sugar donor to an acceptor molecule. The GTs constitute a large family of enzymes. Classification of the GTs into subfamilies has been performed according to the degree of primary sequence identity (Campbell *et al*., 1997; Coutinho *et al*., 2003). Members of these families can be found at the CAZY database (http://www.cazy.org/) comprises 91 different GT families. The primary structure of GTs is poorly conserved (Campbell *et al*., 1997; Hu and Walker, 2002). Despite low sequence conservation, the GTs crystallized so far show a highly conserved secondary and tertiary structure (Hu and Walker, 2002; Unligil and Rini, 2000; Zhang *et al*., 2003), and they all adopt distinct folds designated as GT-A and GT-B fold (Breton et al., 2006).

The 3D glycosyltransferase database (http://www.cermav.cnrs.fr/glyco3d/) holds crystal structures of 53 different GTs representing a total of 26 different CAZY families. Twenty six of the GTs with solved crystal structures adopt the GT-B fold and 10 of these belong to the family 1 GTs (listed in Fig. 5.1).

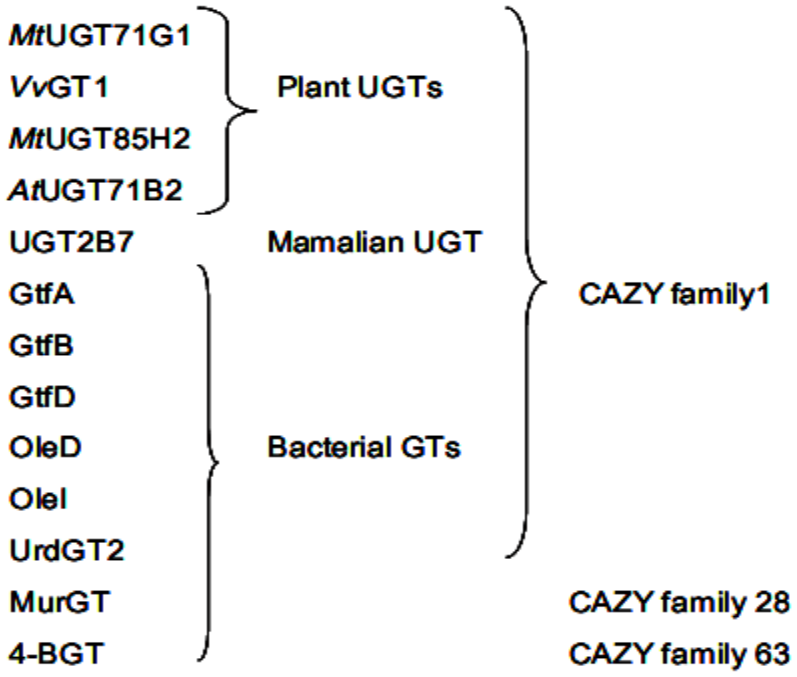


Fig. 5.1: List of GTs with available crystal structures.

The crystal structure of a plant UGT has only been resolved recently, the first being Mt UGT71G1 (Shao *et al*., 2005). Additional three plant UGTs have subsequently been crystallized: VvGT1 (Offen *et al*., 2006), MtUGT85H2 (Li *et al*., 2007) and AtUGT72B1 (Brazier-Hicks *et al*., 2007). The available plant UGT crystal structures are listed in Table 5.1.

Plant UGT	**Crystal name**	**In complex with**	**References**
*Vv*GT1	2CIX	UDP	Offen et al. (2006)
	2CIZ	UDP-2-fluoro-glucose, kaempferol	Offen et al. (2006)
*Mt*UGT71G1	2C9Z	UDP, quercetin	Offen et al. (2006)
	2ACV (molecules A & B)	UDP	Shao et al. (2005)
	2ACW (molecules A & B)	UDP-glucose	Shao et al. (2005)
*Mt*UGT85H2	2PQ6	None	Li et al. (2007)
*At*UGT72B1	2VCE	UDP-glucose	Brazier-Hick et al. (2007)
	2VCH	UDP-Tris buffer	Brazier-Hick et al. (2007)
	2VG8	UDP-2-fluoro-glucose, TCP	Brazier-Hick et al. (2007)

Table 5.1: List of available plant UGT crystal structures, the structures can be found at http:// www.rcsb.org/pdb/home/home.do.

The available crystal-based 3D structures of plant UGTs (Bolam et al., 2007; Li et al., 2007; Offen et al., 2006; Shao et al., 2005) have provided new insights with respect to inter- and intradomain interactions important for enzyme activity and specificity and have served to identify amino acid residues interacting with the sugar donor and acceptor. Unfortunately, only a limited number of UGT crystal structures have been solved and that is why comparative modeling constitutes an attractive alternative for the study of 3D structures. Comparative modeling in the form of homology modeling relies on the conservation of

secondary and tertiary structure between a query protein and a template protein for which the crystal-based 3D structure is available.

Based on available crystal-based 3D structure and computational methods such as secondary structure prediction, homology modeling and structural superimposition, the catalytic properties and identification of amino acids residues involved in active site could be determined. Docking studies were performed to predict its substrate specificity.

5.2 Materials and methods

5.2.1 Homology Modeling of *B. monniera* glycosyltransferases

The 3D models of *B. monniera* glycosyltransferases (BMGT1 and BMGT2) were built by homology modelling based on high-resolution crystal structures of homologous proteins. A basic local alignment search tool (BLAST) search for the sequence similarities with several members of the GT family was used for selecting the 3D models of the closest homologues available in the Brookhaven Protein Data Bank (PDB). The 3D models of BMGT1 and BMGT2 were generated by the automated homology modelling software MODELLER 9v8 (http://salilab.org) on windows operating environment. This program is used for comparative protein structure modelling that optimally satisfies spatial restraints which includes (i) homology-derived restraints on the distances and dihedral angles in the target sequence extracted from its alignment with the template structures (ii) stereochemical restraints such as bond length and bond angle preferences, obtained from the CHARMM-22 molecular mechanics force field (iii) statistical preferences for dihedral angles and non-bonded interatomic distances, obtained from a representative set of known protein structures and (iv) optional manually curated restraints, such as those from NMR spectroscopy, rules of secondary structure packing, cross-linking experiments, fluorescence spectroscopy, image reconstruction from electron microscopy, site-directed mutagenesis and intuition. The spatial restraints are expressed as probability density functions (pdfs) for the features restrained. The pdfs restrain C^{α}-C^{α} distances, main-chain N-O distances, main-chain and side-chain dihedral angles. The 3D model of the proteins were obtained by optimization of the molecular pdf such that the model violates the input restraints as little as possible. The molecular pdf was derived

as a combination of pdfs restraining individual spatial features of the whole molecule. The optimization procedure is a variable target function method that applies the conjugate gradients algorithm to positions of all non-hydrogen atoms. This model building procedure is similar to structure determined by NMR spectroscopy. The modelled structures were visualized using program WinCoot (http://www.chem.gla.ac.uk/~ bernhard/coot/wincoot.htm-l).

5.2.2 Checking the quality of the model with PROCHECK

The geometry of the final refined models of GT proteins was evaluated with Ramachandran's plot calculations computed with the PROCHECK program.

5.2.3 Secondary structure calculation and comparison between *Bacopa monniera* glycosyltransferases and template

Pair wise sequence alignment of modeled *B. monniera* GTs was done with template sequence using ClustalW. Secondary structural elements (α-helices, β-sheets and turns) of GTs were compared with the already known GTs.

5.2.4 Structural super imposition of *B. monniera* GTs with template

Pair wise structural superimposition of modeled *B. monniera* GTs was done with template using Coot program. Overall view of secondary structures and the segment containing residues forming the active site of *B. monniera* GTs with respective template was presented.

5.2.5 Molecular dynamics simulations

The structure with the least modeller objective function, obtained from the modeller was improved by molecular dynamics which was carried out using GROMACS-4.0.7 software. Simulations were run for 100 ps and 50,000 steps at 300 K using OPLS (Optimized Potentials for Liquid Simulations) force field.

5.2.6 Docking of different substrates into *B. monniera* glycosyltransferases

Molecular docking was performed using Autodock software (http://autodock.scripps.edu/). Structure Data File (SDF) of ligand molecule was downloaded from the pubchem (http://pubchem.ncbi.nlm.nih.gov). SDF format was converted to MOL2 format using online server http://openbabel.org/wiki/Main_Pagehttp. Hydrogens were added to receptor as well as to ligand using the built-in program Add Hydrogen in Autodock Software. Acceptor and donor binding sites were already known so accordingly grid was made near to these regions and docking was performed.

5.3 Results and discussions

5.3.1 Bioinformatic analysis of *B. monniera* glycosyltransferases (BMGT1 and BMGT2)

5.3.1.1 Homology Modeling of *B. monniera* BMGT1 protein

In order to find out the homologous sequences in Protein Data Bank, the primary sequence of *B. monniera* BMGT1 was searched against PDB using BLASTP program at NCBI server (http://www.ncbi.nlm.nih.gov/blast). Among all the homologs, glycosyltransferase from *Vitis vinifera* (PDB ID: 2c1z) was closest to BMGT1, with 32% identity at the amino acid sequence level. Glycosyltransferases from *Medicago truncaluta* and *Arabidopsis thaliana* (PDB ID: 3hbf, 2pq6, 2acv and 2vch) showed 25%, 26%, 24% and 26% identity respectively. The three dimensional coordinates of *V. vinifera* (2c1z) and *M. truncaluta* (3hbf) were used as templates to generate the 3D model of the BMGT1 using the program Modeller 9v8 (http://salilab.org). Out of 463 residues submitted for homology model, 458 residues were modeled in the structure. 5 N-terminal residues remained unmodeled because they are not having regular secondary structures and coming in loop region. Of the 20 structures calculated for the same target (BMGT1) and the template (2c1z and 3hbf), the one with the lowest value of the MODELLER objective function was selected as the best model for BMGT1. The structures with least RMSD value were used for further analysis. The RMSD value of BMGT1 with templates *i.e.* 2c1z and 3hbf was found to be 0.63 Å and 1Å respectively. The

RMSD value less than 1.00 showed that although the identity of the template chosen was less *i.e.* 32% and 25%, the model built was good in all the domains modeled in the predicted structure. The final stable structure of BMGT1 is shown in Fig. 5.2.

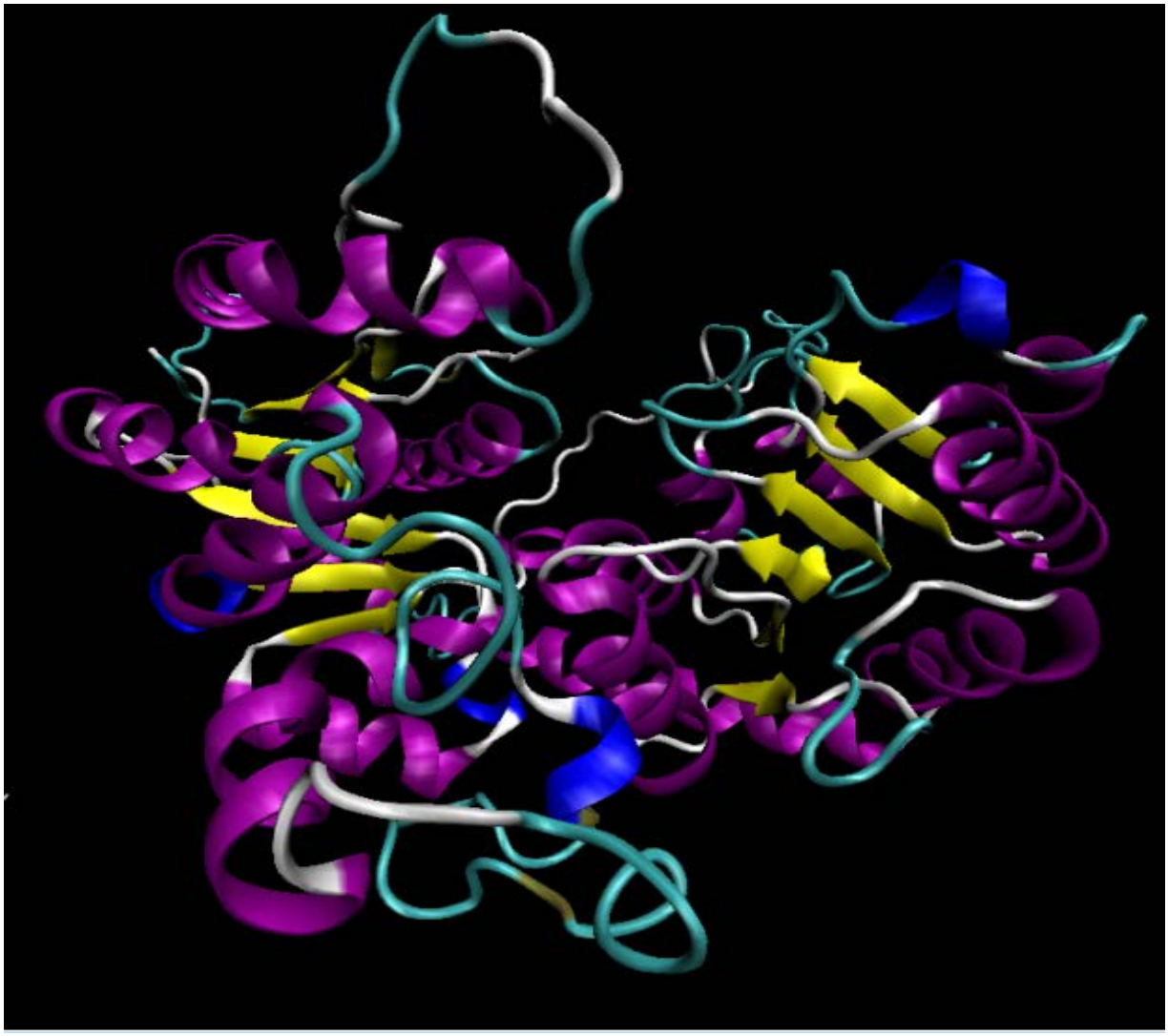


Fig. 5.2: Final 3D structure of BMGT1 enzyme. α-helices are represented in purple and beta sheets in yellow coloured ribbons.

5.3.1.1.1 Checking the quality of the model with PROCHECK

The quality of the model was examined using the program PROCHECK (Laskowski *et al*., 1993). The PROCHECK analysis showed 88%, 10.1%, 1.0% and 1.0% residues in the most favorable, additionally allowed, generously allowed and the disallowed regions of the Ramachandran map respectively (Fig. 5.3 & Table 5.2).

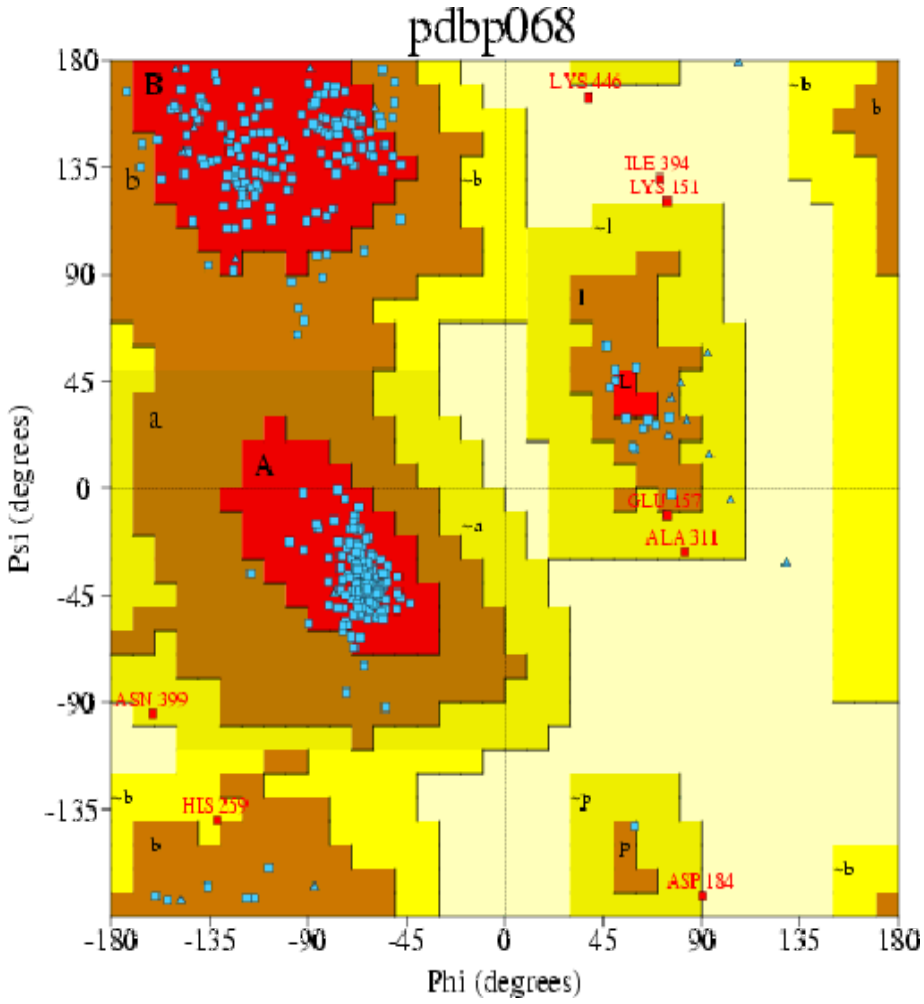


Fig. 5.3: Ramachandran(Φ,ψ) plot for modeled *B. monniera* BMGT1.

Table 5.2: % of residues falling in the core region of the Ramachandran's plot.

	No. of residues	%age
Most favoured regions	367	88.0%*
Additional allowed regions	42	10.1%
Generously allowed regions	4	1.0%
Disallowed regions	4	1.0%*
Non-glycine and non-proline residues	**417**	**100.0%**
End-residues (excl. Gly and Pro)	2	
Glycine residues	24	
Proline residues	20	
Total number of residues	**463**	

5.3.1.1.2 Molecular dynamics simulations

The structure of BMGT1 with the least modeller objective function, obtained from the modeller was improved by molecular dynamics as mentioned in section 5.2.5.

5.3.1.1.3 Secondary structure calculation and comparison between target (*B. monniera* BMGT1) and templates (2c1z and 3hbf)

In order to compare secondary structures of BMGT1 with 2c1z and 3hbf, web based program SPDBV (Guex, N. and Peitsch, M.C. 1997) was used to generate comparative picture of secondary structural elements (α-helices, β-sheets and turns) of BMGT1 with the already known 2c1z and 3hbf. The comparative secondary structural elements of BMGT1 with that of 2c1z and 3hbf are shown in Fig. 5.4.

```
2c1z              6      TNPHVAVLAFPFSTHAAPLLAVVRRLAAAAPHAVFSFFSTSQSNASIFHD-HTMQCNI
3hbf              12     NLLHVAVLAFPFGTHAAPLLSLVKKIATEAPKVTFSFFCTTTTNDTLFSRSNEFLPNI
BMGT1_model2      6    TGKEAHILVFPYPAQGHINPVLPFSKFLASKGLKVTIIVTPSVKKLVNFPP-----NSSI
2c1z                     ..sssssss.......hhhhhhhhhhhhh....ssssss.hhhhhhh...-.....ss
```

```
3hbf                  ..sssssss......hhhhhhhhhhhhhh....ssssss.hhhhhh..........ss
BMGT1_model2          ....ssssss.......hhhhhhhhhhhhhh....ssssss.hhhhh.....------..ss

2c1z            65    KSYDISDGVPEGYVFAGRP-QEDIELFTRAAPESFRQGMVMAVAETGRPVSCLVADAFIW
3hbf            70    KYYNVHDGLPKGY-VSSGNPREPIFLFIKAMQENFKHVIDEAVAETGKNITCLVTDAFFW
BMGT1_model2    61    SIERISDGSE-DVKETEDI-EAYFNRFRREASQNLAKFIDEKKG---WGAKVIVYDSTMP
2c1z                  sssss.............-.hhhhhhhhhhhhhhhhhhhhhhhhh.....ssssss...
3hbf                  sssss........-......hhhhhhhh...hhhhhhhhhhhhhhh....sssssss...
BMGT1_model2          sssss.....-........-.hhhhhhhhhhhhhhhhhhhhh..---....ssssss...

2c1z            124   FAADMAAEMGVAWLPFWTAGPNSLSTHVYIDEIREKIGVSGIQGREDELLNFIPGMSKVR
3hbf            129   FGADLAEEMHAKWVPLWTAGPHSLLTHVYTDLIREKTGSKEVHDVKSID---VLPGFPEL
BMGT1_model2    116   WVLDIAHERGLLGASFFTQS--CFVSAVYCHLHQGTLKYPYEEEEKSTLLSLHPLLPTLQ
2c1z                  hhhhhhhhh..sssss....hhhhhhh...hhhhhhh..........sss........ss
3hbf                  hhhhhhhhh..ssssss...hhhhhhh...hhhhh....hhh.......---........
BMGT1_model2          hhhhhhhhh..sssss.....--hhhhh...hhhhhhh..........sss........ss

2c1z            184   F-RDLQEGIVFGNLNSLFSRMLHRMGQVLPKATAVFINSFEELDDSLTNDLKSKLKTYLN
3hbf            186   KASDLPEGVIK-DIDVPFATMLHKMGLELPRANAVAINSFATIHPLIENELNSKFKLLLN
BMGT1_model2    174   I-NDLPCFSKFDDPKHLVSKHLTDQFINLDKVDWILFNTFYDLETQVAEWMKAKW-PIKT
2c1z                  s-.............hhhhhhhhhhhhhh....sss.......hhhhhhhhhh..ssss
3hbf                  ...........-....hhhhhhhhhhhhhh....ssss......hhhhhhhhh...ssss
BMGT1_model2          s-.............hhhhhhhhhhhhhh....sss.......hhhhhhhhh..-ssss

2c1z            243   IGPFNLIT--------------------GCLQ-WLKERKPTSVVYISFGTVTTPPPAEV
3hbf            245   VGPFNLTTPQRK-----------VSDEHGCLE-WLDQHENSSVVYISFGSVVTPPPHEL
BMGT1_model2    232   IGPTSLLEKHKKLGNDKNQIISLFEQNHKACIDQWLDSMETCSVVYVSLGSIASIGKEEM
2c1z                  s...hh..---------------------.hhh-hhh.......ssss........hhhh
3hbf                  s..........-----------......hhh-hhh.......ssss........hhhh
BMGT1_model2          s..hhhh.......................hhhhhh......sssss........hhhh

2c1z            290   VALSEALEASRVPFIWSLRDKARVHLPEGFLEKTRGYGMVVPWAPQAEVLAHEAVGAFVT
3hbf            292   TALAESLEECGFPFIWSFRGDPKEKLPKGFLERTKTKGKIVAWAPQVEILKHSSVGVFLT
BMGT1_model2    292   EELACGLLMSNCYFLWVVRASEQDKLPSDFMSLASEKGLIVNWCCQTEVLAHPAVACFMT
2c1z                  hhhhhhhhhh..ssssss..hhhh.....hhh....sssss....hhhhhh.....ssss
3hbf                  hhhhhhh.....ssssss...hhhh....hhhh...sssss....hhhhhh.....ssss
BMGT1_model2          hhhhhhhhhh..ssssss...hhh.....hhhhhhh.sss....hhhhhh.....ssss

2c1z            350   HCGWNSLWESVAGGVPLICRPFFGDQRLNGRMVEDVLEIGVRIEG---GVFTKSGLMSCF
3hbf            352   HSGWNSVLECIVGGVPMISRPFFGDQGLNTILTESVLEIGVGVDN---GVLTKESIKKAL
BMGT1_model2    352   HCGWNSTLEAISCGVPLVTMAQWVDQQPNAKCVEDLWKVGVRIKGPENGTFEREEIARCI
2c1z                  ...hhhhhhhhhh..sssss.....hhhhhhhhhhh..sss....---....hhhhhhhh
3hbf                  ...hhhhhhhhhh..sssss.....hhhhhhhhhhh..sssss..---....hhhhhhhh
BMGT1_model2          ...hhhhhhhhhh..sssss.....hhhhhhhhhhh..............hhhhhhhh

2c1z            407   DQILSQEKGKKLRENLRALRETADRAVGPKGSSTE---NFITLVDLVSKPKD--
3hbf            409   ELTMSSEKGGIMRQKIVKLKESAFKAVEQNGTSAM---DFTTLIQIVTS-
BMGT1_model2    412   QQVIGGDNADELRANAWKWKKLAQDAMEENGNSTKNIEDFVVQFFNMSLLLT-
```

```
2c1z                hhhh...hhhhhhhhhhhhhhhhhhhh....hhhh---hhhhhhhhh......--
3hbf                hhhh...hhhhhhhhhhhhhhhhhhhh....hhhh---hhhhhhhhh..-
BMGT1_model2        hhhh...hhhhhhhhhhhhhhhhhhhh............hhhhhhhhh....-
```

Fig 5.4: Pair wise sequence alignment of BMGT1 with 2c1z and 3hbf. The secondary structures of BMGT1 and templates (2c1z & 3hbf) are shown above. Helices and strands are represented by h and s respectively. The Figure was generated using web based program SPDBV.

As shown in Fig. 5.4 it is quite evident that the contents of secondary structural elements in BMGT1 are more or less similar to 2c1z and 3hbf. The aligned protein of BMGT1 with the templates contains 20 α-helices and 15 β-sheets respectively. These secondary structures are highly conserved with the template inspite of several amino acid differences.

5.3.1.1.4 Structural superimposition of *B. monniera* BMGT1 with 2c1z and 3hbf

Pair wise structural alignment of modeled *B. monniera* BMGT1 was done with templates 2c1z and 3hbf using combinatorial extension algorithm at WinCoot. Cartoon view of the superposed structures is shown in the Fig. 5.5.

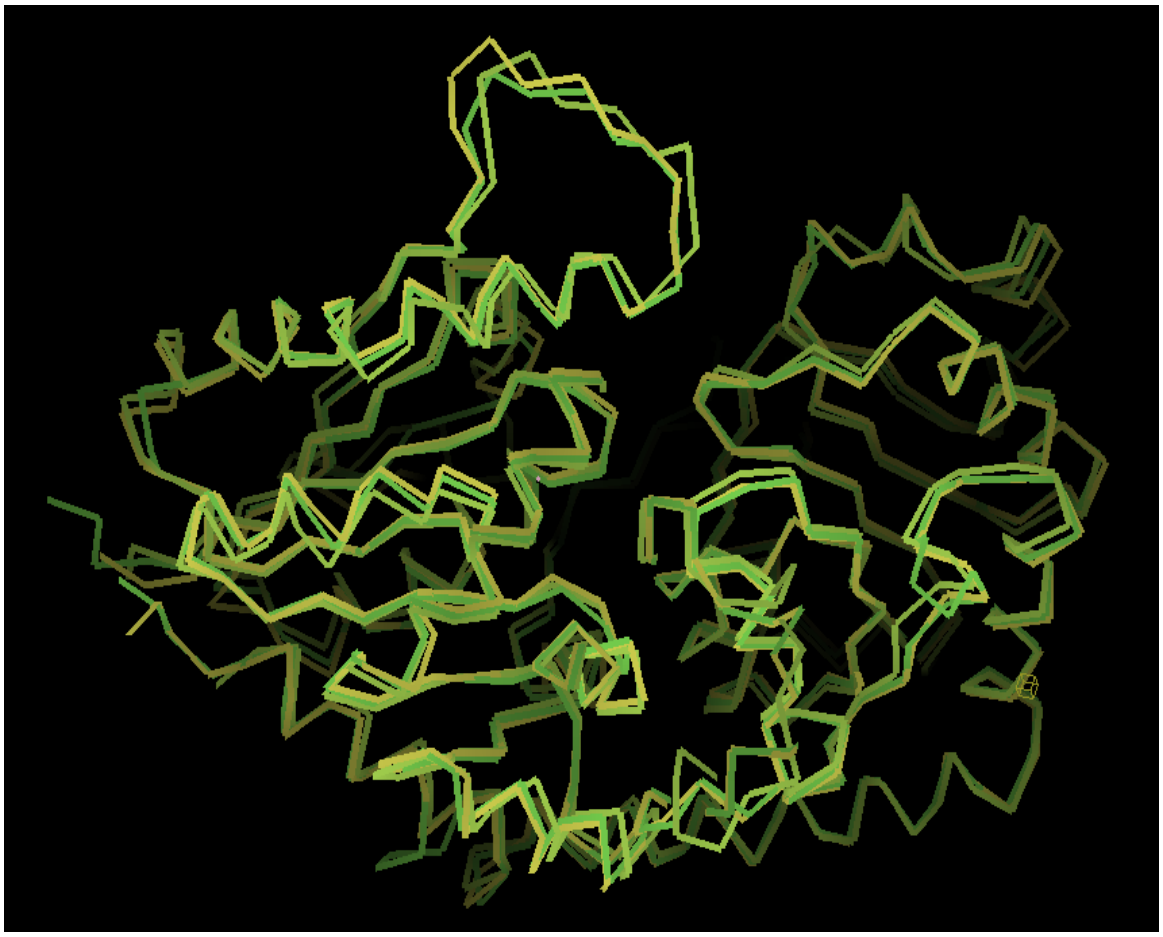


Fig. 5.5: Cartoon view of superimposition of modelled BMGT1 (Dark green) with 2c1z (yellow) and 3hbf (light green).

The RMSD of Cα trace of superimposed BMGT1 with templates *i.e.* 2c1z and 3hbf was found to be 0.63 Å and 1 Å respectively. This refined model was then used further for docking studies.

5.3.1.1.5 Analysis of amino acids involved in acceptor binding and donor binding sites in *B. monniera* glycosyltransferase (BMGT1) with other glycosyltransferases

SPDBV software suite (http://www.expasy.org/spdbv) was used for the alignment of multiple protein structures. 7 structures of glycosyltransferases were aligned with that of modeled BMGT1 to compare the important residues involved in acceptor and donor binding. The

conserved histidine and aspartic acid residues in the acceptor binding cleft and other conserved residues involved in donor binding pocket are shown in Fig. 5.6.

```
             ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                  5         15         25         35         45         55
BMGT2        -------MED ------AIVL YSSAEHLNSM LVPAKFISKH HPSISVIIIS TAAESAAAS-
3HBF         MSTFKNEMNG NNLLHVAVLA FPFGTHAAPL LSLVKKIATE APKVTFSFFC TTTTNDTLFS
2pq6         -----MGNFA NRKPHVVMIP YPVQGHINPL FKLAKLLHLR G--FHITFVN TEYNHKRLLK
2acv         ----MSMSDI NKNSELIFIP APGIGHLASA LEFAKLLTNH DKNLYITVFC IKFPGMPFAD
2C1Z         MSQTTTNP-- ----HVAVLA FPFSTHAAPL LAVVRRLAAA APHAVFSFFS TSQSNASIFH
2vch         -------MEE SKTPHVAIIP SPGMGHLIPL VEFAKRLVHL HGLTVTFVIA GEGPPSKAQR
BMGT1        ----MESKGT GKEAHILVFP YPAQGHINPV LPFSKFLASK GLKVTIIVTP SVKKLVNFPP
Clustal Co                    ..   .   *  .  .   : :          .

             ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                  65         75         85         95        105        115
BMGT2        --VASVP--- -SITYHRLPS APL----PPD LPTSIIELFF EIPRFHN-PF LHEALLEISQ
3HBF         RSNEFLP--- -NIKYYNVHD GLP----KGY VSSGNPREPI FLFIKAMQEN FKHVIDEAVA
2pq6         SRGPKAFDGF TDFNFESIPD GLTPMEGDGD VSQDVPTLCQ SVRKNFLKPY CELLTRLNHS
2acv         SYIKSVLASQ PQIQLIDLPE VEP----PPQ ELLKSPEFYI LTFLESLIPH VKATIKTILS
2C1Z         DSMHTMQC-- -NIKSYDISD GVP----EGY VFAGRPQEDI ELFTRAAPES FRQGMVMAVA
2vch         TVLDSLPS-- -SISSVFLPP VDL----TDL SSSTRIESRI SLTVTRSNPE LRKVFDSFVE
BMGT1        N--------- SSISIERISD GSE------- -DVKETEDIE AYFNRFRREA SQNLAKFIDE
Clustal Co              .:    :                                   .

             ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                 125        135        145        155        165        175
BMGT2        KSNLR--AFL IDFFCNSTFE VSTSLNIPTY FYLSGGACGL CALLYFPTID EAVSPRDIGE
3HBF         ETGKNITCLV TDAFFWFGAD LAEEMHAKWV PLWTAGPHSL LTHVYTDLIR EKTGSKEVHD
2pq6         TNVPPVTCLV SDCCMSFTIQ AAEEFELPNV LYFSSSACSL LNVMHFRSFV ERGIIPFKDE
2acv         NKVVG---LV LDFFCVSMID VGNEFGIPSY LFLTSNVGFL SLMLSLKNRQ IEEVFDDSDR
2C1Z         ETGRPVSCLV ADAFIWFAAD MAAEMGVAWL PFWTAGPNSL STHVYIDEIR EKIGVSGIQG
2vch         GGRLPT-ALV VDLFGTDAFD VAVEFHVPPY IFYPTTANVL SFFLHLPKLD ETVS-CEFRE
BMGT1        KKGWGAKVIV YDSTMPWVLD IAHERGLLGA SFFTQSCFVS AVYCHLHQGT LKYPYEEEEK
Clustal Co            ::  *       :  . .          .

             ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                 185        195        205        215        225        235
BMGT2        LND------- -FLEIPGCPP VHSLDFPKAM WFRRSNT--Y KHFLDTAGNM RRASGIVTNS
3HBF         VKS------- -IDVLPGFPE LKASDLPEGV IK-DIDVPFA TMLHKMGLEL PRANAVAINS
2pq6         SYLTNGCLET KVDWIPGLKN FRLKDIVDFI RTTNPNDIML EFFIEVADRV NKDTTILLNT
2acv         DHQ------- -LLNIPGISN QVPSNVLPDA CFNKDGG--Y IAYYKLAERF RDTKGIIVNT
2C1Z         REDEL----- -LNFIPGMSK VRFRDLQEGI VFGNLNSLFS RMLHRMGQVL PKATAVFINS
2vch         LTE------- -PLMLPGCVP VAGKDFLDPA QDRKDDA--Y KWLLHNTKRY KEAEGILVNT
BMGT1        STL------- -LSLHPLLPT LQINDLPCFS KFDDPKHLVS KHLTDQFINL DKVDWILFNT
Clustal Co                    *        :.                              :  *:

             ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                 245        255        265        275        285        295
BMGT2        FDAIEFRAKE ALSNSLCTPG LATPPVYVIG PLVAETNR-- ---------- KNGGEEHECL
3HBF         FATIHPLIEN ELN------- SKFKLLLNVG PFNLTTPQR- ---------- -KVSDEHGCL
2pq6         FNELESDVIN ALS------- STIPSIYPIG PLPSLLKQTP QIHQLDSLDS NLWKEDTECL
2acv         FSDLEQSSID ALY----DHD EKIPPIYAVG PLLDLKGQPN PK-------- LDQAQHDLIL
2C1Z         FEELDDSLTN DLK------- SKLKTYLNIG PFNLITPPP- ---------- -VVPNTTGCL
2vch         FFELEPNAIK ALQ----EPG LDKPPVYPVG PLVNIGKQE- ---------- AKQTEESECL
BMGT1        FYDLETQVAE WMK------- -AKWPIKTIG PTSLLEKHKK LGNDKNQIIS LFEQNHKACI
Clustal Co   *  :.    .  :               :* *                        :    :
```

```
             ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                305        315        325        335        345        355
BMGT2        K-WLDSQPIK SVIFLCFGRR GL-FSAAQLK EMAIGLENSG HRFLWSVRSP PGPAAAK---
3HBF         E-WLDQHENS SVVYISFGSV VT-PPPHELT ALAESLEECG FPFIWSFR-- ----------
2pq6         D-WLESKEPG SVVYVNFGST TV-MTPEQLL EFAWGLANCK KSFLWIIRP- ----------
2acv         K-WLDEQPDK SVVFLCFGSM GVSFGPSQIR EIALGLKHSG VRFLWSN--- ----------
2C1Z         Q-WLKERKPT SVVYISFGTV TT-PPPAEVV ALSEALEASR VPFIWSLR-- ----------
2vch         K-WLDNQPLG SVLYVSFGSG GT-LTCEQLN ELALGLADSE QRFLWVIRSP SGIANSSYFD
BMGT1        DQWLDSMETC SVVYVSLGSI AS-IGKEEME ELACGLLMSN CYFLWVVR-- ----------
Clustal Co   . **..     **::: :*          ::   :: .*  .    *:*

             ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                365        375        385        395        405        415
BMGT2        ---DPDLDAL LPEGFMERTK DRG-FVIKTW APQKEVLSHE AVGGFVTHCG RSSVLEAVSF
3HBF         ----GDPKEK LPKGFLERTK TKG--KIVAW APQVEILKHS SVGVFLTHSG WNSVLECIVG
2pq6         -DLVIGGSVI FSSEFTNEIA DRG--LIASW CPQDKVLNHP SIGGFLTHCG WNSTTESICA
2acv         ----SAEKKV FPEGFLEWME LEGKGMICGW APQVEVLAHK AIGGFVSHCG WNSILESMWF
2C1Z         ----DKARVH LPEGFLEKTR GYG--MVVPW APQAEVLAHE AVGAFVTHCG WNSLWESVAG
2vch         SHSQTDPLTF LPPGFLERTK KRG-FVIPFW APQAQVLAHP STGGFLTHCG WNSTLESVVS
BMGT1        ----ASEQDK LPSDFMSLAS EKG--LIVNW CCQTEVLAHP AVACFMTHCG WNSTLEAISC
Clustal Co             :.  * .      *   :  * . * ::* *  : . *::*.*  .*  *.:

             ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                425        435        445        455        465        475
BMGT2        GVPMIGWPMY AEQRMQRVFM VEEMKVALPL AEEA---DGF VTAGELEKRV RELMGLPAGK
3HBF         GVPMISRPFF GDQGLNTILT ESVLEIGVG- -VDN----GV LTKESIKKAL ELTMSSEKGG
2pq6         GVPMLCWPFF ADQPTDCRFI CNEWEIGMEI DTNVK----- --REELAKLI NEVIAGDKGK
2acv         GVPILTWPIY AEQQLNAFRL VKEWGVGLGL RVDYRKGSDV VAAEEIEKGL KDLMD--KDS
2C1Z         GVPLICRPFF GDQRLNGRMV EDVLEIGVR- -IEG----GV FTKSGLMSCF DQILSQEKGK
2vch         GIPLIAWPLY AEQKMNAVLL SEDIRAALRP RAGD---DGL VRREEVARVV KGLMEGEEGK
BMGT1        GVPLVTMAQW VDQQPNAKCV EDLWKVGVRI KGPEN---GT FEREEIARCI QQVIGGDNAD
Clustal Co   *:*::  . :  :*  :      .    .:                   :   .    :

             ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                485        495        505        515        525        535
BMGT2        AVTQRVAELR TAAEAAVRKG GSSVVALGKF IETVTRR--- ---------- ----------
3HBF         IMRQKIVKLK ESAFKAVEQN GTSAMDFTTL IQIVTS---- ---------- ----------
2pq6         KMKQKAMELK KKAEENTRPG GCSYMNLNKV IKDVLLKQN- ---------- ----------
2acv         IVHKKVQEMK EMSRNAVVDG GSSLISVGKL IDDITGMSMS DINKNSELIF IPAPGIGHLA
2C1Z         KLRENLRALR ETADRAVGPK GSSTENFITL VDLVSKPKDV ---------- ----------
2vch         GVRNKMKELK EAACRVLKDD GTSTKALSLV ALKWKAHKKE LEQNGNH--- ----------
BMGT1        ELRANAWKWK KLAQDAMEEN GNSTKNIEDF VVQFFNMSLL LT-------- ----------
Clustal Co    :  .    :   :        * *   .  .
```

Fig. 5.6: Multiple structural alignment in sequence form. *B. monniera* (BMGT2), *Medicago truncaluta* (3hbf), *Medicago truncaluta* (2pq6), *Medicago truncaluta* (2acv), *Arabidopsis thaliana* (2vch) and *B. monniera* (BMGT1). Histidine (H) and aspartic acid (D) residues which are highlighted with red are involved in acceptor binding. Tryptophan (W), HCGWNS box and glutamine (Q) shown in green, pink and sky blue respectively are involved in donor binding.

5.3.1.1.6 Docking of different substrates to the *B. monniera* BMGT1

In order to confirm the experimental results of BMGT1 enzyme, docking studies were carried out. BMGT1 was docked with genistein, naringenin and kaempferol as acceptors and UDP-glucose as sugar donor molecule.

5.3.1.1.6.1 Docking of Genistein and UDP-glucose with BMGT1

A homology model of BMGT1 was docked with genistein and UDP-glucose. The structure of the model takes a GT-B-fold conformation comprising two distinct domains of the N- and C-terminal parts, which form a deep cleft accommodating the sugar donor and sugar acceptor substrates. The sugar acceptor is positioned interacting with the amino acid residues in the N-terminal domain while the sugar donor mainly interacts with the residues in the C-terminal region. The best docked pose of Enzyme-ligand complex constructed using Autodock and Flex docking programs is shown in Fig. 5.7.

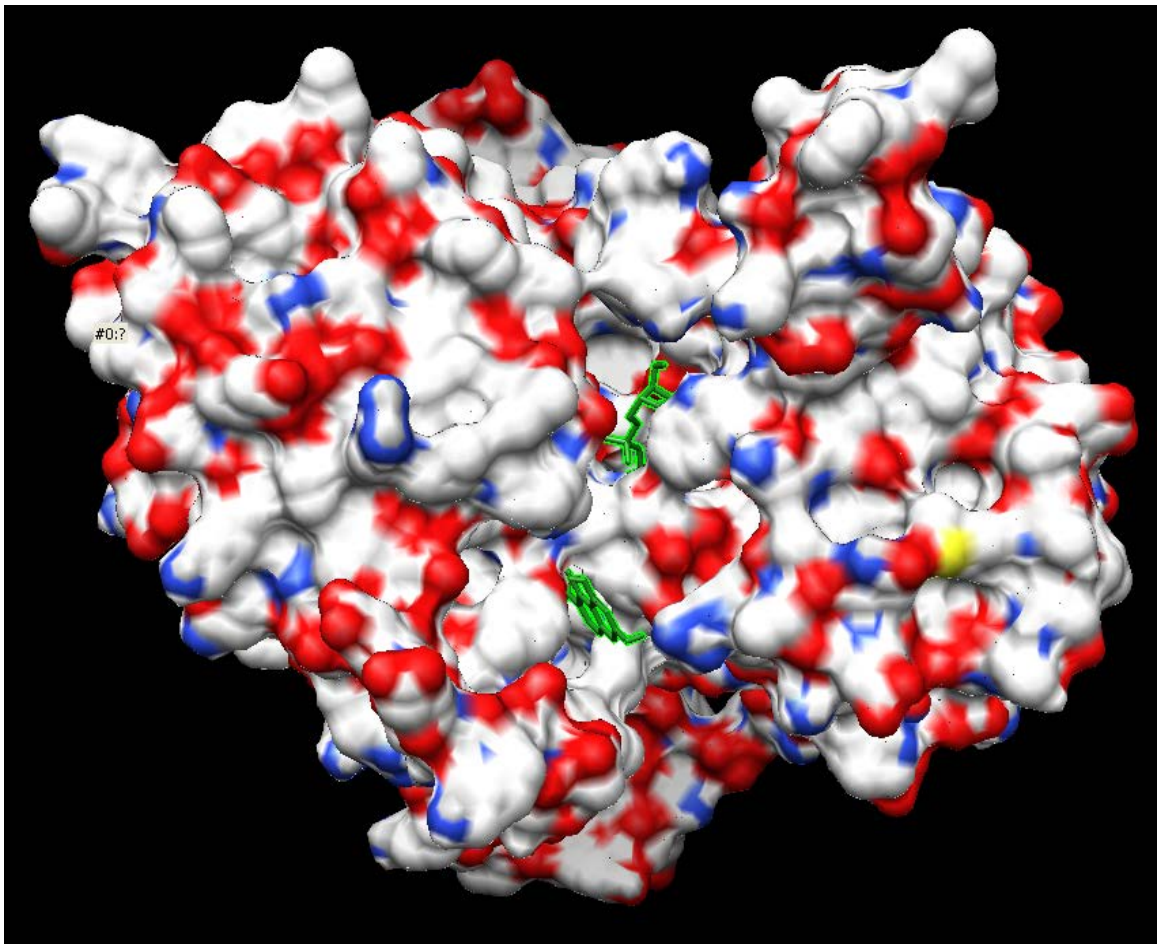


Fig. 5.7: Structure of BMGT1 protein with docked genistein (acceptor) and UDP-glucose (donor) molecules.

Theoretically, the binding energy of genistein with BMGT1 was calculated as -6.68 kcal/mol. Many variations of amino acid residues occur with different GT members, but their very similar active site structures ensure that their analogous residues will have most of the same interactions. Receptor ligand interactions are shown in SPDBV (http://spdbv.vital-it.ch/). HE2 (His 22) interacts with the O (genistein) at 1.84 Å in the acceptor binding site and HN (Trp 355) with O14 (UDP-glucose), HE22 (Gln 377) with O17 (UDP-glucose), OD1 and OD2 (Asp 376) with H20 and H21 (UDP-glucose) in the donor binding site at 1.77 Å, 2 Å and 1.91 Å respectively.

Total 13 residues were found in 4 Å vicinity of the docked ligands in BMGT1 pockets (Fig. 5.8). The residues lining the full active site pockets are Gln 20, His 22, Asn 24, Gln 134, Cys 335, Gln 337, His 352, Trp 355, Ser 357, Asp 376, Gln 377, Glu 360 and Tyr 81.

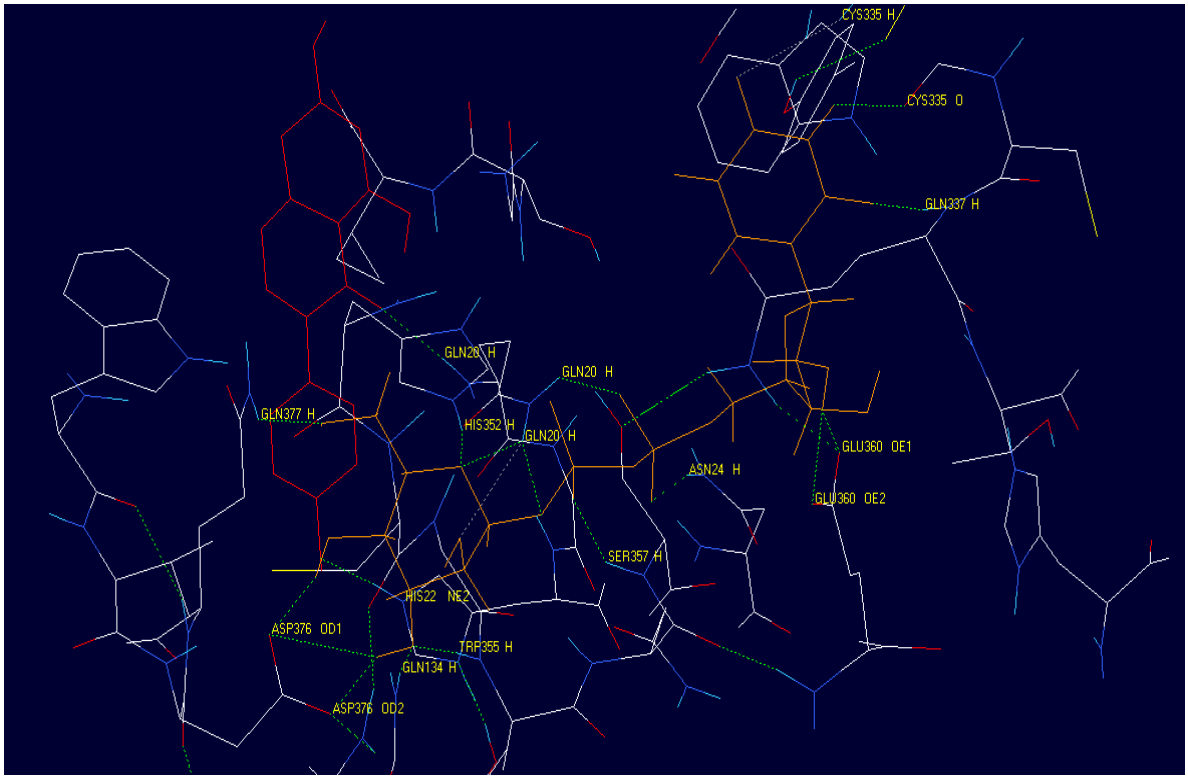


Fig. 5.8: Interactions of catalytic residues of BMGT1 with genistein (shown in red color) and UDP-glucose (shown in orange color) in the docked structure. Hydrogen bond interactions are shown in green color.

Complementing experimental results, docking also shows that histidine (His 22) deprotonates the 4-OH group of the acceptor (genistein) molecule and the deprotonated genistein will then interact with the UDP-glucose. In this process it adds sugar molecule to itself resulting in the formation of genistein 4-glycoside.

Out of 13 residues, histidine residue, His 22 (of BMGT1) is conserved in acceptor binding cleft within GTs and its role in deprotonation of acceptor molecule has been earlier described previously (Shao *et al.*, 2005). The residues conserved in donor binding cleft are Trp 355, Asp 376 and Gln 377 which interact with the donor (UDP-glucose) molecule.

5.3.1.1.6.2 Docking of Naringenin and UDP-glucose with BMGT1

A homology model of BMGT1 was also docked with naringenin and UDP-glucose. The best docked pose of Enzyme-ligand complex constructed using Autodock and Flex docking programs is shown in Fig. 5.9.

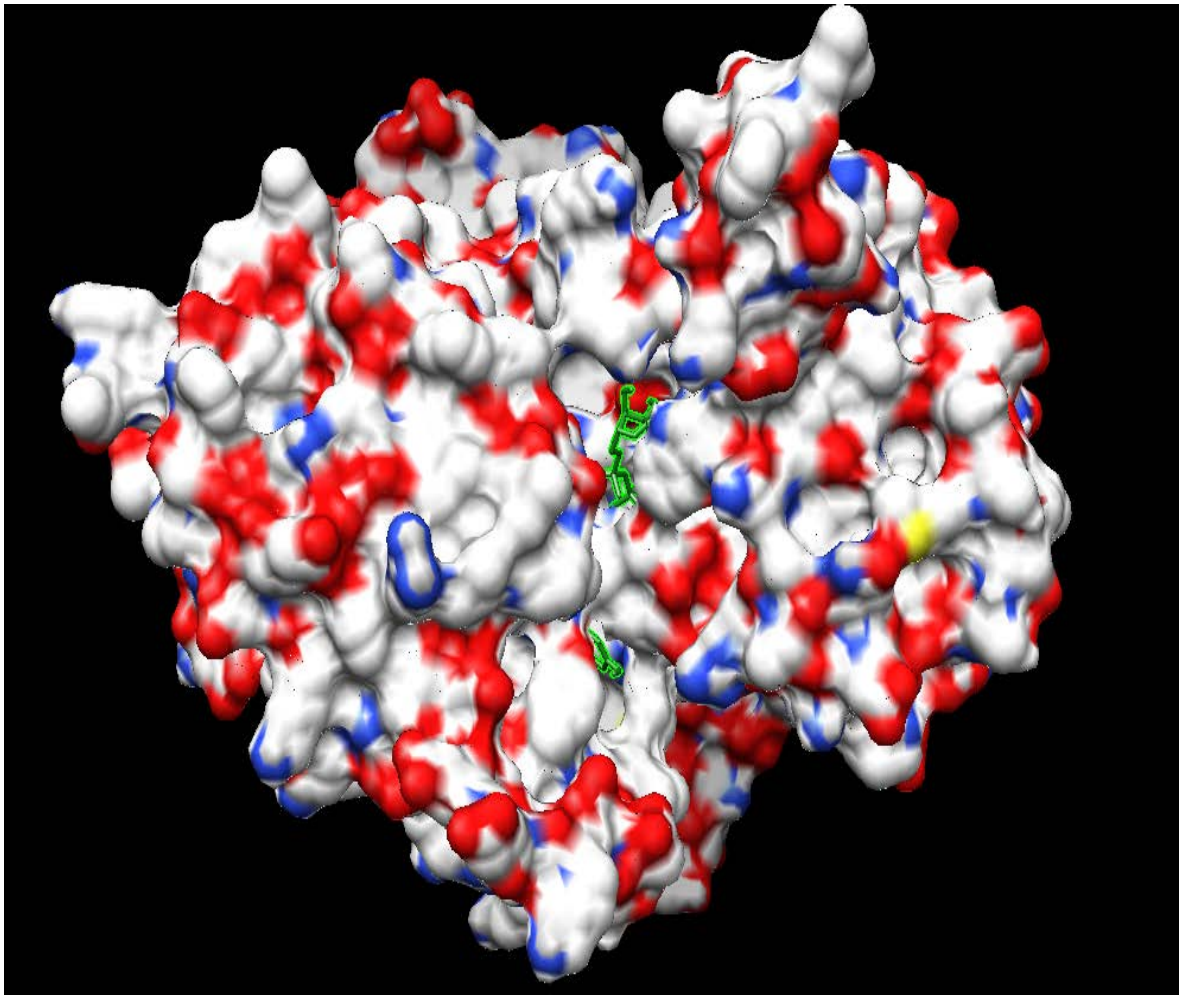


Fig. 5.9: Structure of BMGT1 protein with docked naringenin (acceptor) and UDP-glucose (donor) molecules.

The binding energy of naringenin with BMGT1 was calculated as - 2.23 kcal/mol. HE2 (His 22) interacts with the O (naringenin) at 1.94 Å in the acceptor binding site and HN (Trp 355)

with O14 (UDP-glucose), HE22 (Gln 377) with O (UDP-glucose), OD1 and OD2 (Asp 376) with H (UDP-glucose) in the donor binding site at 1.77 Å, 2 Å and 1.91 Å respectively.

Total 12 residues were found in 4 Å vicinity of the docked ligands in BMGT1 pockets (Fig. 5.10). The residues lining the full active site pockets are Gln 20, His 22, Gln 134, Cys 335, His 352, Trp 355, Ser 357, Asp 376, Gln 377, Glu 360, Asn 356 and Asn 83.

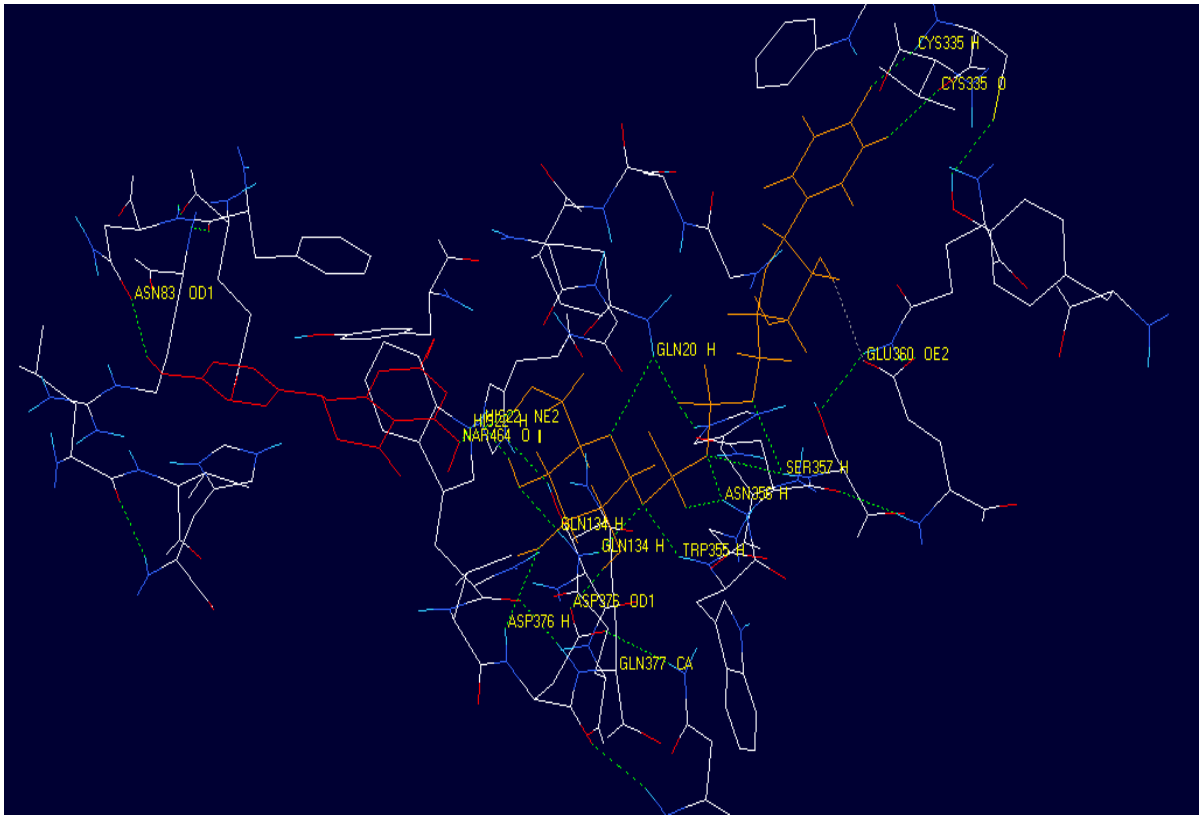


Fig. 5.10: Interactions of catalytic residues of BMGT1 with naringenin (shown in red color) and UDP-glucose (shown in orange color) in the docked structure. Hydrogen bond interactions are shown in green color.

Out of 12 residues, His 22, Trp 355, Asp 376 and Gln 377 residues are present in the acceptor and donor binding clefts of BMGT1 protein.

5.3.1.1.6.3 Docking of kaempferol and UDP-glucose with BMGT1

A homology model of BMGT1 was also docked with kaempferol and UDP-glucose. The best docked model was constructed using Autodock and Flex docking programs as shown in Fig. 5.11.

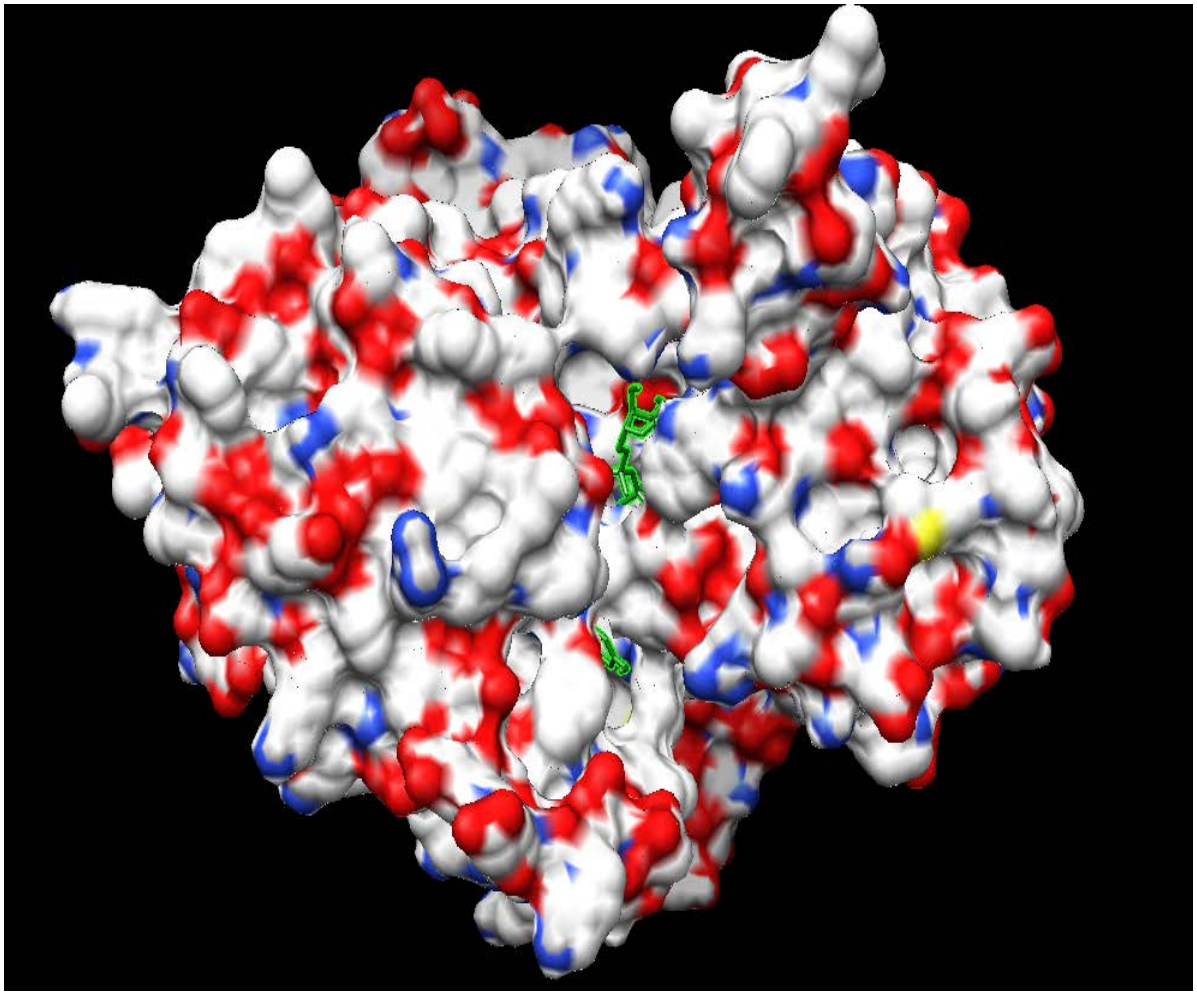


Fig. 5.11: Structure of BMGT1 protein with docked kaempferol (acceptor) and UDP-glucose (donor) molecules.

The binding energy of kaempferol with BMGT1 was calculated as – 3.41 kcal/mol. HE2 (His 22) interacts with the O4 (kaempferol) at 2.86 Å in the acceptor binding site.

Total 12 residues were found in 4 Å vicinity of the docked ligands in BMGT1 pockets (Fig. 5.12). The residues lining the full active site pockets are Gln 20, His 22, Gln 134, Cys 335, His 352, Trp 355, Ser 357, Asp 376, Gln 377, Glu 360, Asn 24 and Gln 337.

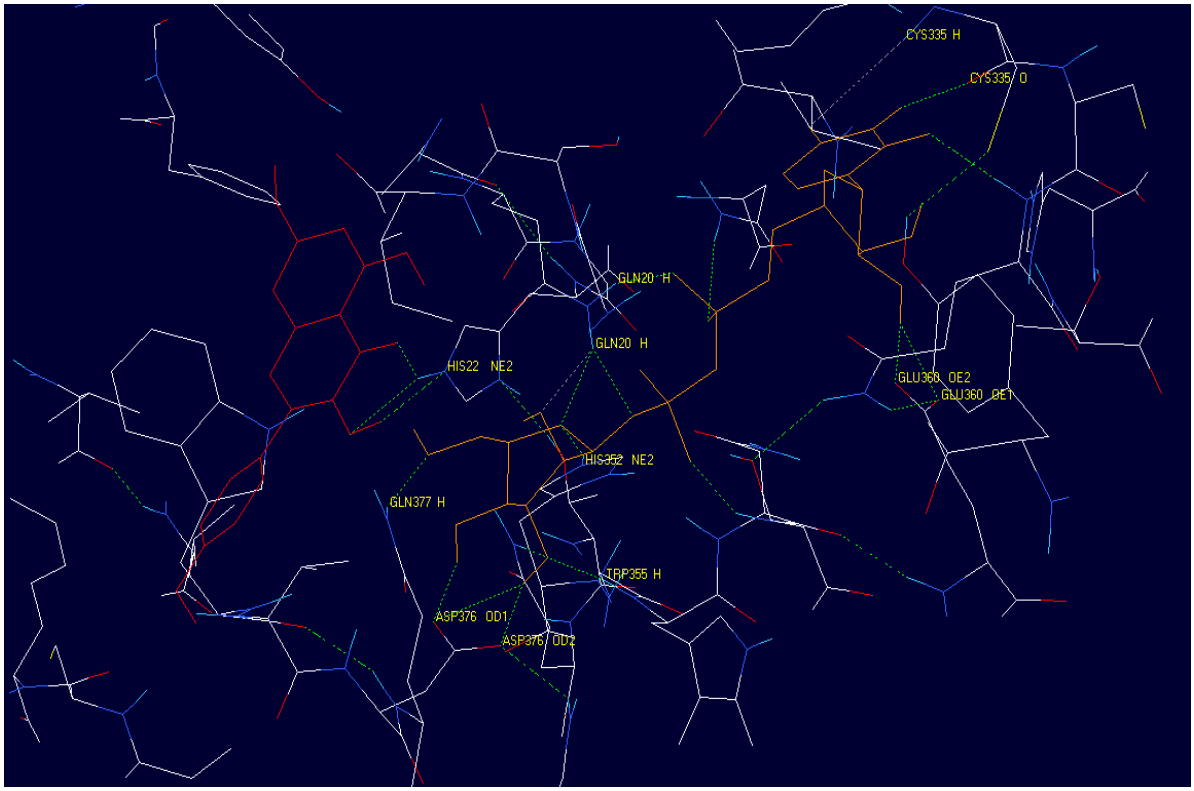


Fig. 5.12: Interactions of catalytic residues of BMGT1 with kaempferol (shown in red color) and UDP-glucose (shown in orange color) in the docked structure. Hydrogen bond interactions are shown in green color.

Out of 12 residues, His 22, Trp 355, Asp 376 and Gln 377 residues are present in the acceptor and donor binding clefts of BMGT1 protein.

5.3.1.2 Homology Modeling of *B.monniera* BMGT2 protein

Similarly in order to find out the homologous sequences in Protein Data Bank, the primary sequence of *B. monniera* BMGT2 was searched against PDB using BLASTP program at NCBI server (http://www.ncbi.nlm.nih.gov/blast). Among all the homologs,

glycosyltransferase from *V. vinifera* (PDB ID: 2c1z) was closest to BMGT2, with 29% identity at the amino acid sequence level. Glycosyltransferases from *M. truncaluta* and *A. thaliana* (PDB ID: 3hbf, 2pq6, 2acv and 2vch) showed 27%, 32%, 36% and 39% identity respectively. The three dimensional coordinates of *V. vinifera* (2c1z) was used as template to generate the 3D model of the BMGT2 using the program Modeller 9v8 (http://salilab.org). Out of 456 residues submitted for homology model, all residues were properly modeled in the structure. Of the 20 structures calculated for the same target (BMGT2) and the template (2c1z), the one with the lowest value of the MODELLER objective function was selected as the best model for BMGT2. The structures with least RMSD value were used for further analysis. The RMSD value of BMGT2 with templates *i.e.* 2c1z was found to be 1.65 Å. The final stable structure of BMGT2 is shown in Fig. 5.13.

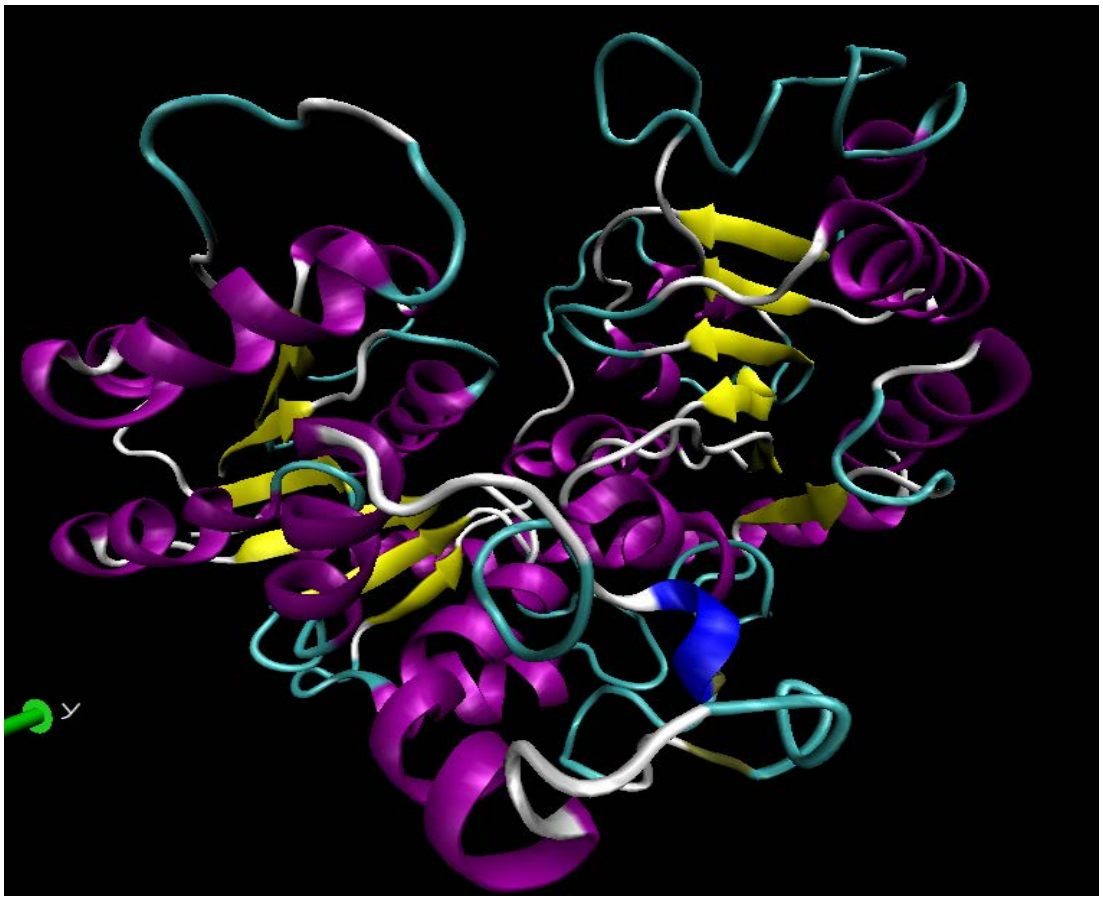


Fig. 5.13: Final 3D structure of BMGT2 enzyme. α-helices are represented in purple and beta sheets in yellow coloured ribbons.

5.3.1.2.1 Checking the quality of the model with PROCHECK

The quality of the model was examined using the program PROCHECK (Laskowski *et al*., 1993). The PROCHECK analysis showed 81.1%, 16.3%, 2.0% and 0.5% residues in the most favorable, additionally allowed, generously allowed and the disallowed regions of the Ramachandran map respectively (Fig. 5.14 & Table 5.3).

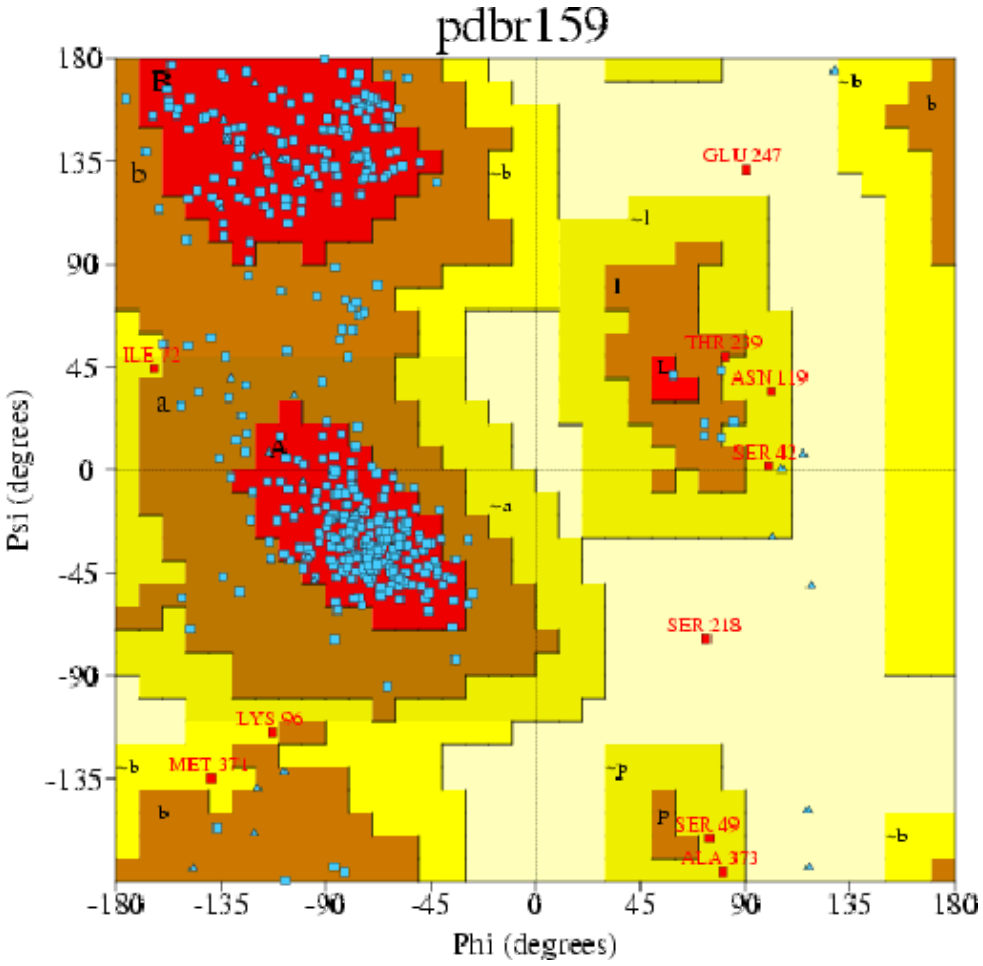


Fig. 5.14: Ramachandran(Φ,ψ) plot for modeled *B. monniera* BMGT2.

Table 5.3: % of residues falling in the core region of the Ramachandran's plot.

	No. of residues	%age
Most favoured regions	318	81.1%*****
Additional allowed regions	64	16.3%
Generously allowed regions	8	2.0%
Disallowed regions	2	0.5%*****
Non-glycine and non-proline residues	**392**	**100.0%**
End-residues (excl. Gly and Pro)	2	
Glycine residues	30	
Proline residues	32	
Total number of residues	**456**	

5.3.1.2.2 Molecular dynamics simulations

The structure of BMGT2 with the least modeller objective function, obtained from the modeller was improved by molecular dynamics as described in section 5.2.5.

5.3.1.2.3 Secondary structure calculation and comparison between target (*B. monniera* BMGT2) and template (2c1z)

In order to compare secondary structures of BMGT2 with 2c1z, web based program SPDBV was used to generate comparative picture of secondary structural elements (α helices, β sheets and turns) of BMGT2 with the already known 2c1z. The comparative secondary structural elements of BMGT2 with that of 2c1z are shown in Fig. 5.15.

```
BMGT2.B99990014   1     MEDAIVLYSSAEHLNSMLVPAKFISKHHPSISVIIISTAAESAAASVA-SVPSI-TYH
2c1z              6   TNPHVAVLAFPFSTHAAPLLAVVRRLAAAAPHAVFSFFSTSQSNASIFHDHTMQCNIKSY
BMGT2.B99990014       sssssss......hhhhhhhhhhhhhh....ssssss.hhhhhhh...-.....-sss
2c1z                ..sssssss.......hhhhhhhhhhhhh....ssssss.hhhhhhh........sssss
```

```
BMGT2.B99990014   57   RLPSAPLP----PDLPTSIIELFFEIPRFHNPFLHEALLEISQKS-NLRAFLIDFFCNST
2c1z              68   DISDGVPEGYVFAGRPQEDIELFTRAAP--ESFRQGMVMAVAETGRPVSCLVADAFIWFA
BMGT2.B99990014        ss..........----.....hhhhhhhhhhhh...hhhhhhhhhhh..-...ssssss...hh
2c1z                   ss..............hhhhhhhhhhh--hhhhhhhhhhhhhhh.....ssssss...hh

BMGT2.B99990014   112  FEVSTSLNIPTYFYLSGGACGLCALLYFPTIDEAVSPRDIGE-LNDFLE-IPGCPPVHSL
2c1z              126  ADMAAEMGVAWLPFWTAGPNSLSTHVYIDEIREKIGVSGIQGREDELLNFIPGMSKVRFR
BMGT2.B99990014        hhhhhhh..sssss....hhhhhhhhhhhhhhhhh.......-..sss.-......sss.
2c1z                   hhhhhhh..sssss....hhhhhhh...hhhhhhh..........sss........sss.

BMGT2.B99990014   170  DFPKAMWFRRSN-TYKHFLDTAGN-MRRASGIVTNSFDAIEFRAKEALSNSLCTPGLATP
2c1z              186  DLQEGIVFGNLNSLFSRMLHRMGQVLPKATAVFINSFEEL----DDSLTNDLK-SKLK--
BMGT2.B99990014        ............-hhhhhhhhh..-......sss..........hhhhhhhhhh.....
2c1z                   ............hhhhhhhhhhhhhh....sss.....----.hhhhhhhh-hh..--

BMGT2.B99990014   228  PVYVIGPLVAETNRKNGGEEHECLKWLDSQPIKSVIFLCFGRRGLFSAAQLKEMAIGLEN
2c1z              239  TYLNIGPFNLIT---------GCLQWLKERKPTSVVYISFGTVTTPPPAEVVALSEALEA
BMGT2.B99990014        sssss...hhh..........hhhhhh......sssss........hhhhhhhhhhhhh
2c1z                   sssss...hh..----------.hhhhhh.......ssss........hhhhhhhhhhhh

BMGT2.B99990014   288  SGHRFLWSVRSPPGPAAAKDPDLDALLPEGFMERTKDRGFVIKTWAPQKEVLSHEAVGGF
2c1z              299  SRVPFIWSLRD-----KARV-----HLPEGFLEKTRGYGMVVPW-APQAEVLAHEAVGAF
BMGT2.B99990014        h..ssssss.....................hhhhhhhh.sss.....hhhhhh.....ss
2c1z                   h..ssssss..-----hhhh-----.....hhh....sssss..-..hhhhhh.....ss

BMGT2.B99990014   348  VTHCGRSSVLEAVSFGVPMIGWPMYAEQRMQRVFMVEEMKVALPLAEEADGFVTAGELEK
2c1z              348  VTHCGWNSLWESVAGGVPLICRPFFGDQRLNGR-MVED-VLEIGVRIEG-GVFTKSGLMS
BMGT2.B99990014        ss...hhhhhhhhhh..sssss.....hhhhhhhhhh.................hhhhhh
2c1z                   ss...hhhhhhhhhh..sssss.....hhhhhh-hhhh-h..sss....-....hhhhhh

BMGT2.B99990014   408  RVRELMGLPAGKAVTQRVAELRTAAEAAVRKGGSSVVALGKFIETVTRR
2c1z              405  CFDQILSQEKGKKLRENLRALRETADRAVGPKGSSTENFITLVDLVSKPKD
BMGT2.B99990014        hhhhhh...hhhhhhhhhhhhhhhhhhhh....hhhhhhhhhhhhh...
2c1z                   hhhhhh...hhhhhhhhhhhhhhhhhhhh....hhhhhhhhhhhhh.....
```

Fig. 5.15: Pair wise sequence alignment of BMGT2 with 2c1z. The secondary structures of BMGT2 and templates (2c1z) are shown above. Helices and strands are represented by h and s respectively. The Figure was generated using web based program SPDBV.

As shown in Fig. 5.14 it is quite evident that the contents of secondary structural elements in BMGT2 are more or less similar to 2c1z. The aligned protein of BMGT2 with the templates contains 18 α-helices and 14 β-sheets respectively. These secondary structures are highly conserved with the template inspite of several amino acid differences.

5.3.1.2.4 Structural superimposition of *B. monniera* BMGT2 with 2c1z

Pair wise structural alignment of modeled *B. monniera* BMGT2 was done with 2c1z using combinatorial extension algorithm at WinCoot. Cartoon view of the superposed structures is shown in the Fig. 5.16.

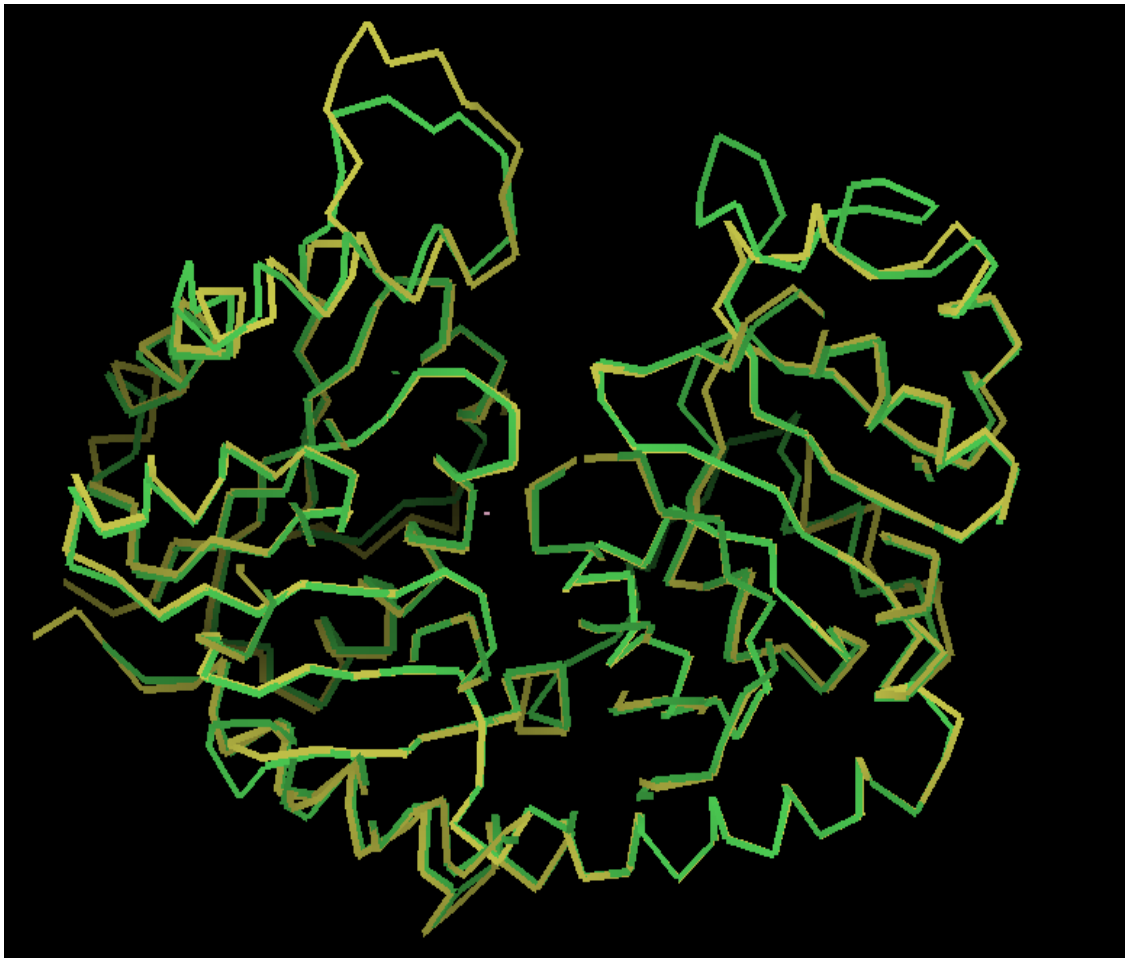


Fig. 5.16: cartoon view of superimposition of modelled BMGT2 (Dark green) with 2c1z (yellow).

The RMSD of Cα trace of superimposed BMGT2 with template (2c1z) was found to be 1.65 Å. This refined model was then used further for docking studies.

5.3.1.2.5 Docking of different substrates to the *B. monniera* BMGT2

Amino acid sequence and phylogenetic analysis of *B. monniera* BMGT2 showed that it is closely related to *Lamiaceae* family showing specificity to UDP-glucuronic acid as donor molecule. However so far, there are no UDP-glucuronic acid glycosyltransferases reported from *Bacopa*. Therefore, docking studies were carried out with UDP-glucuronic acid (donor) and acceptor molecules (baicalein & scutellarein), into the modeled BMGT2 in an attempt to find its probable natural substrate *in vivo*.

5.3.1.2.5.1 Docking of Baicalein and UDP-glucuronic acid with BMGT2

A homology model of BMGT2 was docked with baicalein and UDP-glucuronic acid. The structure of the model takes a GT-B-fold conformation comprising two distinct domains of the N- and C-terminal parts, which form a deep cleft accommodating the sugar donor and sugar acceptor substrates. The sugar acceptor is positioned interacting with the amino acid residues in the N-terminal domain while the sugar donor mainly interacts with the residues in the C-terminal region. The best docked pose of Enzyme-ligand complex constructed using Autodock and Flex docking programs is shown in Fig. 5.17.

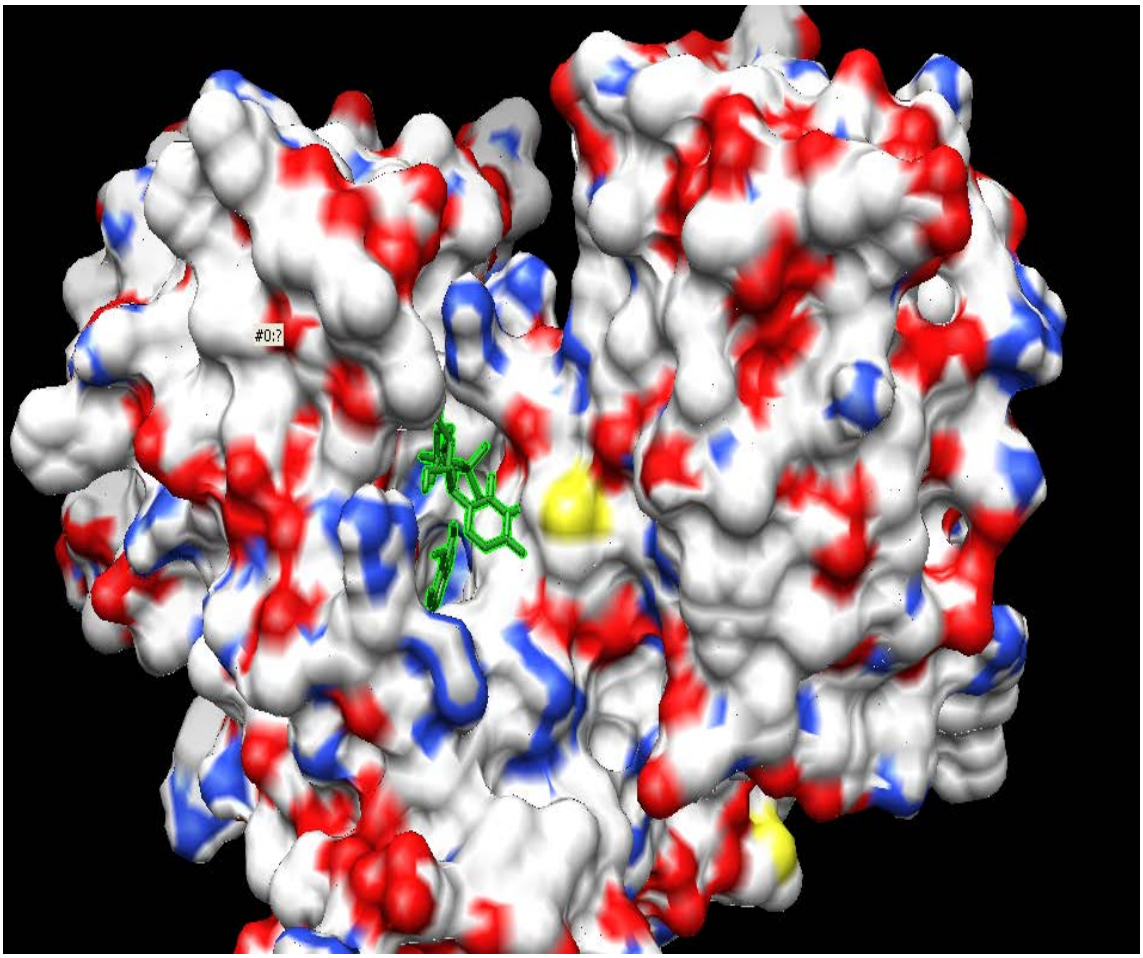


Fig. 5.17: Structure of BMGT2 protein with docked baicalein (acceptor) and UDP-glucuronic acid (donor) molecules.

Similarly theoretical binding energy of baicalein with BMGT2 was calculated as - 7.38 kcal/mol. Many variations of amino acid residues occur with different GT members, but their very similar active site structures ensure that their analogous residues will have most of the same interactions. Receptor ligand interactions are shown in SPDBV (http://spdbv.vital-it.ch/). HE2 (His 13) interacts with the O (baicalein) at 2.15 Å in the acceptor binding site and HH12 & HH22 (Arg 353) with O (UDP-glucuronic acid), HE21 & HE22 (Gln 375) with O (UDP- glucuronic acid), OE1 and OE2 (Glu 374) with H (UDP-glucuronic acid) in the donor binding site at 2.03 Å, 1.67 Å and 1.92 Å respectively.

Total 7 residues were found in 4 Å vicinity of the docked ligands in BMGT2 pockets (Fig. 5.18). The residues lining the full active site pockets are His 13, Arg 269, Arg 270, His 350, Arg 353, Gln 375 and Glu 374.

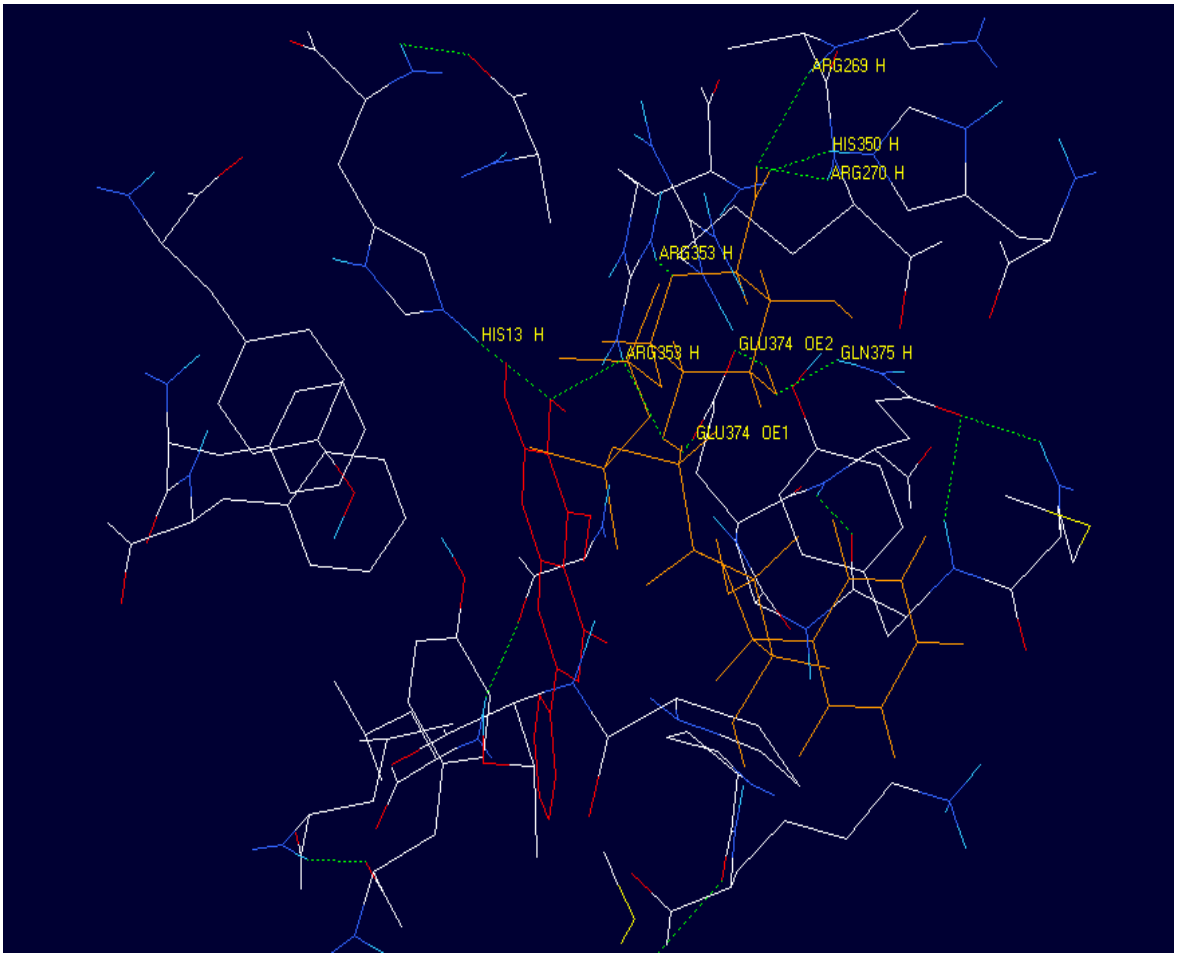


Fig. 5.18: Interactions of catalytic residues of BMGT2 with baicalein (shown in red color) and UDP-glucuronic acid (shown in orange color) in the docked structure. Hydrogen bond interactions are shown in green color.

Out of 7 residues, Histidine residue, His 13 (of BMGT2) is conserved in acceptor binding cleft and its role in deprotonation of acceptor molecule has been earlier described previously (Wang *et al*., 2005). The residues conserved in donor binding cleft are Arg 353, Glu 374 and Gln 375 which interact with the donor (UDP-glucuronic acid) molecule.

5.3.1.2.5.2 Docking of Scutellarein and UDP-glucuronic acid with BMGT2

A homology model of BMGT2 was docked with scutellarein and UDP-glucuronic acid. The best docked pose constructed using Autodock and Flex docking programs is shown in Fig. 5.19.

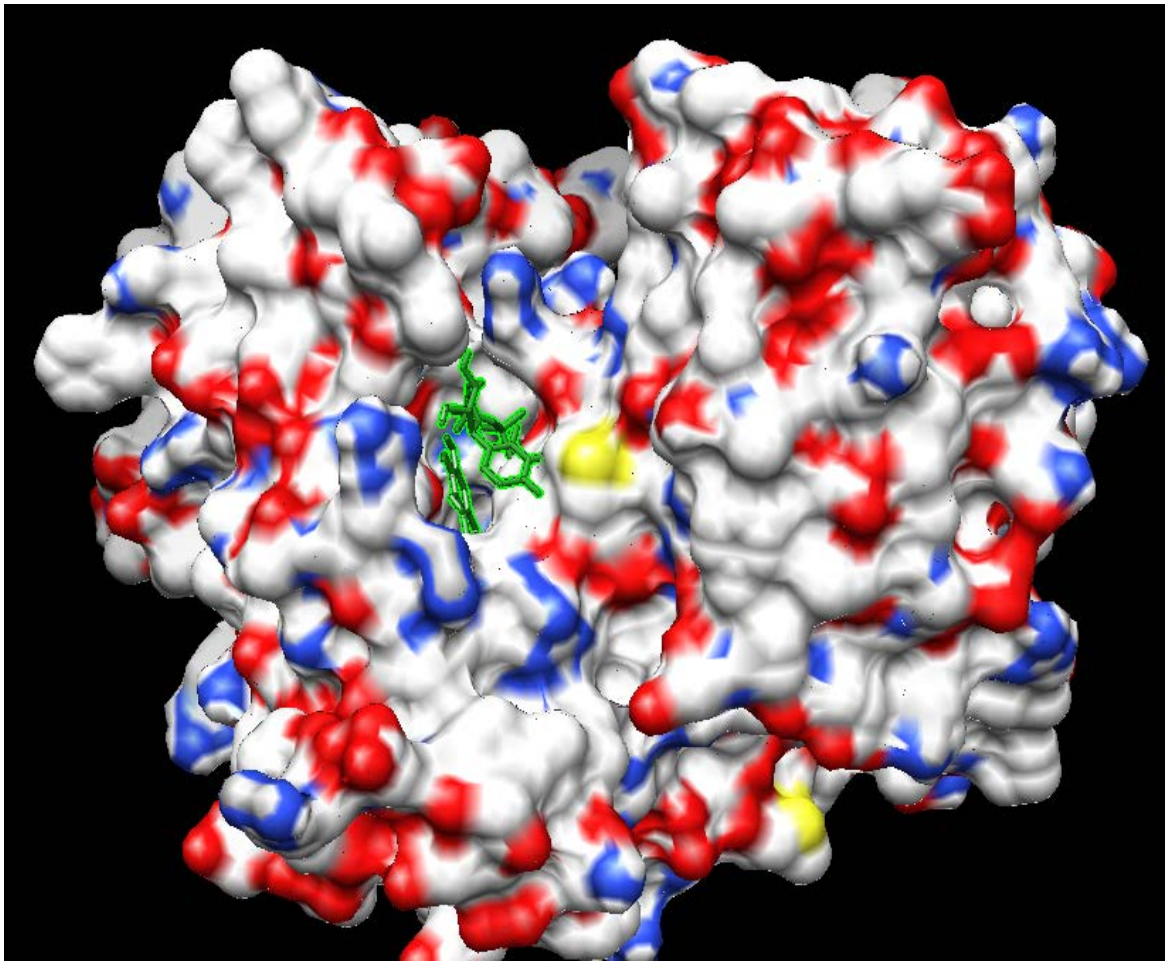


Fig. 5.19: Structure of BMGT2 protein with docked scutellarein (acceptor) and UDP-glucuronic acid (donor) molecules.

The binding energy of scutellarein with BMGT2 was calculated as – 3.96 kcal/mol. HE2 (His 13) interacts with the O (scutellarein) at 2.51 Å in the acceptor binding site and HH12 & HH22 (Arg 353) with O (UDP-glucuronic acid), HE21 & HE22 (Gln 375) with O (UDP-

glucuronic acid), OE1 and OE2 (Glu 374) with H (UDP-glucuronic acid) in the donor binding site at 2.03 Å, 1.67 Å and 1.92 Å respectively.

Total 8 residues were found in 4 Å vicinity of the docked ligands in BMGT2 pockets (Fig. 5.20). The residues lining the full active site pockets are His 13, Arg 269, Arg 270, His 350, Arg 353, Tyr 183, Gln 375 and Glu 374.

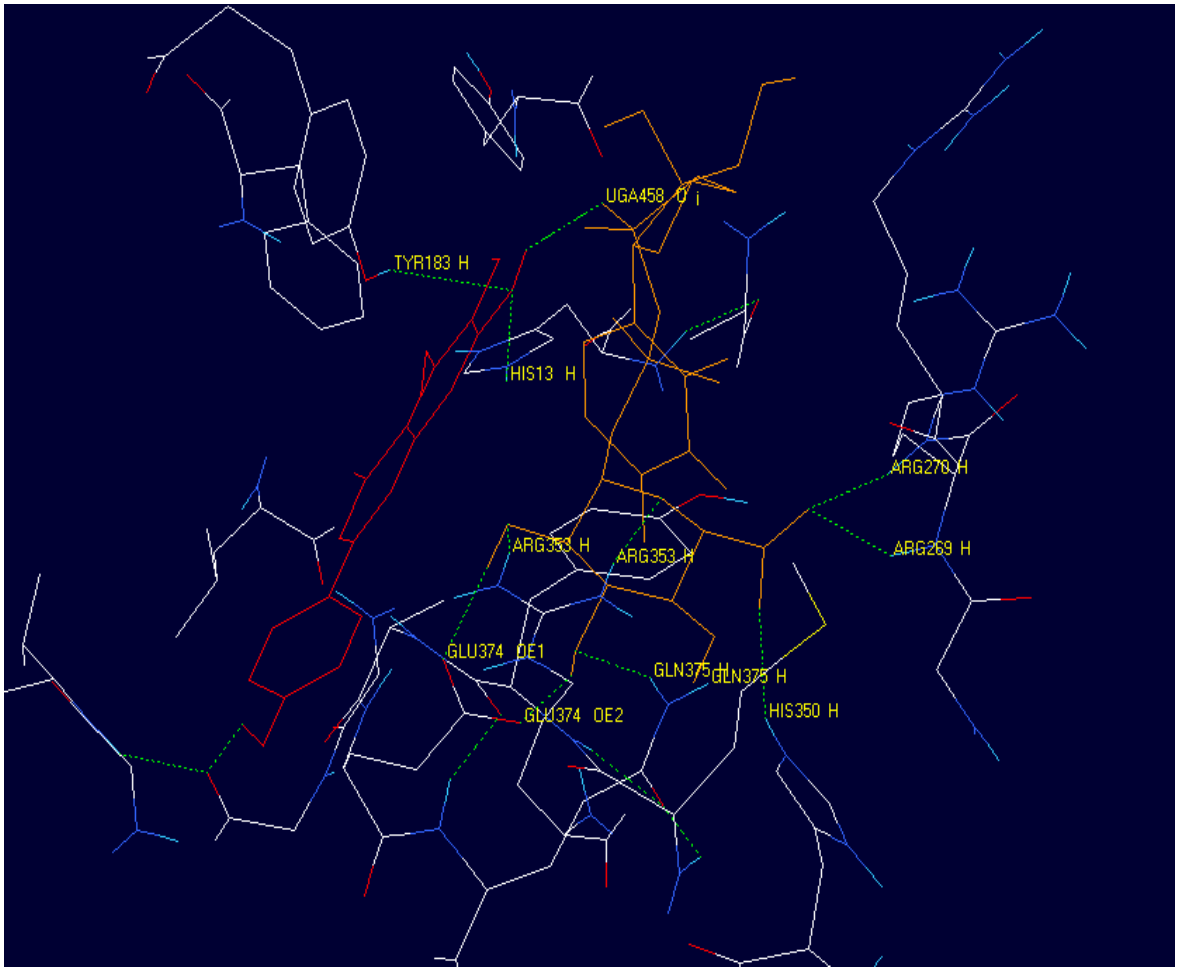


Fig. 5.20: Interactions of catalytic residues of BMGT2 with scutellarein (shown in red color) and UDP-glucuronic acid (shown in orange color) in the docked structure. Hydrogen bond interactions are shown in green color.

Out of 8 residues, histidine residue, His 13 (of BMGT2) is present in acceptor binding cleft and the residues conserved in donor binding cleft are Arg 353, Glu 374 and Gln 375.

Conclusions

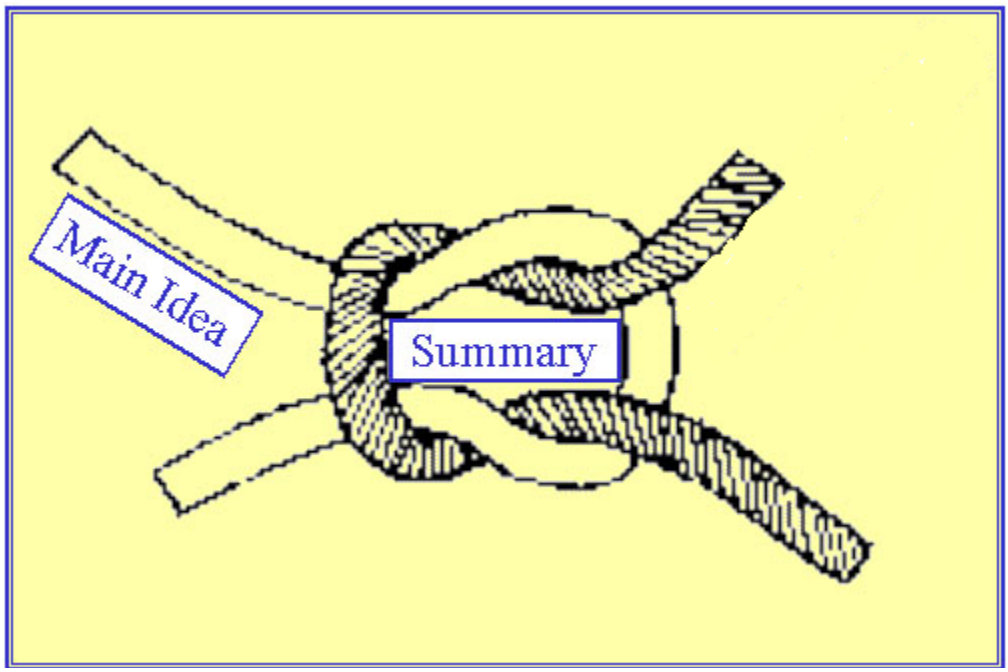
- Homology modelling and docking studies of BMGT1 was done to validate the experimental data on substrate specificity. Homology model of BMGT1 showed 88% residues in the most favorable region.

- Docking studies of BMGT1 was done with UDP-glucose (donor) and acceptor (genistein, naringenin and kaempferol) molecule. Docking studies confirmed the experimental results that BMGT1 is more specific to genistein forming genistein 4-glycoside. It also confirms the important active site residues i.e. His 22 in acceptor binding site and Trp 355, Asp 376, Gln 377 in donor binding cleft.

- In case of BMGT2 homology modelling and docking studies were performed to find out its probable natural substrate *in vivo*.

- Docking of BMGT2 was done with UDP-glucuronic acid (donor) and acceptor (baicalein & scutellarein) molecule. Docking studies showed that BMGT2 is more specific to baicalein (acceptor) and UDP-glucuronic acid (donor). It also confirmed the active site residue which is important for UDP-glucuronic acid specificity (Arg 353) already reported (One *et al*., 2009).

Summary



Summary

Bacopa monniera is known for its medicinal use which is mainly due to the presence of various glycosides in the plant. These glycosides (produced by an enzyme called as glycosyltransferase) are present in the plant in very low quantity so an alternative to increase the yield of such medicinally important glycosides was to isolate the enzyme, glycosyltransferase and over express it in suitable expression system. As a first step, two cDNAs of glycosyltransferase gene from *B. monniera* were isolated and characterized. The two cDNAs encoding GTs were designated as BMGT1 (Acc. No.FJ586244) and BMGT2 (Acc. No.FJ586245). The sequence analysis revealed an Open Reading Frame (ORF) of BMGT1 to be 1391 and BMGT2 to be 1372 bp. The predicted molecular weight and pI for BMGT1 and BMGT2 were estimated to be 52.3 kDa / 5.41 and 50 kDa / 6.21, respectively and were identified as cytosolic proteins (www.expasy.org/tools). Deduced amino acid sequence of BMGT1 and BMGT2 showed 49% and 72% identity at amino acid level with the GTs from *Lycium barbarum* and *Sesamum indicum* respectively. Deduced amino acid sequences of cDNAs from *B. monniera* contain consensus sequences for PSPG box which is conserved in whole of the plant GT family. Phylogenetic analysis of BMGT1and BMGT2 deduced amino acid sequences was done using MEGA 4 software program. The results show that BMGT1 is evolutionarily most similar to UDP-glucose:glycosyltransferases from *Lycium barbarum* of *Lamiaceae* family whereas BMGT2 is evolutionary similar to UDP-glucuronic acid:glycosyltransferases from *Sesamum indicum* of *Lamiaceae* family. Both BMGT1 and BMGT2 genes were over-expressed in *E. coli* and antibodies were raised against the purified protein from BMGT1.

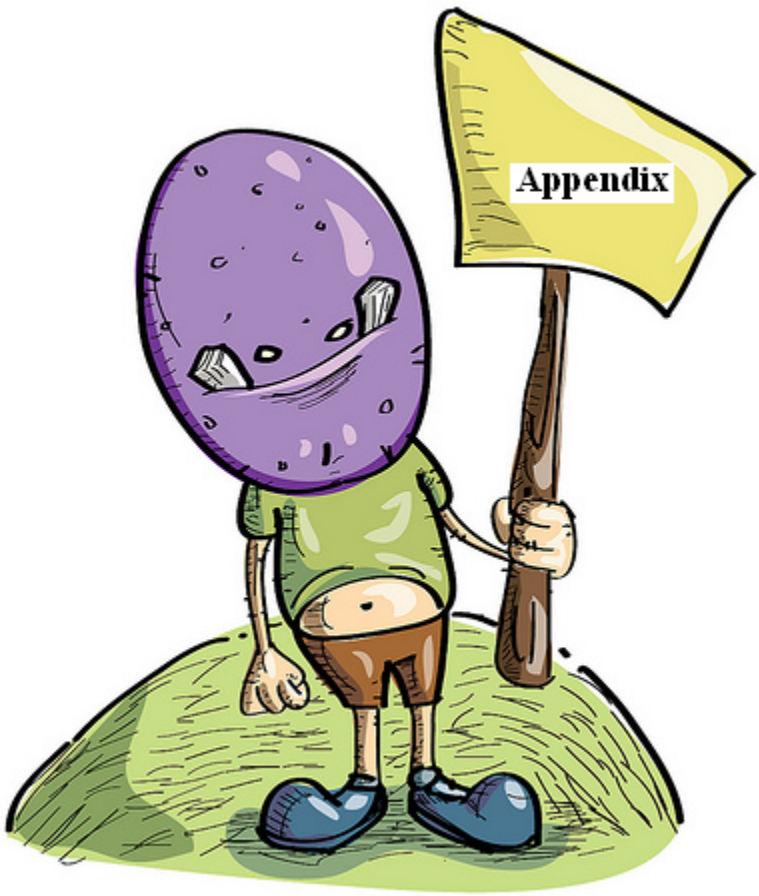
In order to carry out the enzyme assay, both BMGT1 and BMGT2 were expressed in their active form. Out of 22 substrates tested, BMGT1 could show activity with UDP-glucose as donor and genistein, naringenin and kaempferol as acceptor molecules. BMGT1 was more specific to genistein giving rise to genistein 4-glycoside. BMGT2 was also checked for activity with all these substrates and UDP-glucose as donor but it could not provide any activity. This was quite expected as the gene sequence of BMGT2 itself showed that it has conserved region which is specific to UDP-glucuronic acid as donor. The assay with UDP-glucuronic acid could not be performed since the substrates which accept glucuronic acid as

donor molecule were not available so only way to study the specificity of BMGT2 enzyme was through in silico studies i.e. homology modelling and docking studies.

Spatio-temporal expression profile was investigated using real-time PCR analysis under normal and stress induced (NaCl, methyl jasmonate, salicylic acid, mannitol, heat and cold shock treatments). Real-time PCR analysis showed differential expression at transcript level with maximum expression in leaf tissue in control plant for both BMGT1 and BMGT2 genes. In case of mannose stress, expression has increased 5 folds in root whereas in case of salicylic acid stress expression has increased almost 20 folds in root for BMGT1 gene. In case of BMGT2 the relative expression has increased 20 folds in stem incase of methyl jasmonate and salt stress whereas 10 folds increase in expression in leaf was seen in case of salicylic acid stress. Immuno-cytolocalization studies was also done using antibodies raised against BMGT1 protein. Transverse sections of immuno-cytolocalized root, leaf and stem showed that glycosyltransferases were localized in the vascular bundle.

Homology modeling and docking studies was done for both BMGT1 and BMGT2. Complementing experimental data, docking results also showed that BMGT1 is more specific to UDP-glucose (donor) and genistein (acceptor) forming genistein 4-glycoside. Docking results of BMGT2 showed that it is more specific to baicalein as acceptor and UDP-glucuronic acid as donor molecule. This study is the first time report of the glycosyltransferases from *B. monniera.*

Appendix



Appendix

Table 2.1: List of equipments used in present study

S. No	Equipment	Make
1	Balances	Contech/Sartorious
2	Water bath	Fisher Scientific/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 Electrophoresis Units	Bangalore Genei/Bio-Rad
11	Protein Gel Electrophoresis Units	Hoeffer Scientific/BioRad
12	Speed Vac concentrater	Savant/Eppendorf
13	pH-Meter	Microset
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	HPLC	Perkin Elmer
22	ELISA Plate Reader	Amersham (USA)
23	iBlot Gel Transfer System	Invitrogen
24	LC-MS	Waters
25	MALDI-MS	Waters

Table 2.2 Buffers and Solutions for DNA Electrophoresis

Name	Components	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 room temperature.
TBE buffer	90 mM Tris 90 mM Boric acid 2 mM EDTA	Room temperature.
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 (µ) filter and stored at room temperature.

Table 2.3 Buffers and Solutions for g-DNA isolation

Name	Components	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 ß-mercaptoethanol to 0.2 % before use.	Room temperature.

Table 2.4 Stock solutions for *E. coli* transformation and selection

Name	Components	Preparation and Storage
IPTG solution	200 mg/mL in SMQ	Sterile filtration and storage at -20 ºC
X-Gal (***5-bromo-4chloro***	20 mg/mL in N,N'-Dimethyl formamide (DMF)	Light sensitive, covered & store at -20 ºC

-3-indolyl-β-D-galactoside)		
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

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

Name	Components	Preparation and Storage
Solution I or Resuspension buffer	50 mM Glucose, 25 mM Tris-HCl (pH 8.0),10 mM EDTA (pH 8.0),	Store at 4 ºC
Solution II or Lysis buffer	0.2 N NaOH and 1% SDS	Freshly prepared
Solution III or Neutralisation buffer	3M Potassium acetate (pH-5.3)	Store at 4 ºC
RNase A	10 mg/mL	Store at -20 ºC
Other solutions or Reagents	Chloroform, Absolute ethanol, 3.0 M Sodium acetate, 70% ethanol & Deionized sterile water	Room temperature

Table 2.6 Buffers and Solutions for Gel Electrophoresis (native/ SDS-PAGE)

Name	Components	Preparation and Storage
Monomer solution	29.2% acrylamide 0.8% bis-acrylamide in water	Store at 4 ºC (in darkness)
Stacking gel	Distilled water 2.76 mL 1 M Tris-HCl (pH 6.8) 0.50 mL Acrylamide/bis 30% 0.65 mL 10% SDS	Freshly prepared

	0.04 mL 10% (w/v) APS 0.04 mL TEMED 4 µL	
Separating gel (10%)	Distilled water 3.162 mL 1.5 M Tris-HCl (pH 8.8) 2.0 mL Acrylamide/Bis 30% 2.67 mL 10% SDS (**SDS-PAGE**) 0.08 mL 10% (w/v) APS 0.08 mL TEMED 8 µL	Freshly prepared
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% SDS(**SDS-PAGE**) 3.3 mL ß-Mercaptoethanol 0.5 mL 0.5% Bromophenol blue 0.5 mL	Store at 4 ºC
10x SDS-electrode buffer	Tris base 15.1 g Glycine 94.0 g SDS 00.5 g Adjust pH-8.3 ***and make-up the volume up to 500 mL.***	Store at 4ºC Dilute 1:10 before use
Staining solution	Coomassie-blue R 250, 0.25 g in Methanol, 40 mL Acetic acid, 10 mL ***Make-up to 100 mL***	Store at 4 ºC (in darkness)
Destaining solution	Methanol, 40 mL Acetic acid, 10 mL ***Make- up to 100 mL***	Store at 4 ºC (in darkness)

Silver staining Fixer solution	40% Methanol, (150 mL) 10% acetic acid, (50 mL) ***Make-up to 100 mL***	Store at RT
Sensitizing solution	0.2% $Na_2S_2O_3$	Store at RT
Silver solution	0.2% silver nitrate (0.6 g) 0.01 % formaldehyde (225 µL) ***Make-up to 300 mL***	Prepare fresh in darkness
Developing solution	6% Na_2CO_3 (18 g) 0.02% formaldehyde (150 µL) ***Make-up to 300 mL***	Prepare fresh
Stop solution	1.5% Na_2EDTA (4.5 g) ***Make-up to 300 mL***	Store at RT

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

Name	Components	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

Table 2.8 Buffers and solutions for protein extraction under denaturing conditions

Name	Components	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)	Store at 4 ºC
Sonication buffer	100 mM Tris HCl (pH8.0) 50 mM Glycine	Store at 4 ºC
Dispersion buffer	100 mM Tris-HCl (pH8.0) 50 mM Glycine & 8 M Urea /6 M GuHCl	Store at 4 ºC

Table 2.9 substrates and products used for the study

Substrate/Products	Manufacturer
Apigenin	Sigma-Aldrich (USA)
Daidzein	Sigma-Aldrich (USA)
Naringenin	Sigma-Aldrich (USA)
Catechin	Sigma-Aldrich (USA)
Myricetin	Sigma-Aldrich (USA)
Isorhamnetin	Sigma-Aldrich (USA)
Genistein	Sigma-Aldrich (USA)
Luteolin	Sigma-Aldrich (USA)
Hesperetin	Sigma-Aldrich (USA)
Curcumin	Sigma-Aldrich (USA)
Kaempferol	Sigma-Aldrich (USA)
Diosgenin	Sigma-Aldrich (USA)

Capsaicin	Sigma-Aldrich (USA)
Vannilic acid	Sigma-Aldrich (USA)
Salicylic acid	Sigma-Aldrich (USA)
Resorcinol	Sigma-Aldrich (USA)
Thymol	Sigma-Aldrich (USA)
Euginol	Sigma-Aldrich (USA)
Benzoic acid	Sigma-Aldrich (USA)
4-nitro phenol	Sigma-Aldrich (USA)
Methyl vanilate	Sigma-Aldrich (USA)
3,4dimethoxyphenol	Sigma-Aldrich (USA)
Genistein 4-glycoside	Chromadex (USA)
Genistein 7-glycoside	Chromadex (USA)
Naringenin 7-glycoside	Chromadex (USA)
Kaempferol 7-glycoside	Chromadex (USA)

Table 2.10 Buffers and Solutions used for ELISA and Immuno-cytolocalization

Name	Components	Preparation and Storage
Crude protein extraction buffer	100 mM Tris-HCl pH.7.5, 2% PVPP, 2% PEG 4000, 5 mM DTT 1 mM PMSF	Store at room temperature
PBST	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	Store at room temperature
Antibody dilution Buffer	PBS with 0.25% BSA	Store at room temperature

Substrate buffer	200 mM Tris-HCl pH 9.5 0.5 mM $MgCl_2$	Store at room temperature
1X PBS	10 mM NaH_2PO_4- Na_2HPO 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 $MgCl_2$	Store at 4 ºC
BCIP/NBT mix	0.577 mM BCIP 0.122 mM NBT	Store at 4 ºC in dark
Stop Solution	10 mM EDTA	Store at room temperature

Table 2.11 Component of Murasighe & Skoog media/Different inducing media and hormones

Name	Components	Preperation 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	Store at 4 ºC
Minor components	0.147 mM $MnSO_4$ 5.3 x 10^{-2} mM $ZnSO_4$ 1.56 x 10^{-4} mM $CuSO_4$ 1.05 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 $Na_2Mo_4.2H_2O$	Store at 4 ºC
Vitamins	5.55 x10^{-2} mM Myoionsitol 4.06 x 10^{-3} mM Nicotinic acid 2.43 x 10^{-3}mM Pyridoxine HCl 2.96 x 10^{-4} mM Thymine HCl 2.66 x 10^{-2} mM Glycine	Store at 4 ºC
Iron	0.1 mM $FeSO_4.7H_20$	Store at 4 ºC

	0.1 mM Na_2EDTA	
BAP	1.776 mM BAP (*dissolve in NaOH and make up the volume by adding ethanol*)	Store at 4 ºC

Table 2.12 Different media used for studies bacterial growth

Name	Components	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 4 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 sterlized.

Table 2.13 Bacterial cell lines used

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) f80lacZΔM15 ΔlacX74 recA1 araD139

	Δ(araleu) 7697 galU galK rpsL (StrR) endA1 nupG
BL 21	F-, ompT hsdSB (rB – mB -) gal dcm (DE3) pLysS (CamR)

BRADFORD PROTEIN ASSAY

Principle

The assay is based on the observation that the absorbance maximum for an acidic solution of Coomassie Brilliant Blue G-250 shifts from 465 nm to 595 nm when binding to protein occurs. Both hydrophobic and ionic interactions stabilize the anionic form of the dye, causing a visible color change. The assay is useful since the extinction coefficient of a dye-albumin complex solution is constant over a 10-fold concentration range.

Reagent required

1. Bradford reagent: Dissolve 10 mg Coomassie Brilliant Blue G-250 in 5 mL 95% ethanol, add 10 mL 85% (w/v) phosphoric acid. Dilute to 100 mL when the dye has completely dissolved, and filter through Whatman #1 paper just before use.
2. (Optional) 1 M NaOH (to be used if samples are not readily soluble in the color reagent).

The Bradford reagent should be a light brown in color. Filtration may have to be repeated to rid the reagent of blue components.

ASSAY

1. Warm up the spectrophotometer 15 min. before use.
2. Dilute samples with buffer to an estimated concentration of 20 to 200 micrograms/mL
3. Prepare standards containing a range of 20 to 200 micrograms protein (albumin or gamma globulin are recommended) to a standard volume (generally 1 mL or less).
4. Prepare unknowns to estimated amounts of 20 to 200 micrograms protein per tube, same volume as the unknowns.
5. Add 0.25 mL 1 M NaOH (Optional) to each sample and vortex.
6. Add 5 mL dye reagent and incubate 5 min.
7. Measure the absorbance at 590 nm

ANALYSIS

Prepare a standard curve of absorbance versus micrograms protein (or *vice versa*), and determine amounts from the curve. Determine concentrations of original samples from the amount protein, volume/sample, and dilution factor, if any.

Publications

Publications

Research Papers under preparation

1. **Ruby**, R. J. Santosh Kumar, R.K. Vishwakarma, Somesh singh, Dr. B.M. Khan* (2011). “Molecular cloning, characterization and expression analysis of glycosyltransferase gene from *Bacopa monniera* (Manuscript under preparation).

2. **Ruby,** Ranu Sharma, Dr. C.G. Suresh*, Dr. B.M. Khan* (2011). In silico studies of two glycosyltransferase genes from *Bacopa monniera* (Manuscript under preparation.

3. R. J. Santosh Kumar, **Ruby,** R.K. Vishwakarma, Somesh singh, Dr. B.M. Khan* (2011). “Molecular cloning, characterization and expression studies of glycosyltransferase gene from *Withania somnifera* (Manuscript under preparation).

4. R.K. Vishwakarma, **Ruby,** Somesh singh, R. J. Santosh Kumar, Dr. B.M. Khan* (2011). “Isolation, cloning, characterization and expression studies of Acetyl CoA transferase gene from *Bacopa monniera* (Manuscript under preparation).

5. Somesh singh, R. J. Santosh Kumar, R.K. Vishwakarma, **Ruby**, Dr. B.M. Khan* (2010). “Isolation, cloning, and characterization of Flavanoid Glucosyltransferase gene from *Withania somnifera* (Manuscript under preparation).

Abstracts / Proceedings published:-

1. "Two different glycosyltransferases from *Bacopa monniera* involved in flavanoid biosynthesis". **Ruby**, R. J. Santosh Kumar, R.K. Vishwakarma, Somesh singh, Dr. B.M. Khan*. Abstract of the poster presented in National Symposium on recent advances in plant tissue culture and biotechnological researches in India (2011).

2. "Secondary metabolite flavonoid glycosyltransferases from *Withania somnifera*". R. J. Santosh Kumar, **Ruby,** Somesh singh, R.K. Vishwakarma, Dr. B.M. Khan*. Abstract of the poster presented in National Symposium on ecent advances in plant tissue culture and biotechnological researches in India (2011).

3. "Development of transgenic *Bacopa monniera* plants for altering mevalonate pathway gene(s) with reference to increased bacoside content". R.K. Vishwakarma, **Ruby,** Somesh singh, R. J. Santosh Kumar, Dr. B.M. Khan*. Abstract of the poster presented in National Symposium on recent advances in plant tissue culture and biotechnological researches in India (2011).

4. "Metabolic engineering of terpenoids biosynthesis pathway in *Bacopa monniera*". R.K. Vishwakarma, **Ruby**, S. Singh, R. J. Santosh Kumar, D. Prashant, P. Sharma and Dr. B.M. Khan*. International symposium on Aeromatic and medicinal plants (AROMED), CIMAP, Lucknow, India. (2010).

5. "Glycosyltransferase involved in flavonol glycoside biosynthesis in *Withania somnifera*". S. Singh, R. J. Santosh Kumar, R.K. Vishwakarma, **Ruby**, Dr. B.M. Khan*. International symposium on Aeromatic and medicinal plants (AROMED), CIMAP, Lucknow, India. (2010).

6. “Role of squalene synthesis in withanolide biosynthesis”. P. Sharma, R.J.Santosh Kumar, R.K. Vishwakarma, **Ruby**, S.Singh and Dr. B.M. Khan*. International symposium on Aeromatic and medicinal plants (AROMED), CIMAP, Lucknow, India. (2010).

7. “Molecular approaches for production of pharmacologically active plant glycosides, A valuable strategy for finding new lead compounds”. **Ruby**, R. J. Santosh Kumar, R.K. Vishwakarma, Somesh singh, Dr. B.M. Khan*. Abstract of the poster presented in INSA Platinum jubilee international symposium on Research in Molecular medicine based on natural resources and traditional knowledge (2009).

8. “Molecular study of Glycosylation of polyphenolic compounds from Withania somnifera and its applications for human health”. Somesh singh, R. J. Santosh Kumar, R.K. Vishwakarma, **Ruby**, Dr. B.M. Khan*. (INSA 2009).

9. "An approach for genetic modification of Leucaena leucocephala for Eco-friendly pulp and paper production". Manish Arha, Gupta K. Sushim, , Mohd Noor, Yadav Arun, Srivastava Sameer, K. Pallavi, OU. Abhilash, RJ. Santosh, Malini Kaul, Jay Kumar, Azfar Quraishi, Ulfat Iqbal, **Ruby**, Khan B. M., Rawal S. K*. Abstract of the Poster presented in the International symposium on Frontiers in Genetic and Biotechnology- Retrospect and Prospect at Osmania University (2005).

References

References

Abdalla MF, Saleh NAM, Gabr S, Abu-Eyta AM, El-Said H (1983). Flavone glycosides of *Salvia triloba*. Phytochemistry. 22:2057–2060.

Achnine L, Huhman DV, Farag MA, Sumner LW, Blount JW, and Dixon RA (2005) Genomics-based selection and functional characterization of triterpene glycosyltransfererases from the model legume *Medicago truncatula*. Plant. J 41: 875–887.

Aloe A, Alleve E, Fiore M (2002) Stress and nerve growth factor findings in animal models and humans. Pharmacol. Biochem. Behav. 73:159-66.

Anbarasi K, *et al*. (2005a) Effects of bacoside-A on membrane-bound ATPases in the brain of rats exposed to cigarette smoke. J. Biochem. Mole. Toxicol. 19:59-65.

Anbarasi K, Vani G, Shyamala Devi CS (2005b) Protective effects of bacoside-A on cigarette smoking-induced brain mitochondrial dysfunction in rats. J. Env. Pathol. Toxicol. Oncol. 24:225-234.

Anbarasi K, *et al*. (2005c) Induction of Hsp70 and apoptosis during cigarette smoking in rat brain: Modulation by bacoside A. Presented at the 6th IBRO School of Neuroscience, Bangalore, India, Aug 8-20.

Anbarasi K, *et al*. (2006) Effects of bacoside-A on brain antioxidant status in cigarette smoke exposed rats. Life Sci. 78:1378-84.

Anon (2004) *Bacopa monnieri*. Monograph. Altern. Med. Rev. 9:79-85.

Aoshiba K, Tamaoki, J, Nagai A (2001) Acute cigarette smoke exposure induces apoptosis of alveolar macrophages. AJP- Lung Cellular and Molecular Physiology. 281:L1392-L1401.

APG II (2003) An update of the angiosperm phylogeny group classification for the orders and families of flowering plants: APG II. Bot. J. Linn. Soc. 141:399–436.

Asen, S., Norris KH, Stewart RN (1972) Copigment of aurone and flavone from petals of *Antirrhinum majus*. Phytochemistry. 11:2739–2741.

Autonomous (2002) Indian Herbal Pharmacopoeia. Mumbai.

Bafna PA, Balaraman R (2005) Antioxidant activity of DHC-1, an herbal formulation, in experimentally-induced cardiac and renal damage. Phytother Res. 19:216-21.

Barrett SCH, Strother JL (1978) Taxonomy and natural history of *Bacopa* in California. Syst. Bot. 5:408–419.

Basu N, Rastogi RP, Dhar ML (1967) Chemical examination of *Bacopa monniera* Wettst: Part III-Bacoside B. Indian J. Chem. 5:84–95.

Bhakuni DS, Dhar ML, Dhar MM, Dhawan BN, Mehrotra BN (1969) Screening of Indian plants for biological activity: Part II. Ind. J. Exp. Biol. 7:250-62.

Bhattacharya, SK, Ghosal S (1998) Anxiolytic activity of a standardized extract of *Bacopa monniera*: an experimental study. Phytomedicine. 5:77–82.

Bhattacharya SK, Kumar A, Ghosal S (1999) Effect of *Bacopa monniera* on animal models of Alzheimer's disease and perturbed central cholinergic markers of cognition in rats. In: Siva Sankar, D.V. (Ed.), Molecular Aspects of Asian Medicines. PJD Publications, New York.

Bhattacharya SK, *et al*. (2000) Antioxidant activity of *Bacopa monniera* in rat frontal cortex, striatum and hippocampus. Phytotherapy Res. 14:174-179.

Bischoff SJ (2008) Quecetin: potentials in the prevention and therapy of diseases. Curr. Opin. Clin. Nutr. Metab. Care. 11:733 –740.

Bolam DN, Roberts S, Proctor MR, Turkenburg JP, Dodson EJ, Martinez-Fleites C, Yang M, Davis BG, Davies GJ, Gilbert HJ (2007) The crystal structure of two macrolide glycosyltransferases provides a blueprint for host cell antibiotic immunity. Proc. Natl. Acad. Sci. USA. 104:5336–5341.

Bone K, (1996) Clinical applications of Ayurvedic and Chinese herbs: Monographs for the western herbal practitioner. Warwick, Queensland: Physiotherapy press.

Bose KC, Bose NK (1931) Observations on the actions and uses of *Herpestis monniera*. J. Indian Med. Assoc. 1:60.

Bourne Y, Henrissat B (2001) Glycoside hydrolases and glycosyltransferases:families and functional modules. Curr. Opin. Struct. Bio. l11:593-600.

Bowles D, Isayenkova J, Lim EK, Poppenberger B (2005) Glycosyltransferases: managers of small molecules. Curr. Opin. Plant Biol. 8:254–263.

Bradford MM (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein dye binding. Anal. Biochem. 72:248–254.

Brazier-Hicks M, Offen WA, Gershater MC, Revett TJ, Lim EK, Bowles DJ, Davies GJ, Edwards R (2007) Characterization and engineering of the bifunctional N- and O-glucosyltransferase involved in xenobiotic metabolism in plants. Proc. Natl. Acad. Sci. USA. 104:20238–20243.

Bremer B, Bremer K, Chase MW, Reveal JL, Soltis DE, Soltis PS, Stevens PF (2003) An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG II. Bot. J. Linn. Soc. 141:399–436.

Burge CB, Karlin S (1998) Finding the genes in genomic DNA. Current Opinion in Structural Biology. 8:346-354.

Burge CB (1998) Modeling dependencies in pre-mRNA splicing signals. In: Salzberg S, Searls D, Kasif S, eds. Computational Methods in Molecular Biology, Amsterdam: Elsevier Science. 127-163.

Butterweck V, Juergenliemk G, Nahrstedt A, Winterhoff H (2000) Flavonoids from *Hypericum perforatum* show antidepressant activity in the forced swimming test. Planta Med. 66:3–6.

Campbell WH, Gowri G. (1990) Codon usage in higher plants, green algae, and cyanobacteria. Plant Physiology. 92:1–11.

Chopra RN, Nayar L, Chopra IC (1956) Glossary of Indian Medicinal Plants, vol. 32. Council of Scientific and Industrial Research, New Delhi.

Chopra RN, Indigenous drugs of India. 2nd ed. Calcutta, India: U.N. Dhur and sons;1958:341.

Chatterji N, Rastogi RP, Dhar ML (1965) Chemical examination of *Bacopa monniera* Wettst: Part II. The constitution of Bacoside A. Indian J Chem. 3:24–30.

Campbell JA, Davies GJ, Bulone V, Henrissat B (1997) A classification of nucleotide-diphospho-sugar Glycosyltransferases based on amino acid sequence similarities. Biochem J. 326(Pt3):929-939.

Chandel RS, Kulshreshtha DK, Rastogi RP (1977) Bacogenin A3: a new sapogenin from *Bacopa monniera*. Phytochemistry. 16:141–143.

Chakravarty AK, Sarkar T, Masuda K, Shiojima K, Nakane T, Kawahara N (2001) Bacopaside I and II: two pseudojujubogenin glycosides from *Bacopa monniera*. Phytochemistry. 58:553–556.

Chakravarty AK, Sarkar T, Nakane T, Kawahara N, Masuda K (2002) New phenylethanoid glycosides from *Bacopa monniera*. Chem. Pharm. Bull. 50:1616–1618.

Chakravarty AK, Garai S, Masuda K, Nakane T, Kawahara N (2003) Bacopasides III–V: three new triterpenoid glycosides from *Bacopa monniera*. Chem. Pharm. Bull. 51:215–217.

Chowdhuri DK, *et al*. (2002) Antistress effects of bacosides of *Bacopa monniera*: modulation of Hsp70 expression, Superoxide dismutase and cytochrome P450 activity in rat brain. Phytotherapy Res. 16:639-645.

Channa S, *et al.* (2003) Broncho-vasodilatory activity of fractions and pure constituents isolated from *Bacopa monniera.* J. Ethnopharmacol. 86:27-35.

Channa S, Dar A, Anjum S, Yaqoob M, Rahman A (2006) Anti-inflammatory activity of *Bacopa monniera* in rodents. J Ethanopharmacol. 104:286-9.

Chunekar KC (1960) Bhav Prakasa Nighantu, vol. 372. Chaukhamba Bharati Publications, Varanasi (Hindi translation).

Colasanti M, Suzuki H (2000) The dual personality of NO. TPS. 21:249–252.

Cao PJ, Bartley LE, Jung KH, Ronald PC (2008) Construction of a rice glycosyltransferase phylogenomic database and identification of rice-diverged glycosyltransferases. Molecular Plant. 1:858–877.

Chong J, Baltz R, Schmitt C, Beffa R, Fritig B, Saindrenan P (2002) Downregulation of a pathogen-responsive tobacco UDP-Glc:phenylpropanoid glucosyltransferase reduces scopoletin glucoside accumulation, enhances oxidative stress, and weakens virus resistance. Plant Cell. 14:1093–1107.

Creelman RA, Mullet JE (1997) Biosynthesis and action of jasmonates in plants. Plant Mol. Biol. 48:355–381.

Chou TC, Chang LP, Li CY, Wong CS, Yang SP (2003) The anti-inflammatory and analgesic effects of baicalin in carrageenan-evoked thermal hyperalgesia. Anesth. Analg. 97:1724–1729.

Coutinho PM, Deleury E, Davies GJ, Henrissat B (2003) An evolving hierarchical family classication for glycosyltransferases. J. Mol. Biol. 328:307-317.

Dar A, Channa S (1997a) Relaxant effect of ethanol extract of *Bacopa monniera* on trachea, pulmonary artery and aorta from rabbit ad guinea-pig. Phytother. Res. 11:323–325.

Dar A, Channa S (1997b) Bronchodilatatory and cardiovascular effects of an ethanol extract of *Bacopa monniera* in anaesthetized rats. Phytomedicine. 4:319–323.

Dar A, Channa S (1999) Calcium antagonistic activity of *Bacopa monniera* on vascular and intestinal smooth muscles of rabbit and guinea-pig. J. Ethnopharmacol. 66:167–174.

Das A, Shanker G, Nath C, Pal R, Singh S, Singh H (2002) A comparative study in rodents of standardized extracts of *Bacopa monniera* and *Ginkgo biloba*. Pharmacol. Biochem. Behav. 73:893–900.

Dave UP, Chauvan V, Dalvi J (1993) Evaluation of BR-16 A (Mentat) in cognitive and behavioural dysfunction of mentally retarded children: a placebo-controlled study. Indian J. Pediatr. 60:423–428.

Deb DD, *et al*. (2008) In vitro safety evaluation and anticlastogenic effects of BacoMind™ on human lymphocytes. Biomed. Environ. Sci. 21:7-23.

Des Marais DL, Rausher MD (2008) Escape from adaptive conflict after duplication in an anthocyanin pathway gene. Nature. 454:762–765.

Dhawan BN, Singh HK (1996) Pharmacology of ayurvedic nootropic *Bacopa monniera*, Abstr. No. NR 59. Int. Conv. Biol. Psychiat. Bombay.

Dharmani P, Palit G (2006) Exploring Indian medicinal plants for antiulcer activity. Ind. J. Pharmacol. 38:95-9.

Di Carlo G, Mascolo N, Izzo AA, Capasso F (1999) Flavonoids: olds and new aspects of a class of natural therapeutic drugs. Life Sci. 65:337–353.

Dixon RA, Harrison MJ, Paiva NL (1995) The isoflavonoid phytoalexin pathway: from enzymes to genes to transcription factors. Physiol. Plant. 93:385–392.

Dodson E, Harding MM, Hodgkin DC, Rossmann MG (1966) The crystal structure of insulin.3. Evidence for a 2-fold axis in rhombohedral zinc insulin. J Mol Biol. 16:227-241.

Dorababu M, Prabha T, Priyambada S, Agarwal VK, Aryya NC, Goel RK (2004) Effect of *Bacopa monniera* and *Azadirachta indica* on gastric ulceration and healing in experimental NIDDM rats. Ind. J. Exp. Biol. 42:389-97.

Dugas AJ, Castaneda Acosta J, Bonin GC, Price KL, Fischer NH, Winston GW (2000) Evaluation of the total peroxyl radical-scavenging capacity of flavonoids: Structure-activity relationships. J. Nat. Prod. 63:327–331.

Elangovan V, *et al*. (1995) *In vitro* studies on the anticancer activity of *Bacopa monnieri*. Fitoterapia 66: 211-215.

Enz A, Amstutz R, Boddeke H *et al*. (1993) Brain selective inhibition of acetocholinesterase: a novel approach to therapy for Alzheimer's disease. Progress in Brain Res. 98:431-438.

Felsenstein J (1985) Confidence limits on phylogenies: An approach using the bootstrap. Evolution. 39:783-791.

Ford CM, Boss PK, Hoj PB (1998) Cloning and characterization of *Vitis vinifera* UDP-glucose:flavonoid 3-O glucosyltransferase, a homolog of the enzyme encoded by the maize Bronze-1 locus that may primarily serve to glucosylate anthocyanindins in vivo. J. Biol. Chem. 273:9224–9233.

Freeman WM, Walker SJ, Vrana KE (1999) Quantitative RT-PCR: pitfalls and potential. Biotechniques. 26:112-125.

Frydman A, Weisshaus O, Bar-Peled M, Huhman D V, Sumner L W, Marin F R, Lewinsohn E, Fluhr R, Gressel J, Eyal Y (2004) Citrus fruit bitter flavors: isolation and functional characterization of the gene Cm1, 2RhaT encoding a 1, 2 rhamnosyltransferase, a key enzyme in the biosynthesis of the bitter flavonoids of citrus. Plant J. 40:88–100.

Fujioka S, Yokota T (2003) Biosynthesis and metabolism of brassinosteroids. Plant Biol. 54:137–164.

Fukuchi-Mizutani M, Okuhara H, Fukui Y, Nakao M, Katsumoto Y, Yonekura-Sakakibara K, Kusumi T, Tanaka Y (2003) Biochemical and molecular characterization of a novel UDP-glucose:anthocyanin 39-O-glucosyltransferase, a key enzyme for blue anthocyanin biosynthesis, from *Gentian*. Plant Physiol. 132: 1652–1663.

Gachon CM, Langlois-Meurinne M, Saindrenan P (2005) Plant secondary metabolism glycosyltransferases: The emerging functional analysis. Trends Plant Sci. 10:542–549.

Gandia-Herrero F, Lorenz A, Larson T, Graham IA, Bowles DJ, Rylott EL, Bruce NC (2008) Detoxification of the explosive 2,4,6-trinitrotoluene in *Arabidopsis*: Discovery of bifunctional O- and C-glucosyltransferases. Plant J. 56:963–974.

Gang DR, Lavid N, Zubieta C, Chen F, Beuerle T, Lewinsohn E, Noel JP, Pichersky E (2002) Characterization of phenylpropene O-methyltransferases from sweet basil:

Facile change of substrate specificity and convergent evolution within a plant O-methyltransferase family. Plant Cell. 14:505–519.

Ganguly DK, Malhotra CL (1967) Some behavioural effects of an active fraction from *Herpestine monniera*, Linn. (Brahmi). Ind. J. Med. Res. 55:473-82.

Gao Z, Huang K, Yang X, Xu H (1999) Free radical scavenging and antioxidant activities of flavonoids extracted from the radix of *Scutellaria baicalensis* Georgi. Biochim. Biophys. Acta. 1472:643–650.

Garay S, Mahato SB, Ohtani K, Yamasaki K (1996a) Dammarane-type triterpenoid saponins from *Bacopa monniera*. Phytochemistry. 42:815–820.

Garay S, Mahato SB, Ohtani K, Yamaski K. (1996b) Bacosaponin D-a pseudojujubogenin glycoside from *Bacopa monniera*. Phytochemistry. 43:447–449.

Gershenzon J, Dudareva N (2007) The function of terpene natural products in the natural world. Nat. Chem. Biol. 3:408–414.

Goel RK, Sairam K, Babu MD, Tavares IA, Raman A (2003) *In vitro* evaluation of *Bacopa monniera* on anti-*Helicobacter pylori* activity and accumulation of prostaglandins. Phytomedicine. 10:523–527.

Goel RK, Sairam K. Antiulcer drugs from indigenous sources with emphasis on *Musa Sapientum, tambrabhasma, Asparagus racemosus* and *Zinzibar officinale*. Ind. J. Pharmacol. 2002;34:100-10.

Govindrajan R, Vijayakumar M, Pushpangadan P (2005) Antioxidant approach to disease management and the role of ‘Rasayana’ herbs. Ayur. J. Ethanopharmacol. 99:165-78.

Greenhagen BT, O’Maille PE, Noel JP, Chappell J (2006) Identifying and manipulating structural determinates linking catalytic specificities in terpene synthases. Proc. Natl. Acad. Sci. USA 103:9826–9831.

Guex N, Peitsch MC (1997). SWISS-MODEL and the Swiss-PdbViewer: An environment for comparative protein modeling. Electrophoresis. 18:2714-2723.

Hackman R (1998) Antioxidants that Entertain the Brain. Nutrition Science News. 10:530-538.

Harborne JB (1963) Plant polyphenols. X. Flavone and aurone glycosides of *Antirrhinum*. Phytochemistry. 2:327–334.

Harborne JB, Baxter H (1999) The Handbook of Natural Flavonoids, Vol. 2. (New York: John Wiley & Sons).

Harborne]B, Williams CA (2000) Advances in flavonoid research since 1992. Phytochem. 55:481-504.

He XZ, Wang X, Dixon RA (2006) Mutational analysis of the *Medicago* glycosyltransferase UGT71G1 reveals residues that control regioselectivity for (Iso)flavonoid glycosylation. J. Biol. Chem. 281:34441–34447.

Hengen P (1995) Purification of His-Tag fusion proteins from *Escherichia coli*. Trends in Biochemical Science. 20:285-286.

Hertog MGL, Hollman PCH (1996) Potential health effects of the dietary flavonol quercetin. Eur. J. Clin. Nutr. 50:63–71.

Hirotani M, Kuroda R, Suzuki H, Yoshikawa T (2000) Cloning and expression of UDP-glucose:flavonoid 7-O-glucosyltransferase from hairy root cultures of *Scutellaria baicalensis*. Planta. 210:1006–1013.

Hirotani M, Nagashima S, Yoshikawa T (1998) Baicalin and baicalein productions of cultured *Scutellaria baicalensis* cells. Nat. Med. 52:440–443.

Holcomb LA, Dhanasekaran M, Hitt AR, Young KA, Riggs M, Manyam BV (2006) *Bacopa monniera* extracts reduces amyloid levels in PSAPP mice. J Alzheimers Dis. 9:243-51.

Hou CC, Lin SJ, Cheng JT, Hsu FL (2002) Bacopaside III, bacopasaponin G, and bacopasides A, B, and C from *Bacopa monniera*. J. Nat. Prod. 65:1759–1763.

Hou B, Lim E K, Higgins G S, Bowles D J (2004). N-glucosylation of cytokinins by glycosyltransferases of *Arabidopsis thaliana*. J. Biol. Chem. 279:47822–47832.

Hu Y, Walker S (2002) Remarkable structural similarities between diverse glycosyltransferases. Chem. Biol. 9:1287–1296.

Hughes J, Hughes M A (1994). Multiple secondary plant product UDP–glucose glucosyltransferase genes expressed in *Cassava*. DNA Seq. 5: 41–49.

Huang Y, De Bruyne T, Apers S, Ma Y, Claeys M, Pieters L, Vlietinck A (1999) Flavonoid glucuronides from *Picriafel-terrae*. Phytochemistry. 52:1701–1703.

Hwang SB, Chang MN, Garcia ML, Han QQ, Huang L, King VF, Kaczorowski GJ, Winquist RJ (1987) L652, 469-a dual receptor antagonist of platelet activating factor and dihydropyridines from Tussilago farfara L. Eur. J. Pharmacol. 141:269–281.

Izzo AA, Ernst E (2001) Interactions between herbal medicines and prescribed drugs:a systematic review. Drugs. 61:2163–2175.

Jackson RG, Lim EK, Li Y, Kowalczyk M, Sandberg G, Hoggett J, Ashford D A, Bowles DJ (2001) Identification and biochemical characterization of an *Arabidopsis* indole-3-acetic acid glucosyltransferase. J. Biol. Chem. 276:4350–4356.

Jackson RG, Kowalczyk M, Li Y, Higgins G, Ross J, Sandberg G, Bowles D J (2002) Over-expression of an *Arabidopsis* gene encoding a glucosyltransferase of indole-3-acetic acid: phenotypic characterization of transgenic lines. Plant J. 32:573–58.

Jain P, *et al.* (1994) Antiinflammatory effects of an Ayurvedic preparation, Brahmi Rasayan, in rodents. Indian J. Expt. Biol. 32:633-636.

Jones P, Messner B, Nakajima J, Schaffner A R, Saito K (2003) UGT73C6 and UGT78D1, glycosyltransferases involved in flavonol glycoside biosynthesis in *Arabidopsis thaliana*. J. Biol. Chem. 278:43910–43918.

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

Juergenliemk G, Boje K, Huewel S, Lohmann C, Galla HJ, Nahrstedt A (2003) *In Vitro* studies indicate that miquelianin (quercetin 3-O-b-D-glucuronopyranoside) is able to reach the CNS from the small intestine. Plant. Med. 69:1013–1017.

Jyoti A, Sharma D (2006) Neuroprotective role of *Bacopa monniera* extract against aluminium-induced oxidative stress in the hippocampus of rat brain. Neuro. Toxicol. 27:451-457.

Kala CP (2002) Medicinal Plants of Indian Trans-Himalaya. Dehradun.

Kala CP, Mathur VB (2002) Patterns of plant species distribution in the trans-Himalayan region of Ladakh, India. Journal of Vegetation Science. 13:751-754.

Kala CP (2004) Revitalizing traditional herbal therapy by exploring medicinal plants: A case study of Uttaranchal State in India. In Indigenous Knowledges: Transforming the Academy, Proceedings of an International Conference Pennsylvania: Pennsylvania State University.15-21.

Kapoor KR, Srivastava SS, Kakkar P (2008) *Bacopa monnieri* modulates antioxidant responses in brain and kidney of diabetic rats. Environ. Toxicol. Pharmacol. (In Press).

Karim MR, Hashinaga F (2002) Preparation and properties of immobilized *Pummelo* limonoid glucosyltransferase. Process Biochem. 38:809–814.

Kar A, Panda S, Bharti S (2002) Relative efficacy of three medicinal plant extracts in the alteration of thyroid hormone concentrations in male mice. J Ethanopharmacol. 81:281-5.

Katsumoto Y, *et al*. (2007). Engineering of the rose flavonoid biosynthetic pathway successfully generated blue-hued flowers accumulating delphinidin. Plant Cell Physiol. 48:1589–1600.

Kawasaki M, Hayashi T, Arisawa M, Morita N, Berganza LH (1988) 8-Hydroxytricetin 7-glucuronide, a b-glucuronidase inhibitor from *Scoparia dulcis*. Phytochemistry. 27:3709–3711.

Kida A, Nishi K, Nagai H, Matsuura N, Tsuchiya H (1982) Anti-allergic actions of crude drugs and blended Chinese traditional medicines. Effects on type I and type IV. Yakugaku Zasshi. 80:31–41.

Kikuchi N, Kwon YD, Gotoh M, Narimatsu H (2003) Comparison of glycosyltransferase families using the profile hidden Markov model. Biochem Biophys Res Commun. 310:574-579.

Kim JH, Kim BG, Park Y, Ko JH, Lim CE, Lim J, Lim Y, Ahn JH (2006) Characterization of flavonoid 7-O-glucosyltransferase from *Arabidopsis thaliana*. Biosci. Biotechnol. Biochem. 70:1471–1477.

King CD, Rios GR, Green MD, Tephly TR (2000) UDP-glucuronosyltransferases. Curr. Drug Metab. 1:143–161.

Kleczkowski K, Schell J (1995) Phytohormone conjugates: nature and function. Plant Sci. 14:283–298.

Kliebenstein DJ (2008) A role for gene duplication and natural variation of gene expression in the evolution of metabolism. PLoS One. 3:e1839.

Knofel HD, Schwarzkopf E, Muller P, Sembdner G (1984) Enzymic glucosylation of gibberellins. J. Plant Growth Regul. 3:127–140.

Koeduka T, Louie GV, Orlova I, Kish CM, Ibdah M, Wilkerson CG, Bowman ME, Baiga TJ, Noel JP, Dudareva N, Pichersky E (2008) The multiple phenylpropene synthases in

both *Clarkia breweri* and *Petunia hybrida* represent two distinct protein lineages. Plant J. 54:362–374.

Kramer CM, Prata RTN, Willits MG, Luca VD, Steffens JC, Gracer G (2003) Cloning and regiospecificity studies of two flavonoid glucosyltransferases from *Allium cepa*. Phytochemistry. 64:1069–1076.

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

Kroemer HK, Klotz U (1992) Glucuronidation of drugs. Clin. Pharmacokinet. 23:292–310.

Kubo A, Arai Y, Nagashima S, Yoshikawa T (2004) Alteration of sugar donor specificities of plant glycosyltransferases by a single point mutation. Arch. Biochem. Biophys. 429:198–203.

Kulshreshtha DK, Rastogi RP (1973) Identification of ebelin lactone from Bacoside-A and the nature of its genuine sapogenin. Phytochemistry.12:2074–81.

Kulshreshtha DK, Rastogi RP (1973) Bacogenin A1: a novel dammerane triterpene sapogenin from *Bacopa monniera*. Phytochemistry. 12:887–892.

Kulshreshtha DK, Rastogi RP (1974) Bacogenin A2: a new sapogenin from bacosides. Phytochemistry. 13:1205–1206.

Kurosawa Y, Takahara H, Shiraiwa M (2002) UDP-glucuronic acid:soyasapogenol glucuronosyltransferase involved in saponin biosynthesis in germinating soybean seeds. Planta. 215:620–629.

Kyte J, Doolittle RF (1982) A simple method for displaying the hydropathic character of a protein. Journal of Molecular Biology. 157:105–132.

Laemmli UK (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature. 227:680–685.

Lanot A, Hodge D, Jackson RG, George GL, Elias L, Lim EK, Vaistij FE, Bowles DJ (2006) The glucosyltransferase UGT72E2 is responsible for monolignol 4-O-glucoside production in *Arabidopsis thaliana*. Plant J. 48:286–295.

Laskowski RA, McArthur MW, Moss DS, Thornton JM (1993) PROCHECK: aprogram to check the stereo-chemical quality of protein structures. J. Applied Cryst. 26:283-291.

Leatherbarrow RJ (1990) Using linear and non-linear regression to fit biochemical data. Trends Biochem. Sci. 15:455–458.

Lee C, O'Neill MA, Tsumuraya Y, Darvill AG, Ye Z (2007)The irregular xylem9 mutant is deficient in xylan xylosyltransferase activity. Plant Cell Physiol. 48:1624-1634.

Li Y, Baldauf S, Lim EK, Bowles DJ (2001) Phylogenetic analysis of the UDP-glycosyltransferase multigene family of *Arabidopsis thaliana*. J. Biol. Chem. 276:4338–4343.

Li P (2001) Purification of inclusion bodies and refolding of protein. Basic strong lab protocol. Immunogenetics. 53:279-287.

Li L, Modolo LV, Escamilla-Trevino LL, Achnine L, Dixon RA, Wang X (2007) Crystal structure of *Medicago truncatula* UGT85H2 – insights into the structural basis of a multifunctional (Iso)flavonoid glycosyltransferase. J. Mol. Biol. 370:951–963.

Lim EK, Li Y, Parr A, Jackson R, Ashford DA, Bowles DJ (2001) Identification of glucosyltransferase genes involved in sinapate metabolism and lignin synthesis in *Arabidopsis*. J. Biol. Chem. 276:4344–4349.

Lim EK, Doucet CJ, Li Y, Elias L, Worrall D, Spencer SP, Ross J, Bowles DJ (2002) The activity of *Arabidopsis* glycosyltransferases toward salicylic acid, 4-hydroxybenzoic acid, and other benzoates. J. Biol. Chem. 277:586–592.

Lim EK, Baldauf S, Li Y, Elias L, Worrall D, Spencer SP, Jackson RG, Taguchi G, Ross J, Bowles DJ (2003) Evolution of substrate recognition across a multigene family of glycosyltransferases in *Arabidopsis*. Glycobiology. 13:139–145.

Lim EK, Bowles DJ (2004) A class of plant glycosyltransferases involved in cellular homeostasis. EMBO J. 23:2915–2922.

Lim EK, Ashford DA, Hou B, Jackson RG, Bowles DJ (2004) *Arabidopsis* glycosyltransferases as biocatalysts in fermentation for resioselective synthesis of diverse quercetin glucosides. Biotechnol. Bioeng. 87:623–631.

Lim EK (2005) Plant glycosyltransferases: their potential as novel biocatalysts. Chem. Eur. J. 11:5486–5494.

Lim EK, Doucet CJ, Hou B, Jackson RG, Abrams SR, Bowles DJ (2005a) Resolution of (+)-abscisic acid using an *Arabidopsis* glycosyltransferase. Tetrahedron: Asymmetry. 16:143–147.

Lim EK, Jackson RG, Bowles DJ (2005b) Identification and characterisation of *Arabidopsis* glycosyltransferases capable of glucosylating coniferyl aldehyde and sinapyl aldehyde. FEBS Lett. 579:2802–2806.

Liu J, Mushegian A (2003) Three monophyletic superfamilies account for the majority of the known glycosyltransferases. Protein Sci. 12:1418-1431.

Lodhi MA, Ye GN, Norman FW, Bruce IR (1994) A simple and efficient method for DNA extraction from grapevine cultivars, *Vitis* species and Ampelopsis. Plant Molecular Biology Reporter. 12:6-13.

Loutre C, Dixon DP, Brazier M, Slater M, Cole DJ, Edwards R (2003) Isolation of a glucosyltransferase from *Arabidopsis thaliana* active in the metabolism of the persistent pollutant 3,4-dichloroaniline. Plant J. 34:485–493.

Macı´as FA, Galindo JL, Galindo JC (2007) Evolution and current status of ecological phytochemistry. Phytochemistry. 68:2917–2936.

Mackenzie P, Little JM, Radominska-Pandya A (2003) Glucosidation of hyodeoxycholic acid by UDP-glucuronosyltransferase 2B7. Biochem. Pharmacol. 65:417–421.

Mackenzie PI, *et al*. (1997) The UDP glycosyltransferase gene superfamily: Recommended nomenclature update based on evolutionary divergence. Pharmacogenetics. 7:255–269.

Maher BF, Stough C, Shelmerdine A, Wesnes K, Nathan PJ (2002) The acute effects of combined administration of *Ginkgo biloba* and *Bacopa monniera* on cognitive function in humans. Hum. Psychopharmacol. 17:163–164.

Malhotra CL, Das PK (1959) Pharmacological studies of *Herpestis monniera* Linn. (Brahmi). Indian J. Med. Res. 47:294–305.

Marchler-Bauer A, Anderson JB, Derbyshire MK, DeWeese-Scott C, Gonzales NR, Gwadz M, Hao L, He S, Hurwitz DI, Jackson JD, Ke Z, Krylov D, Lanczycki CJ, Liebert CA, Liu C, Lu F, Lu S, Marchler GH, Mullokandov M, Song JS, Thanki N, Yamashita RA, Yin JJ, Zhang D, Bryant SH (2007) A conserved domain database for interactive domain family analysis. Nucleic Acids Research. 35:237–240.

Martis G, Rao A, Karanth KS (1992) Neuropharmaological activity of *Herpestis monniera*. Fitoterapia. 63:399–404.

Mathur S, *et al*. (2002) Herb yield and bacoside A content of field grown *Bacopa monniera* accessions. J. Herbal, Spices and Medicinal Plants. 9:11-18.

Mathew KM (1984) The flora of Tamil Nadu and Carna. Rapinat Herbarium St. Joseph's College, Tiruchirapalli, India.

Matros A, Mock H P (2004) Ectopic expression of a UDP-glucose:phenylpropanoid glucosyltransferase leads to increased resistance of transgenic tobacco plants against infection with Potato Virus Y. Plant Cell Physiol. 45:1185–1193.

Messner B, Thulke O, Schaffner A R (2003) *Arabidopsis* glucosyltransferases with activities toward both endogenous and xenobiotic substrates. Planta. 217:138–146.

Miley MJ, Zielinska AK, Keenan JE, Bratton SM, Radominska-Pandya A, Redinbo MR (2007) Crystal structure of the cofactor-binding domain of the human phase II drug-metabolisi enzyme UDP-glucuronosyltransferase 2B7. J. Mol. Biol. 369:498–511.

Miller KD, Guyon V, Evans JN, Shuttleworth WA, Taylor LP (1999) Purification, cloning, and heterologous expression of a catalytically efficient flavonol 3-O-galactosyltransferase expressed in the male gametophyte of *Petunia hybrida*. J. Biol. Chem. 274:34011–34019.

Mok DWS, Mok MC (2001) Cytokinin metabolism and action. Annu Rev Plant Physiol Plant Mol. Biol. 52:89–118.

Moraga AR, Nohales PF, Perez JA, Gomez-Gomez L (2004) Glucosylation of the saffron apocarotenoid crocetin by a glucosyltransferase isolated from *Crocus sativus* stigmas. Planta. 219:955–966.

Morita Y, Hoshino A, Kikuchi Y, Okuhara H, Ono E, Tanaka Y, Fukui Y, Saito N, Nitasaka E, Noguchi H, Iida S (2005) Japanese morning glory dusky mutants displaying reddish-brown or purplish-gray flowers are deficient in a novel glycosylation enzyme for anthocyanin biosynthesis, UDP-glucose:anthocyanidin 3-O-glucoside-2''-O-glucosyltransferase, due to 4-bp insertions in the gene. Plant J. 42:353–363.

Mulichak AM, Losey HC, Lu W, Wawrzak Z, Walsh CT, Garavito RM (2003) Structure of the TDP-epi-vancosaminyltransferase GtfA from the chloroeremomycin biosynthetic pathway. Proc. Natl. Acad. Sci. USA. 100:9238-9243.

Mullis K, Faloona F, Scharf S, Saiki R, Horn G, Erlich H (1992) Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction. Biotechnology. 24:17-27.

Nadkarni KM (1976) Indian Materia Medica. Popular Prakashan Private, Bombay. 624–625.

Nagashima S, Hirotani M, Yoshikawa T (2000) Purification and characterization of UDP-glucuronate:baicalein 7-O-glucuronosyltransferase from *Scutellaria baicalensis* Georgi. cell suspension cultures. Phytochemistry. 53:533–538.

Nagashima S, Inagaki R, Kubo A, Hirotani M, Yoshikawa T (2004) cDNA cloning and expression of isoflavonoid-specific glucosyltransferase from *Glycyrrhiza echinata* cell-suspension cultures. Planta. 218:456–459.

Nathan PJ, Clarke J, Lloyd J, Hutchison CW, Downey L, Stough C (2001) The acute effect of an extract of *Bacopa monniera* (Brahmi) on cognitive function in healthy normal subjects. Hum. Psychopharmacol. 16:345–351.

Nathan PJ, Ricketts E, Wesnes K, Mrazek L, Greville W, Stough C (2002) The acute nootropic effects of an extract of *Gingko biloba* in healthy elderly human subjects: a preliminary investigation. Hum. Psychopharmacol. 17:45–49.

Nathan PJ, Tanner S, Lloyd J, Harrison B, Curran L, Oliver C *et al*. (2004) Effects of a combined extracts of *Ginko biloba* and *Bacopa monniera* on cognitive function in healthy humans. Human Psycopharmacology. 19:91-96.

Negi KS, Singh YD, Kushwaha KP, Rastogi CK, Rathi AK, Srivastava JS, *et al*. (2000) Clinical evaluation of memory enhancing properties of Memory plus in children with attention deficit hyperactivity disorder. Ind. J. Psychiatry. 42:Supplement.

Nishioka T, Kawabata J, Aoyama Y (1998) Baicalein, an α-glucosidase inhibitor from *Scutellaria baicalensis*. J. Nat. Prod. 61:1413–1415.

Noguchi A, Fukui Y, Iuchi-Okada A, Kakutani S, Satake H, Iwashita T, Nakao M, Umezawa T, Ono E (2008) Sequential glucosylation of a furofuran lignan, (+)-sesaminol by *Sesamum indicum* UGT71A9 and UGT94D1 glucosyltransferases. Plant J. 54:415–427.

Noguchi A, Saito A, Homma Y, Nakao M, Sasaki N, Nishino T, Takahashi S, Nakayama T (2007) A UDP-glucose:isoflavone 7-O-glucoyaltransferase from the roots of soybean (*Glycine max*) seedling. J. Biol. Chem. 282:23581–23590.

Ober D (2005) Seeing double: Gene duplication and diversification in plant secondary metabolism. Trends Plant Sci. 10:444–449.

Offen W, Martinez-Fleites C, Yang M, Lim EK, Davis BG, Tarling CA, Ford CM, Bowles DJ, Davies GJ (2006) Structure of a flavonoid glucosyltransferase reveals the basis for plant natural product modification. EMBO J. 25:1396–1405.

Ogata J, Kanno Y, Itoh Y, Tsugawa H, Suzuki M (2005) Anthocyanin biosynthesis in roses. Nature. 435:757–758.

O'Leary KA, Day AJ, Needs PW, Mellon FA, O'Brien NM, Williamson G (2003) Metabolism of Quercetin-7- and Quercetin-3-glucuronides by an in vitro hepatic model: The role of human β-glucuronidase, sulfotransferase, catechol-O-methyltransferase and multi-resistant protein 2 (MRP2) in flavonoid metabolism. Biochem. Pharmacol. 65:479–491.

O'Maille PE, Malone A, Dellas N, Andes Hess B, Jr. Smentek L, Sheehan I, Greenhagen BT, Chappell J, Manning G, Noel JP (2008) Quantitative exploration of the catalytic landscape separating divergent plant sesquiterpene synthases. Nat. Chem. Biol. 4:617–623.

Ono E, Fukuchi-Mizutani M, Nakamura N, Fukui Y, Yonekura-Sakakibara K, Yamaguchi M, Nakayama T, Tanaka T, Kusumi T, Tanaka Y (2006) Yellow flowers generated by expression of the aurone biosynthetic pathway. Proc. Natl. Acad. Sci. USA. 103:11075–11080.

Ono E, Nakayama T (2007) Molecular breeding of novel yellow flowers by engineering the aurone biosynthetic pathway. Transgen. Plant J. 1:66–80.

Ono E, et al. (2010) Functional differentiation of the glycosyltransferases that contribute to the chemical diversity of bioactive flavonol glycosides in grapevines (*Vitis vinifera*). The Plant Cell. 22:2856–2871.

Osmani S, Bak S, Imberty A, Olsen CE, Møller BL (2008) Catalytic key anmino acids and UDP-sugar donor specificity of a plant glucuronosyltransferase, UGT94B1. Plant Physiol. 148:1295–1308.

Osmani SA, Bak S, Møller BL (2009) Substrate specificity of plant UDP-dependent glycosyltransferases predicted from crystal structures and homology modeling. Phytochemistry. 70:325–347.

Ouzzine M, Gulberti S, Levoin N, Netter P, Magdalou J, Fournel-Gigleux S (2002) The donor substrate specificity of the human b1,3-glucuronosyltransferase I toward UDP-glucuronic acid is determined by two crucial histidine and arginine residues. J. Biol. Chem. 277:25439–25445.

Packer L (2000) Flavonoids and Other Polyphenols. Academic Press, Inc., Tokyo.

Paquette S, Moller B L, Bak S (2003) On the origin of family 1 plant glycosyltransferases. Phytochem. 62:399–413.

Pawar R, Gopalakrishnan C, Bhutani KK (2001) Dammarane triterpene saponin from *Bacopa monniera* as the superoxide inhibitor in polymorphonuclear cells. Planta Med. 67:752–754.

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

Pichersky E, Gang DR (2000) Genetics and biochemistry of secondary metabolites in plants: An evolutionary perspective. Trends Plant Sci. 5:439–445.

Pinot F, *et al*. (1997) Induction of stress proteins by tobacco smoke in human monocytes: modulation by antioxidants. Cell Stress Chaperones. 2:156-161.

Poppenberger B, Berthiller F, Lucyshyn D, Sieberer T, Schuhmacher R, Krska R, Kuchler K, Gloss J, Luschnig C, Adam G (2003) Detoxification of the fusarium mycotoxin deoxynivalenol by a UDPglucosyltransferase from *Arabidopsis thaliana*. J. Biol. Chem. 278:47905–47914.

Poppenberger B, Fujioka S, Soeno K, George G L, Vaistij F E, Hiranuma S, Seto H, Takatsuto S, Adam G, Yoshida S, Bowles D J (2005) The UGT73C5 of *Arabidopsis thaliana* glucosylates brassinosteroids. Proc. Natl. Acad. Sci. USA. 102:15253–15258.

Prajapati ND, Purohit SS, Sharma AK, Kumar T (2003) A Handbook of Medicinal Plants. Jodhpur: Agrobios.

Pravina K, *et al*. (2007) Safety evaluation of BacoMind^{TM} in healthy volunteers: A phase I study. Phytomedicine 14:301-308.

Quiel JA, Bender J (2003) Glucose conjugation of anthranilate by the *Arabidopsis* UGT74F2 glucosyltransferase is required for tryptophan mutant blue fluorescence. J. Biol. Chem. 278:6275–6281.

Rai D, Bhatia G, Palit G, Pal R, Singh S, Singh HK (2003) Adaptogenic effect of *Bacopa monniera* (Brahmi). Pharmacol. Biochem. Behav. 75:823–830.

Rao MR, Palada MC, Becker BN (2004) Medicinal and aromatic plants in agro-forestry systems. Agroforestry Systems. 61:107-122.

Rao CHV, Sairam K, Goel RK (2000) Experimental evaluation of *Bacopa monniera* on rat gastric ulceration and secretion. Indian J. Physiol. Pharmacol. 44:435–441.

Rastogi RP (1990) Compendium of Indian Medicinal Plants, vol. 1. CSIR, New Delhi, pp. 118–122.

Rastogi S, *et al*. (1994b) Bacoside A3-A triterpenoid saponin from *Bacopa monniera*. Phytochem. 36:133-137.

Richman A, Swanson A, Humphrey T, Chapman R, McGarvey B, Pocs R, Brandle J (2005) Functional genomics uncovers three glucosyltransferases involved in the synthesis of the major sweet glucosides of *Stevia rebaudiana*. Plant J. 41:56–67.

Rodo AP, Brugiere N, Vankova R, Malbeck J, Olson J M, Haines SC, Martin RC, Habben JE, Mok DWS, Mok MC (2008) Overexpression of a zeatin O-glucosylation gene in maize leads to growth retardation and tassel seed formation. Journal of Experimental Botany. 59:2673–2686.

Rohini G, Sabitha, KE., Devi, CS (2004) *Bacopa monniera* Linn. Extract modulates antioxidant and marker enzyme status in fibrosarcoma bearing rats. Ind. J. Exp. Biol. 42:776-80.

Roodenrys S, Booth D, Bulzoni S, Phipps A, Micallef C, Smoker J (2002) Chronic effects of Brahami (*Bacopa monnieri*) on human memory. Neuropsychopharmacology. 27:279–281.

Rosen ML, Edman M, Sjostrom M, Wieslander A (2004) Recognition of fold and sugar linkage for glycosyltransferases by multivariate sequence analysis. J. Biol. Chem. 279:38683-38692.

Ross J, Li Y, Lim E K, Bowles D J (2001) Higher plant glycosyltransferases. Genome Biol. 2:30041–30046.

Ross JA, Kasum CM (2002) Dietary flavonoids: Bioavailability, metabolic effects, and safety. Annu. Rev. Nutr. 22:19-34.

Russo A, Borrelli F (2005) *Bacopa monniera*, a reputed nootropic plant: an overview. Phytomed. 12:305-17.

Russo A, *et al*. (2003a) Nitric oxide-related toxicity in cultured astrocytes: effects of *Bacopa monniera*. Life Sci. 73:1517-1526.

Russo A, *et al*. (2003b) Free radical scavenging capacity and protective effects of *Bacopa monniera* on DNA damage. Phytotherapy Res. 17:870-875.

Saitou N, Nei M (1987) The neighbor-joining method: A new method for reconstructing phylogenetic trees. Molecular Biology and Evolution. 4:406-425.

Sairam K, Rao CV, Babu MD, Goel RK (2001) Prophylactic an curative effects of *Bacopa monniera* in gastric ulcer models. Phytomedicine. 8:423–430.

Sairam K, Dorababu M, Goel RK, Bhattacharya SK (2002) Antidepressant activity of standardized extract of *Bacopa monniera* in experimental models of depression in rats. Phytomedicine. 9:207–211.

Samant SS, Dhar U, Palni LMS (1998) Medicinal Plants of Indian Himalaya: Diversity Distribution Potential Values. Almora: G. B. Pant Institute of Himalayan Environment and Development.

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning: a laboratory manual, 2nd edn. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press.

Samiulla DS, Prashnath D, Amit A (2001) Mast cell stabilizing activity of *Bacopa monniera*. Fitoterapia. 72:284-285.

Sandrock RW, VanEtten HD (1998) Fungal sensitivity to and enzymatic degradation of the phytoanticipin alpha-tomatine. Phytopathol. 88:137–143.

Sara SJ (1989) Noradrenergic–cholinergic interaction: its possible role in memory dysfunction associated with senile dementia. Arch. Gerontol. Geriatr. 1S:99–108.

Sastri MS, Dhalla NS, Malhotra CL (1959) Chemical investigation of *Herpestis monniera* Linn (Brahmi). Indian J. Pharmacol. 21:303–304.

Satyavati GV, Raina MK, Sharma M (1976) Medicinal Plants of India, vol I. Indian Council of Medical Research, New Delhi. pp. 112–118.

Sawada S, Suzuki H, Ichimaida F, Yamaguchi MA, Iwashita T, Fukui Y, Hemmi H, Nishino T, Nakayama T (2005) UDPglucuronic acid:anthocyanin glucuronosyltransferase from red daisy (*Bellis perennis*) flowers. Enzymology and phylogenetics of a novel

glucuronosyltransferase involved in flower pigment biosynthesis. J. Biol. Chem. 280:899–906.

Schneider G, Schliemann W (1994) Gibberellin conjugates: an overview. Plant Growth Regul. 15:247–260.

Schulz M, Strack D, Weissenbo¨ ck G, Markham KR, Dellamonica G, Chopin J (1985) Two luteolin O-glucuronides from primary leaves of *Secale* cereal. Phytochemistry. 24:343–345.

Segel IH (1975) Enzyme Kinetics. Behavior and Analysis of Rapid Equilibrium and Steady-State Enzyme Systems. (New York: John Wiley & Sons).

Seiss H (1993) Strategies of antioxidant defense. Euro. J. Biochem. 215:213-9.

Seki H, Ohyama K, Sawai S, Mizutani M, Ohnishi T, Sudo H, Akashi T, Aoki T, Saito K, Muranaka T (2008) *Licorice* beta-amyrin 11-oxidase, a cytochrome P450 with a key role in the biosynthesis of the triterpene sweetener glycyrrhizin. Proc. Natl. Acad. Sci. USA. 105:14204–14209.

Senafi SB, Clarke DJ, Burchell B (1994) Investigation of the substrate specificity of a cloned expressed human bilirubin UDPglucuronosyltransferase:UDP-sugar specificity and involvement in steroid and xenobiotic glucuronidation. Biochem. J. 303:233–240.

Shao H, He X, Achnine L, Blount JW, Dixon RA, Wang X (2005) Crystal structures of a multifunctional triterpene/flavonoid glycosyltransferase from *Medicago truncatula*. Plant Cell. 17:3141–3154.

Shader RI, Greenblatt DJ (1995) Pharmacotherapy of acute anxiety. In Bloom FE, Kupfer DJ, editors. Psycopharmacology: Fourth generation of progress. New York: Raven Press. 1341-8.

Shanker G, Singh HK (2000) Anxiolytic profile of standardized Brahmi extract. Indian J. Pharmacol. 32:152.

Shanmugasundaram ER, Akbar GK, Shanmugasundaram KR (1991) Brahmighritham, an Ayurvedic herbal formula for the control of epilepsy. J. Ethnopharmacol. 33:269–276.

Sharma R, Chaturvedi C, Tewari PV (1987) Efficacy of *Bacopa monniera* in revitalizing intellectual functions in children. J. Res. Educ. Indian Med. 1:12.

Shiva MP (1996) Inventory of Forestry Resources for Sustainable Management and Biodiversity Conservation. New Delhi: Indus Publishing Company.

Shibata S, Hattori S (1931) U¨ ber die konstitution des baicalein und des wogonin. Yakugaku Zasshi. 51:15–17.

Shiono M, Matsugaki M, Takeda K (2005) Structure of the blue cornflower pigment. Nature. 436:791.

Shukia B, Khanna NK, Godhwani JL (1987) Effect of Brahmi Rasayan on the central nervous system. J. Ethnopharmacol. 21:65–74.

Sinha J, *et al*. (2002) Bacosaponin C.: critical evaluation of anti-leishmania properties in various delivery models. Drug Delivery 9:55-62.

Singh DK, Hajra PK (1996) Floristic diversity. In Biodiversity Status in the Himalaya New Delhi: British Council. 23-38.

Singh HK, Dhawan BN (1982) Effect of *Bacopa monniera* extract on avoidance responses in rat. J. Ethnopharmacol. 5:205–214.

Singh HK, *et al*. (1988) Effect of Bacoside A and B on avoidance responses in rats. Phytotherapy Res. 2:70-75.

Singh S, Eapen S, D'Souza SF (2006) Cadmium accumulation and its influence on lipid peroxidation and antioxidative system in an aquatic plant, *Bacopa monnieri* L. Chemosphere. 62:233-46.

Singh HK, Dhawan BN (1992) Drugs affecting learning and memory. In: Tandon, P.N., Bijiani, V., Wadhwa, S. (Eds.), Lectures in Neurobiology, vol. 1. Wiley Eastern, New Delhi, pp. 189–207.

Singh HK, Dhawan BN (1997) Neuropsychopharmacological effects of the Ayurvedic nootropic *Bacopa monniera* Linn (Brahmi). Indian J. Pharmachol. 29:S359–S365.

Singh RH, Singh L (1980) Studies on the anti-anxiety effect of the medyha rasayana drug Brahmi (*Bacopa monniera* Wettst.). Res. Ayur. Siddha. 1:133–148.

Singh HK, Srimal RC, Srivastava AK, Garg NK, Dhan BN (1990) Neuropsychopharmacological effects of bacosides A and B. Proceedings of the Fourth Conference on Neurobiology Learning Memory, Abstract No. 79. Irvine California.

Singh HK, Shanker G, Patnaik GK (1996) Neuropharmacological and anti-stress effects of bacosides: a memory enhancer. Indian J. Pharmacol. 28:47.

Sivarajan VV, Balachandran I (1994) Ayurvedic drugs and their plant sources, Oxford and IBH Publishing Co. New Delhi. Pp. 97-99.

Sivaramakrishna C, *et al*. (2005) Triterpenoid glycosides from *Bacopa monniera*. Phytochemistry. 66:2719-2728.

Steinmetz KA, Potter JD (1996) Vegetables, fruit, and cancer prevention: A review. J. Am. Diet. Assoc. 96:1027–1039.

Stough C, Lloyd J, Clarke J, Downey AL, Hutchison CW, Rodgers T, Nathan PJ (2001) The chronic effects of an extract of *Bacopa monniera* (Brahmi) on cognitive function in healthy human subjects. Psychopharmacology. 156:481–484.

Studier FW, Moffatt BA. 1986. Journal of Molecular Biology 189, 113.

Studier FW, Rosenberg AH, Dunn JJ, Dubendorff JW (1990) Methods in Enzymology 185:60–89.

Subramanian SS, Nair AGR (1973) Scutellarin and hispidulin-7-O-glucuronide from the leaves of *Clerodendrum indicum* and *Clerodendron infortunatum*. Phytochemistry. 12:1195.

Sumathy T, Subramanian S, Govidasamy S, Balakrishna K, Veluchamy G (2001) Protective role of *Bacopa monniera* on morphine induced hepatotoxicity in rats. Phytother. Res. 15:643–645.

Sumathy T, Govindasamy S, Balakrishna K, Veluchamy G (2002) Protective role of *Bacopa monniera* on morphineinduced brain mitochondrial enzyme activity in rats. Fitoterapia. 73:381–385.

Sumathy T, Nongbri A (2008) Hepatoprotective effects of Bacoside-A, a major constituent of *Bacopa monniera* Linn. Phytomedicine. Online.

Tripathi YB, Chaurasia S, Tripathi E, Upadhyay A, Dubey GP (1996) *Bacopa monniera* Linn. as an antioxidant:mechanism of action. Indian J. Exp. Biol. 34:523–526.

Taguchi G, Yazawa T, Hayashida N, Okazaki M (2001) Molecular cloning and heterologous expression of novel glucosyltransferases from tobacco cultured cells that have broad substrate specificity and are induced by salicylic acid and auxin. Eur. J. Biochem. 268:4086–4094.

Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Molecular Biology and Evolution. 24:1596-1599.

The French-Italian Public Consortium for Grapevine Genome Characterization (2007) The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla. Nature. 449:463–468.

Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. 22:4673–4680.

Tohge T, Nishiyama Y, Hirai M Y, Yano M, Nakajima J, Awazuhara M, Inoue E, Takahashi H, Goodenowe D B, Kitayama M, Noji M, Yamazaki M, Saito K (2005) Functional genomics by integrated analysis of metabolome and transcriptome of *Arabidopsis* plants over-expressing an MYB transcription factor. Plant J. 42:218–235.

Tomimori T, Miyaichi Y, Imoto Y, Kizu H, Tanabe Y (1984) Studies on the constituents of *Scutellaria* species III. On the flavonoid constituents of the root of *Scutellaria baicalensis* GEORGI. Yakugaku Zasshi. 104:524–528.

Unligil UM, Rini JM (2000) Glycosyltransferase structure and mechanism. Curr. Opin. Struct. Bio. l10:510-517.

Uphof JC (1968) The dictionary of economic plants, vol. 62. Verlag von J Cramer, NewYork.

Vijayan VA, Helen A (2007) Protective activity of *Bacopa monniera* Linn. On nicotine-induced toxicity in mice. Phytother. Res. 21:378-81.

Vogt T, Zimmermann E, Grimm R, Meyer M, Strack D (1997) Are the characteristics of betanidin glucosyltransferases from cell-suspension cultures of *Dorotheanthus bellidiformis* indicative of their phylogenetic relationship with flavonoid glucosyltransferases? Planta. 203:349–361.

Vogt T, Jones P (2000) Glycosyltransferases in plant natural product synthesis: Characterization of a supergene family. Trends Plant Sci. 5:380–386.

Vohora D, Pal SN, Pillai KK (2000) Protection from phenytoin-induced cognitive deficit by *Bacopa monniera*, a reputed nootropic plant. J. Ethnopharmacol. 71:383–390.

Von RU, Huttl R, Lottspeich F, Gierl A, Frey M (2001) Two glucosyltransferases are involved in detoxification of benzoxazinoids in maize. Plant J. 28:633–642.

Weis M, Lim E K, Bruce N C, Bowles D J (2008) Engineering and kinetic characterisation of two glucosyltransferases from *Arabidopsis thaliana*. Biochimie. 90:830–834.

Woodward AW, Bartel B (2005) Auxin: regulation, action, and interaction. Ann Bot. 95:707–735.

Wortley AH, Rudall PJ, Harris DJ, Scotland RW (2005) How much data are needed to resolve a difficult phylogeny? Case study in *Lamiales*. Syst. Biol. 54:697–709.

Xu ZJ, Nakajima M, Suzuki Y, Yamaguchi I (2002) Cloning and characterization of the abscisic acid-specific glucosyltransferase gene from adzuki bean seedlings. Plant Physiol. 129:1285–1295.

Yamazaki M, Gong Z, Fukuchi-Mizutani M, Fukui Y, Tanaka Y, Kusumi T, Saito K (1999) Molecular cloning and biochemical characterization of a novel anthocyanin 5-O-glucosyltransferase by mRNA differential display for plant forms regarding anthocyanin. J. Biol. Chem. 274:7405–7411.

Yamazaki M, Nakajima J, Yamanashi M, Sugiyama M, Makita Y, Springob K, Awazuhara M, Saito K (2003) Metabolomics and differential gene expression in anthocyanin chemo-varietal forms of *Perilla frutescens*. Phytochemistry. 62:987–995.

Yano S, Horiuchi H, Horie S, Aimi N, Sakai S, Watanabe K (1991) Ca2+ channel blocking effects of hirsutine, an indole alkaloid from *Uncaria* genus, in the isolated rat aorta. Planta Med. 57:403–405.

Yonekura-Sakakibara K, Tanaka Y, Fukuchi-Mizutani M, Fujiwara H, Fukui Y, Ashikari T, Murakami Y, Yamaguchi M, Kusumi T (2000) Molecular and biochemical characterization of a novel hydroxycinnamoyl-CoA: anthocyanin 3-O-glucoside-6”-Oacyltransferase from *Perilla frutescens*. Plant Cell Physiol. 41:495–502.

Yonekura-Sakakibara K, Tohge T, Matsuda F, Nakabayashi R, Takayama H, Niida R, Watanabe-Takahashi A, Inoue E, Saito K (2008) Comprehensive flavonol profiling and transcriptome coexpression analysis leading to decoding gene-metabolite correlations in *Arabidopsis.* Plant Cell. 20:2160–2176.

Yonekura-Sakakibara K, Tohge T, Niida R, Saito K (2007) Identification of a flavonol 7-O-rhamnosyltransferase gene determining flavonoid pattern in *Arabidopsis* by transcriptome coexpression analysis and reverse genetics. J. Biol. Chem. 282:14932–14941.

Yoshida K, Kameda K, Kondo T (1993) Digulucuronoflavones from purple leaves of *Perilla ocimoides*. Phytochemistry. 33:917–919.

Yoshikawa M, Morikawa T, Nakamura S, Li N, Li X, Matsuda H (2007) Bioactive saponins and glycosides. XXV. Acylated oleanane-type triterpene saponins from the seeds of tea plant (*Camellia sinensis*). Chem. Pharm. Bull. (Tokyo) 55:57–63.

Zakim D, Hochman Y, Kenny WC (1983) Evidence for an active site arginine in UDP-glucuronosyltransferase. J. Biol. Chem. 258:6430–6434.

Zhang L, Lin G, Kouacs B, Jani M, Krajcsi P, Zuo Z (2007) Mechanistic study on the intestinal absorption and disposition of baicalein. Eur. J. Pharm. Sci. 31:221–231.

Zhang YY, Guo YZ, Ageta H, Harigaya Y, Onda M, Hashimoto K, Ikeya Y, Okada M, Maruno M (1997) Studies on the constituents of roots of *Scutellaria planipes*. Planta Med. 63:536–539.

Zhang Z, Kochhar S, Grigorov M (2003) Exploring the sequence-structure protein landscape in the glycosyltransferase family. Protein Sci. 12**:**2291–2302.

Zubieta C, He XZ, Dixon RA, Noel JP (2001) Structures of two natural product methyltransferases reveal the basis for substrate specificity in plant O-methyltransferases. Nat. Struct. Biol. 8:271–279.

Zubieta C, Ross JR, Koscheski P, Yang Y, Pichersky E, Noel JP (2003) Structural basis for substrate recognition in the salicylic acid carboxyl methyltransferase family. Plant Cell. 15:1704–1716.

Zuckerkandl E, Pauling L (1965) Evolutionary divergence and convergence in proteins. In: Bryson V, Vogel HJ, eds. Evolving Genes and Proteins, New York: Academic Press. 97-166.