HETEROLOGOUS EXPRESSION, PURIFICATION AND BIOPHYSICAL CHARACTERIZATION OF RGP FROM GUAR

Ph.D. THESIS

by

SHILPI KUMARI

DEPARTMENT OF BIOTECHNOLOGY
INDIAN INSTITUTE OF TECHNOLOGY ROORKEE
ROORKEE – 247667 (INDIA)
DECEMBER, 2014

HETEROLOGOUS EXPRESSION, PURIFICATION AND BIOPHYSICAL CHARACTERIZATION OF RGP FROM GUAR

A THESIS

Submitted in partial fulfilment of the requirements for the award of the degree of

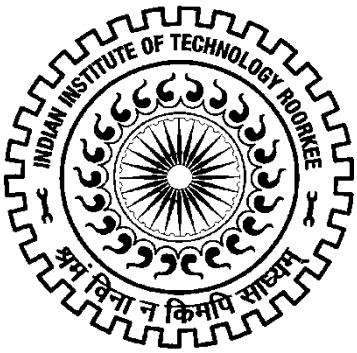
DOCTOR OF PHILOSOPHY

in

BIOTECHNOLOGY

by

SHILPI KUMARI



DEPARTMENT OF BIOTECHNOLOGY
INDIAN INSTITUTE OF TECHNOLOGY ROORKEE
ROORKEE – 247667 (INDIA)
DECEMBER, 2014

INDIAN INSTITUTE OF TECHNOLOGY ROORKEE
ROORKEE

CANDIDATE'S DECLARATION

I hereby certify that the work which is being presented in this thesis entitled, **"HETEROLOGOUS EXPRESSION, PURIFICATION AND BIOPHYSICAL CHARACTERIZATION OF RGP FROM GUAR"** in partial fulfilment of the requirements for the award of the Degree of Doctor of Philosophy and submitted in the Department of Biotechnology of the Indian Institute of Technology Roorkee is an authentic record of my own work carried out during a period from December, 2009 to December, 2014 under the supervision of Dr. G. S. Randhawa, Professor, Department of Biotechnology and Dr. S. K. Tripathi, Professor, Department of Water Resources Development and Management, Indian Institute of Technology Roorkee.

The matter presented in this thesis has not been submitted by me for the award of any other degree of this or any other institute.

(SHILPI KUMARI)

This is to certify that the above statement made by the candidate is correct to the best of our knowledge.

(S. K. Tripathi) (G. S. Randhawa)

Supervisor Supervisor

Date:

ABSTRACT

Molecular sub-cloning and heterologous expression of the cDNA of reversibly glycosylated protein from *Cyamopsis tetragonoloba* (CtRGP) was carried out in pET expression vector (pET 29a (+) and the sub-cloned DNA was over-expressed in *E. coli* BL21 (DE3). The recombinant protein purification was carried out by using immobilized metal ion chromatography (IMAC). Molecular weight of purified protein was found to be approximately 42 kDa. The purification was further confirmed by western blotting with anti-RGP antibody. Biophysical characterization of *Ct*RGP was done by analytical gel filtration, MALDI-TOF/MS, peptide mass fingerprinting, circular dichroism (CD), fluorescence spectroscopy (FS) and isothermal titration calorimetry (ITC). CtRGP protein was found to be in tetrameric form. The molecular mass of CtRGP protein was found to be 42,123.21Da. Peptide mass fingerprinting (PMF) analysis showed a total of 20 hits and most of these proteins were the enzymes involved in mutase activity for plant cell wall synthesis. The CD spectra of native *Ct*RGP revealed the presence of 32.45% α-helix and 12.92% β-sheet. Thermal and chemical denaturation of *Ct*RGP protein showed high thermostability and chemostability. The fluorescence study indicated that the tryptophan residue is located in the hydrophobic core of the protein. The ITC studies with the substrate UDP-glucose and inhibitors diisothiocyanostilbenedisulfonic acid (DIDS) and dithiobisnitrobenzoic acid (DTNB) showed that purified CtRGP protein has a single binding site for its substrate and inhibitors. The protein-ligands energetics revealed that the CtRGP protein has a strong, moderate and weak binding affinities with inhibitor DIDS, substrate UDP-glucose and inhibitor DTNB, respectively. *In silico* results showed that CtRGP is a stable protein and it belongs to the RGP superfamily. The secondary structure prediction of the protein revealed the presence of high helical content. Docking results showed that substrate UDP-glucose interacts with the loop section of the CtRGP protein structure.

ACKNOWLEDGMENTS

Completing a PhD is truly a marathon event, and I would not have been able to complete this journey without the aid and support of countless people over the past five years. I would like to extend my gratitude to the many people who helped to bring this research project to fruition. It gives me immense pleasure in acknowledging all the technical help that I have received during the period of research.

I must first express my deep sense of gratitude towards my advisor, Professor G. S. Randhawa, Department of Biotechnology, Indian Institute of Technology, Roorkee, for his guidance, patience and encouraging support, without his persistent help this thesis would not have been possible. Prof. Randhawa is not only a teacher and guide to me, but my mentor and well–wisher too. His leadership, support, attention to detail and hard work, have set an example. I hope to match some day. His persistent encouragement, perpetual motivation, everlasting patience, constructive criticism and valuable technical inputs in research have benefitted me to an extent which is beyond expression. He has not only trained me in science but in all aspects of life. I would like to take this opportunity to also thank Dr. Mrs. Surinder Randhawa who has been supportive in all the efforts during my research.

I would also like to thank Prof. S. K. Tripathi, Department of Water Resource Development and Management, Indian Institute of Technology, Roorkee, for his constant support and guidance throughout the period of research.

It gives me immense pleasure to acknowledge my heartiest thanks to Dr. Pravindra Kumar, Department of Biotechnology, Indian Institute of Technology, Roorkee, for giving me the exciting opportunity to be a member of his group and giving me a place in his lab where the equipment and some resources have made this project possible. Furthermore, his door has always been open and his continuous enthusiasm and involvement in this project has been greatly appreciated. Without their valuable guidance and constant support throughout the period of research the completion of this thesis would have been impossible.

I would like to thanks Prof. Ritu. Barthwal, Department of Biotechnology, Indian Institute of Technology, Roorkee, for allowing me to use the NMR facility.

I would like to my warm thanks to Prof. Partha Roy, Department of Biotechnology, Indian Institute of Technology, Roorkee for his constant help and support in my research work.

I wish to acknowledge my deep sense of gratitude for Prof. R. Prasad. HOD, Department of Biotechnology, Indian Institute of Technology, Roorkee, for providing me all the necessary resources and lab facility.

I would like to give warm compression of thanks to Dr. Kanwarpal S. Dhugga, and Dr. Rajeev Gupta, Scientists, Crop Genetics and research and Development, Pioneer Hi-Bred International, Johnston, USA for their valuable help and guidance. Without their contribution the completion of this thesis would have been impossible. I express my deep sense of gratitude to them for sparing time from their busy schedule. I would also like to thank Pioneer Hi-Bred International Company for the materials provided for my research work.

I would like to thank Dr. Shivendra Singh for his constant help. He deserves all the credit for training me in cloning, expression, purification and allowing me to try my hand at crystallography with protein. I would also thanks to Dr. Aditya and Dr. Pramod for guidance.

I take this opportunity to extend my gratitude to my seniors Dr. Durga Prasad Panigrahi, Dr. Nagesh Kuravadi and Dr. Pranita Bhatele for their useful advices, encouragement, guidance and goodwill which enabled me successfully complete my work.

I would like to special thanks Pallavi for beginning of my research work with her. I would also like to thank my labmates Umesh Tanwar, Manisha Choudhary, Swati Verma, Shalini Pareek, , Navneet Kaur, Deepa Dewan, Nishu Mittal and Poonam Jaiswal for assisting me at every stage of work and maintaining a friendly environment in the lab. I express my gratitude to University Grant Commission (UGC) Government of India for the financial support in the form of JRF and SRF during the period of this research.

I expressed my unbound gratitude to my parents Mrs. Punam Prasad and Late Lalit Mohan Prasad whose constant inspiration and support have helped me to attain the academic position which I am presently at. I would also like to warm thanks to my loving husband Mr. Ashok Kumar, who has stood beside me in thick and thin, and is the man behind my success. Very special thanks go to my sweet son Hridyansh. I also express my gratitude to all my family members my elder sister Mrs. Anamika, yonger sisters Celesty, Indu and younger brother Aakash for their constant support and encouragement.

Finally I thank the Almighty God for leading me all the way towards successful completion of this work.

Date: *Shilpi Kumari*

Table of contents **Page No.**

LIST OF ABBREVIATIONS USED

˚C	Degree centigrade
ΔG	Change in Gibb's free energy
ΔS	Change in entropy
ΔH	Change in enthalpy
µl	Microlitre
µm	Micrometer
APS	Ammonium per sulphate
BLAST	Basic Local Alignment Search Tool
bp	Base pair
BSA	Bovine Serum Albumin
CASP	Critical assessment of structure prediction
CD	Circular Dichroism
cDNA	Complementary DNA
cm	Centimeter
Csl	Cellulose synthase like
CtRGP	Cyamopsis tetragonoloba Reversibly Glycosylated Polypeptide
DIDS	Diisothiocyanostilbenedisulphonic acid
DNA	Deoxyribose nucleic acid
DNase	Deoxyribose nuclease
dNTPs	Deoxy nucleotide tri phosphates
DTNB	Dithiobis nitrobenzoic acid
E.coli	Escherichia coli
EDTA	Ethylenediaminetetracetic acid
eg.	Example
EMBL	European molecular biology lab
ESNP	Electronic SNP
EST	Expressed sequence tags
et al.	et alia
FASTA	Fast all
FOR	Forward
Fru	Fructose
g	gram
G/Gal	Galactose
GB	Giga bases
GDP-MP	GDP-mannose pyrophosphorylase
GDP-MS	GDP mannan dependent-mannosyl transferase
GenALEx	Genetic AnaLysis in Excel
Glu	Glucose
GMGT	Galactomannan galactosyltransferase
GRAVY	Grand Average Hydropathicity
GT	Glycosyl transferase
HCCA	Alpha-cyano-4-hydroxy cinamic acid
Hr	Hour
IMAC	Immobilized Metal Affinity Chromatography
IPTG	Isopropyl-Beta-D-Thiogalactoside
ITC	Isothermal Titration Calorimetry

KB	Kilo bases
K_D	Dissociation constant
kDa	Kilo Daltons
L.	Linnaeus
LB	Luria-Bertani
m	meter
M/G ratio	Mannose/Galactose ratio
M/Man	Mannose
MALDI-ToF	Mass Adsorption Laser Desorption Ionization- Time Of Flight
ManS	Mannan synthase
MASCOT	Modular Approach to Software Construction Operation and Test
mg	Milligram
$MgCl_2$	Magnesium Chloride
min	minute
ml	Millilitre
mm	Milli meter
mM	Millimolar
mRNA	messenger RNA
MS	Mass Spectrometry
MS	Mannan synthase
n	Binding stoichiometry/ stoichiometry of the interaction
Nacl	Sodium chloride
NCBI	National Center for Biotechnology Information
ng	Nanogram
nm	Nanometer
NTA	Nitrilotriacetic acid
O.D.	Optical Density
ORF	Open Reading Frame
PAGE	Poly Acrylamide Gel Electrophoresis
PCR	Polymerase chain reaction
PDB	Pritein DataBase
PEG	Poly Ethylene Glycol
PHGG	Partially hydrolysed guar gum
pM	Picomolar
PMF	Peptide Mass Fingerprinting
PMI	Phosphomanno isomerase
PMM	Phosphomanno mutase
R	puRine
REV	Reverse
RGP	Reversibly Glycosylated Polypeptide
RNA	Ribose nucleic acid
RNaseA	RibonucleaseA
rpm	Revolutions per minute
SAVES	Structure Analysis and Verification Server
SDS	Sodium Dodecyl Sulphate
SEC	Size Exclusion Chromatography
sec	Second
Taq	*Thermus aquaticus*
TBE	Tris borate EDTA
TBST	Tris-Buffered Saline and Tween 20
$TE/T_{10}E_1$	Tris EDTA
TEMED	Tetramethylethylenediamine

Tris	*tris*-(hydroxymethyl)-aminomethane
U	Unit
UDP	Uridine Di Phosphate
UDP-GE	UDP- galactose 4-epimerase
UDP-Glc	UDP-Glucose
UPGMA	Unweighted Pair-Group Average
USA	United States of America
UV	Ultraviolet
V	volt
V_e	Elution volume
V_o	Void volume
w/v	Weight/volume
Y	pYrimidine
α	Alpha
β	Beta

1. INTRODUCTION

Guar (*Cyamopsis tetragonoloba* [L.] Taub.), is commonly known as annual legume crop belonging to family Fabaceae. It has been used as green manure, vegetable and fodder since ancient times. Recently guar has attained the status of important industrial crop due to the gum content in the seeds of its endosperm. Guar gum is synthesized in the endosperm part of the developing seeds [111]. Guar gum is soluble in cold water and forms a high viscous gel. It is used in a wide range industrial applications as thickening, emulsifying and stabilizing agent [146]. Guar gum is a galactomannan, composed of a linear chain of β-D-mannopyranosyl units linked (1→4) with α-D-galactopyranosyl units (1→6) as side branchess in a ratio of approximately 2:1. Galactomannans are synthesized in Golgi lumen by the combined process of two enzymes: mannan synthase (ManS) and α-galactosyltransferase (GMGT) [113].

All plant cells are surrounded by a strong layer of cell wall, comprising of complex carbohydrates and glycoprotein. Plant cell wall carbohydrates are the most abundant biopolymers on Earth. These are composed of such cellulose, hemicelluloses, and pectins, structural polysaccharides. The cell wall is important for the shape, growth, development and mechanical properties of plant [99]. Plant cell wall is useful for human and animals nutrition in the form of dietary fiber. It is also valuable for industrial uses such as renewable biofuels generation.

The cell wall polysaccharide is divided in two main fractions, cellulose and matrix polysaccharides. The synthesis of cellulose occurs at the plasma membrane whereas the matrix polysaccharides are synthesized in the Golgi apparatus and transported to the cell wall by exocytosis [129]. The synthesis, assembling and integration of cellulose and the non-cellulosic polysaccharides occurs with the synthesis of lignin [23]. The most important functional properties of the plant Golgi apparatus is its ability to synthesize complex matrix

polysaccharides of the cell wall. The cell wall matrix polysaccharides (pectins and hemicelluloses) are assembled in the Golgi-cisternae and then transported to the cell surface within Golgi-derived vesicles [22, 99]. In the case of eukaryotic cell, the processing, sorting and transport of proteins to intra- and extra- cellular matrix is done by Golgi body. Another important function of Golgi apparatus in plants is assembling and exporting the non-cellulosic polysaccharides of the cell wall matrix, which in turn plays an important role in plant protection and development. [47].

In dicots, twenty percent of the primary cell wall is composed of xyloglucan. Xylosyl, xylosyl-galactosyl or fucosyl–galactosyl–xylosyl are residues side chain attached to β-1,4-glucan backbone in xyloglucan [28]. The main enzymes involved in xyloglucan biosynthesis are classified into different groups which is based on donor, acceptor and linkage [167]. The side chains are added to xyloglucan by specific transferases such as xylosyltransferase, fucosyltransferase and galactosyltransferase [143].

Reversibly glycosylated protein (RGP) is plant specific proteins which have ability to perform auto-glycosylation. These are cytosolic proteins but have a tendency to get associated with Golgi membranes. The first report of RGPs was the detection of a reversibly glycosylatable 40 kDa doublet protein in pea membranes by labeling it with radiolabeled UDP sugars as substrate [44]. Subsequently this protein was purified from pea tissue and named as RGP1 [43]. Immunogold labeling localized this protein specifically to trans-Golgi dictyosomal cisternae which suggests that xyloglucan and possibly other hemicelluloses are synthesized because of the involvement of RGP1. The isolation of a cDNA clone encoding RGP1 and the generation of antiRGP1 antibody have been carried out by Dhugga and coworker [43]. Some other RGP orthologs have been identified in *Arabidopsis thaliana* [37], cotton [168], maize [138, 148], Potato[14], wheat and rice [95], and tomato [144].

The expression study on cotton RGP gene indicates the role of RGP in non-cellulosic cell wall polysaccharide biosynthesis [168]. Interaction studies of tomato RGP with tomato leaf curl virus V1 protein point out the possible role of RGP in plant defence responses[144]. Recently, RGP proteins of *Arabidopsis* have been shown to act as mutases which interconvert pyranose and furanose forms of arabinose and this interconversion has been found to be essential for plant development and cell wall establishment [127]. In rice UDP-arabinopyranose mutase, an arginyl residue was found to be essential for catalytic activity and autoglycosylation [83, 84]. A native StRGP from potato tuber is connected with membranes as an oligomeric protein [14, 108]. Incubation at a certain specific ionic strength favours the formation of high molecular weight containing RGP forms [34]. The biochemical approaches have been used to study OsRGP oligomerization during developmental condition [35]. RGP may be a multifunctional protein involved in several unrelated activities [40]. Despite the above reports, the exact biological function (s) of RGPs still remains a mystery.

The biophysical characterization of RGP is expected to pave the way to crystal structure determination of this protein. The knowledge about the structure of RGP would be very helpful in knowing its exact biological role. Purification of RGP has been reported from tissues of pea [44], maize [139], *Arabidopsis thaliana* [46], rice [83-85] and potato [34]. RGP has also been purified from *E.coli* cells in which cDNA encoding RGP proteins of *A. thaliana* [37, 127, 166], cotton [169] and wheat and rice [95] were expressed. Despite the purification of RGP by several workers, reports are not available on biophysical characterization of this protein. This may be due to failure in obtaining highly stable and required quantity of pure RGP protein.

With the above information in view, the current research work was focused on the biophysical characterization of reversibly glycosylated polypeptide in guar. The biophysical

characterization and the 3D structure analysis together revealed valuable information about *Ct*RGP protein.

The current research work was carried out with following objectives:

1. Molecular sub-cloning, heterologous expression and purification of cDNA encoding RGP in guar
2. Biophysical characterization of guar RGP
3. Crystallization trial of guar RGP
4. Structural analysis, 3D Modeling and docking of guar RGP

2. REVIEW OF LITERATURE

The literature has been reviewed under the various headings.

2.1. Cultivation and guar production

Cluster bean (*Cyamopsis tetragonoloba* [L.] Taub.), commonly known as guar, is an annual legume crop. The common name guar is derived from Sanskrit word 'gau ahaar' which means the fodder of cow. Guar belongs to the family Fabaceae. Earlier, it was mainly used as a forage, green manure and vegetable crop, but now it is an important industrial crop due to galactomannan gum contained in the endosperm of its seeds [91]. It is an important leguminous herb, highly adapted to arid and semi-arid parts of the world requiring low inputs and care [116]. Guar is mainly cultivated in dry tracts of Rajasthan, Gujrat, Punjab, Madhya Pradesh and Uttar Pradesh in India. It is also cultivated in some part of Pakistan and has been introduced as a cash crop in Oklahoma and Texas states of USA [5, 125, 146]. It has been introduce as a new crop for western agriculture practices in many countries including Italy, Germany, Spain and Greece [5].

Mutagenesis is an important tool for favourable genetic variability in guar [6]. Some regeneration work in wild species of guar has been done for the improved varieties [1, 103] . For improvement and molecular breeding of cluster bean can be possible through hybridization for obtaining the desired traits [117, 124]. Genetic diversity among several genotype of cluster bean have been collected from different geographical regions of India for crop improvement and conservation [118]. India is the largest producer of guar in the world as total production of guar bean in India is estimated to have crossed 2.7 million metric tons during the year 2013-14. Presently, India accounts for more than three-fourth (or nearly 80 percent) of the total world guar production [10] (Fig. 2.1)

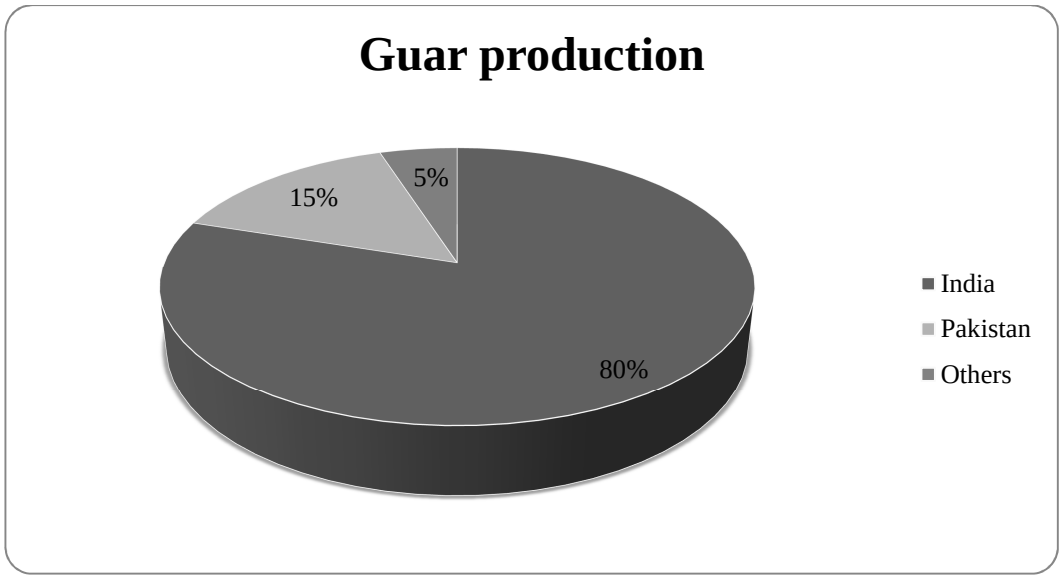


Fig. 2.1 India's share in world guar production in 2013-14 [10].

Guar is a self-pollinated diploid plant with chromosome numbers 2n=14 [2]. Guar plant has single stem, fine branching, or basal branching, depending on the variety, and height to be 18 to 40 inch tall. It has opposite leaves, and small flowers purple, pink to white in colour. Guar plant may produce around 30-90 pods per plant. Pods are generally 1½ to 4 inch long and contain 5 to 12 seeds each [10]. Fig. 2.2 shows the green pods and dried seeds of guar plant. Green pods of guar are a good source of vitamin A and C, calcium, irons and phosphorous. Pods of guar are also used in traditional medicine for controlling constipation , body pains, arthritis, anorexia and diabetes [66]. The guar seed is comprised of three parts: the endosperm (35-42%), embryo of the seed (43-47%) and the seed coat (14-17%). Seeds vary in color and size is about 4 mm [10]. Guar seeds have a large endosperm that contains 80-85% of guar gum/galactomannan. The remaining portion is composed of water, proteins, enzymes, and several other nonpolymer carbohydrates. Guar gum has ability to form hydrogen bond with water molecules which makes it industrially important as thickener and stabilizer [111]. Guar gum is extensively used in foods, cosmetics, textile, paper, mining, petroleum, explosive and pharmaceutical industries [124]. Apart from this guar gum is also useful in the control of many diseases like bowel syndrome, diabetes, heart disease and colon cancer [9, 88].

(a) (b)

Fig. 2.2 (a) Guar green pods grown in a field at IIT Roorkee and (b) Petri plates containing dried seeds of guar.

2.2. Historical background

The early history of old world legume guar is unknown, however established records and circumstantial evidence reveal that guar has been cultivated in Indo-Pakistan subcontinent for numerous generations. Cluster bean is considered to be originated by domestication of the African wild species, *C. senegalensis* which appears to be the ancestor of the *C. tetragonoloba.* The latter species was originally brought to South Asian subcontinent from Africa by Arab traders as fodder for horses probably during the 9th and 13th centuries A.D. [111] The domestication process may have emerged in the dry areas of the northwestern region of the Indo-Pakistan Subcontinent. It has been cultivated as a minor crop in India since ancient times as a vegetable and feed for cattle [77].

Guar was introduced in South western part of USA for the first time in 1903, where the hot climate with long growing seasons suits its adaptation [78, 159]. Carob (locust bean) seed (*Ceratonia siliqua*) from Mediterranean was used to extract carob gum which was extensively used in before the World War II. During World War II the supply of imported carob seeds from the Mediterranean region was cut off, so the Institute of Paper Chemistry,

USA initiated the search for domestic source of galactomannan gum. As a result guar was found to be the alternative source for gum [4]. For the first time guar was investigated as a source of gum in 1945 [159].

2.3 Guar gum

Guar gum or galactomannan is a heterogeneous polysaccharide and belongs to the hemicellulosic cell wall component of the plant cell. The galactomannans assume the role of storage polysaccharides in seeds of some plants, which is functionally similar to that of starch in cereal grains [41]. Galactomannans also play important role in water-retention and defense functions in plants [53]. Guar gum is accumulated in the form of secondary wall thickenings in the endosperm of guar seeds. Commercial guar gum is obtained by separating the endosperm from the germ by milling process. The final milled endosperm, which is commercial gum, contains 70-80% galactomannan, 2.5% of crude fiber, 10-15% of moisture, 5-6% of protein and 0.5-0.8% of ash [31]. The gum is further purified by ethanol precipitation for specific purpose. Figure 2.3 shows the flowchart of guar gum processing [111].

Guar gum is a high molecular weight hyrocolloidal polysaccharide, composed of galactan and mannan units linked through glycosidic linkages [76]. It is white to yellowish-white in color, nearly odorless, free-flowing powder[165][165][165][165][165][165][165]. The gum is soluble in cold or hot water and stable in a broad range of pH from 4 - 10, but insoluble in most organic solvents [164]. At low concentration it hydrates rapidly in cold water to attain uniform and very high viscosity [111]. The high viscosity is due to its long chain structure and high molecular weight. It has strong hydrogen bonding, excellent emulsion, film forming, stabilizing and thickening properties. A number of different gum derivatives due to presence of many hydroxyl groups made for specific applications [86]. Guar gum has high nutritional value which makes it an important natural food supplement. It

plays indispensable role in lowering serum cholesterol and glucose levels, and also considered helpful in weight loss programs [21]. Guar gum is extensively used in various industries because of their natural abundance, low cost and other desirable properties [110, 150].

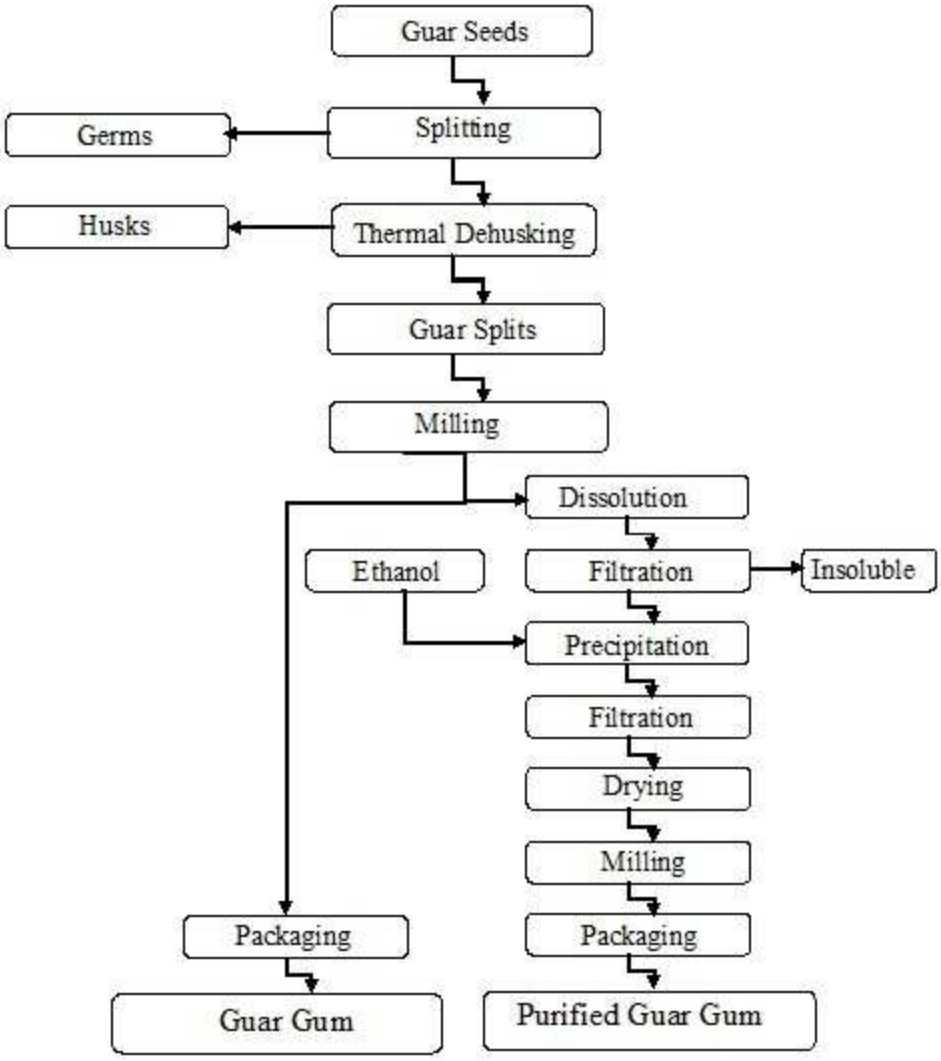


Fig. 2.3 Flowchart of extraction of guar gum [111]

2.3.1 Structure of guar gum

Guar gum has high molecular weight (100-1000 kDa) and a natural non-ionic galactomannan [112]. A guar gum molecule is composed of approximately 10,000 residues, which are polydisperse rod-shaped polymers, non-ionic in nature. The guar gum structure is a linear chain of β-D-mannopyranosyl units linked (1→4) with single member α-D-

galactopyranosyl units (1→6) occurring as side branches [102] (Fig. 2.4). Guar hydroxyls are situated at the cis position that plays a key role in reinforcing each other in hydrogen bonding. For the synthesis of galactomannan, a linear chain of β-l,4-mannan act as a backbone to which galactosyl residues are attached through -1,6 linkages [104]. The degree of galactose (G) substitution to the mannan (M) backbone differs depending upon the source of the gum M/G ratio; like M/G ratio for guar gum is 1.3-2 and for locust bean gum is 3.75-4. The ratio of mannose to galactose units ranges from 1.6 to 1.8 [153]. Figure 2.4 shows the general structure of galactomannans. The change in M/G ratio changes the structure of the gum thereby determining various properties and applications of the galactomannans [151]

Guar gum hydrates readily in aqueous solutions, but the other factors like solution clarity, thermal stability and alcohol solubility led to the production of guar gum with chemical changes. Three hydroxyl groups are available for derivatization on D-galactose or D-mannose sugar units. The molar substitutions (MS) in guar gum can exceed 3 because of the additional availability of hydroxyl groups [87].

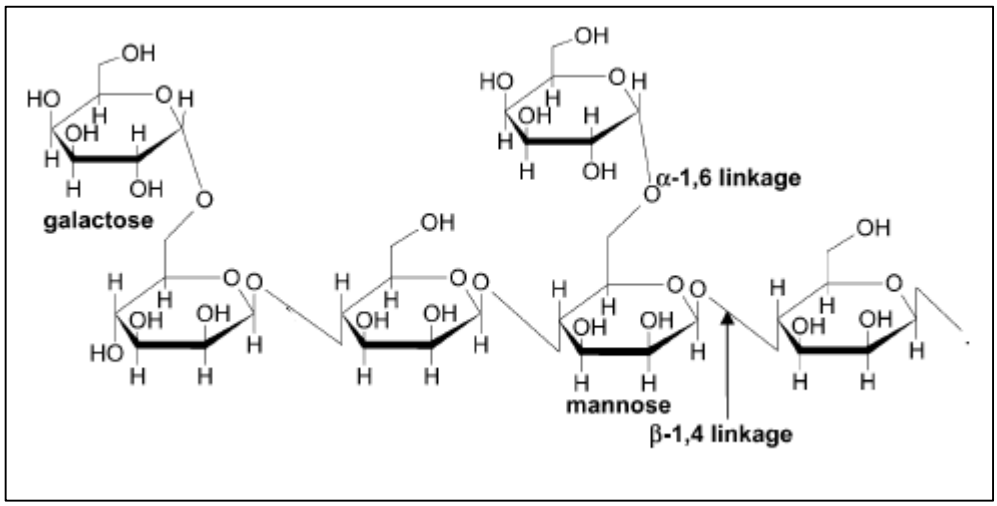


Figure 2.4: General chemical structure of guar galactomannan [102]

Homo-polymers made of long chains of mannose is insoluble in cold water and the increase in galactosylation leads to an increase in solubility [19]. Locust bean gum with M/G ratio of 4 is considered superior in its viscosity characteristics in comparison to guar gum.

Although it is superior in quality yet difficult to produce [20]. Different galactomannans differ from each other in distribution pattern of the galactose residues along the main chain. It is found that in guar galactomannan the distribution is blockwise while in locust bean gum the distribution may be blockwise, random, or ordered [123]. Fig 2.5 shows different kinds of distribution of galactose residues in galactomannans. In recent studies the degree of polymerization and the degree of substitution in galactomannan was studied. Unmodified guar galactomannan has been reported to have aggregates, but this tendency was found to be significantly decreased after partial removal of galactose units. The chain length in combination with the degree of substitution is believed to affect the solubility of molecules.

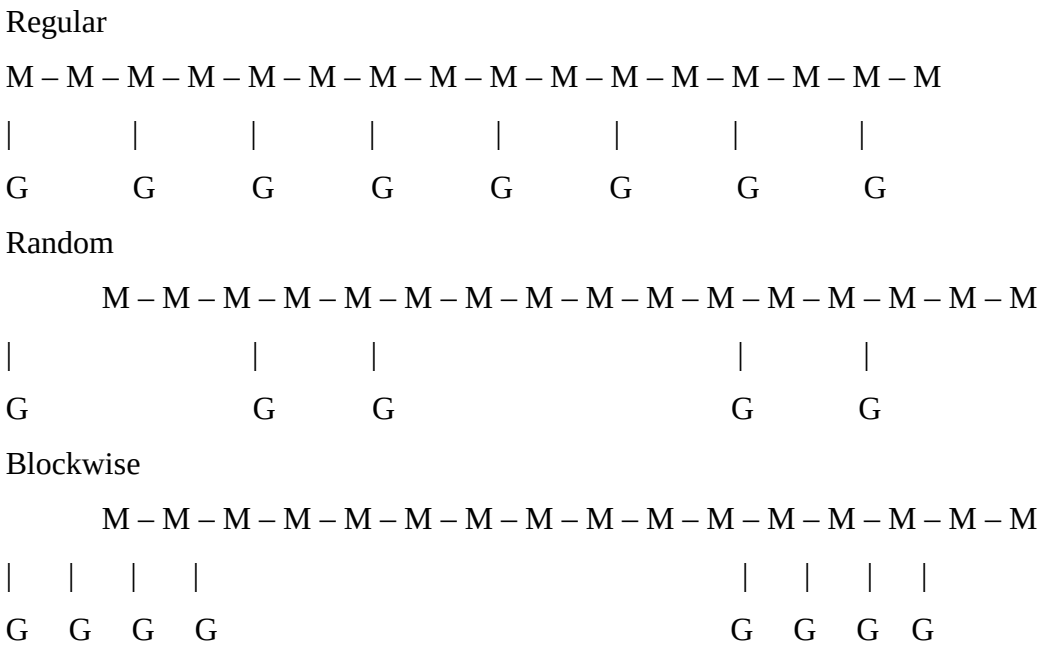


Fig. 2.5: Arrangement of the galactosyl side chains in galactomannan [126]

2.3.2 Biosynthesis of guar galactomannan

Galactomannans are synthesized in endosperm cells by the combined action of two enzymes, mannan synthase (ManS) which makes β-1, 4-linked mannan backbone, and α-galactosyltransferase, which adds galactosyl residues to the mannan backbone [42, 50, 131]. Mannan synthase, the enzyme that forms the β-1,4-linked backbone of galactomannan, has

been assayed *in vitro* by using the membrane particles which are derived from developing endosperm. The results showed independent activity leading to mannose polymer [49]. It showed a high affinity for GDP-mannose (substrate), and divalent cations wre required for the activity [141]. But galactosyltransferase has been found to be dependent on mannan synthase for activity; it shows no activity in presence of the substrate UDP-galctose alone.

In vitro study shows that preformed long mannan chains cannot be galactosylated by the enzyme. However the combined action of both mannan synthase and galactosyl tranferase has produced polymers similar to galactomannan [49]. They have proposed a model for the interaction of GDP mannose mannosyltransferase and UDP-galactose galactosyltransferase in galactomannan biosynthesis [49]. The study on galactosyltransferase from fenugreek has shown that the enzyme acts on D-manno-oligosaccharides of length greater than or equal to 5 acts as acceptor. But the D-manno-oligosaccharides longer than 9 monomers require heating for dissolving [52].

The degree of galactosylation of natural galactomannans is also believed to be determined by two methods. One in which the control is at biosynthesis level i.e. in guar. Second kind of regulation involves the alteration of galactomannan formed by the action of α-galactosidase later in seed development [50] like that in locust bean gum. However both the mechanisms of mannose:galactose ratio control in plants are genetic [130]. The pathway for biosynthesis of guar galactomannans is shown in Fig 2.6.

The direct precursors, GDP-D-mannose and UDP-D-galactose, for the biosynthesis of galactomannan are produced by of GDP mannose phosphorylase (EC 2.7.7.22) and UDP-galactose 4-epimerase (EC 5.1.3.2) enzymes. The M:G ratio of the galactomannan has also been found to be affected by the relative concentrations of these precursors as confirmed by *in vitro* studied [50].

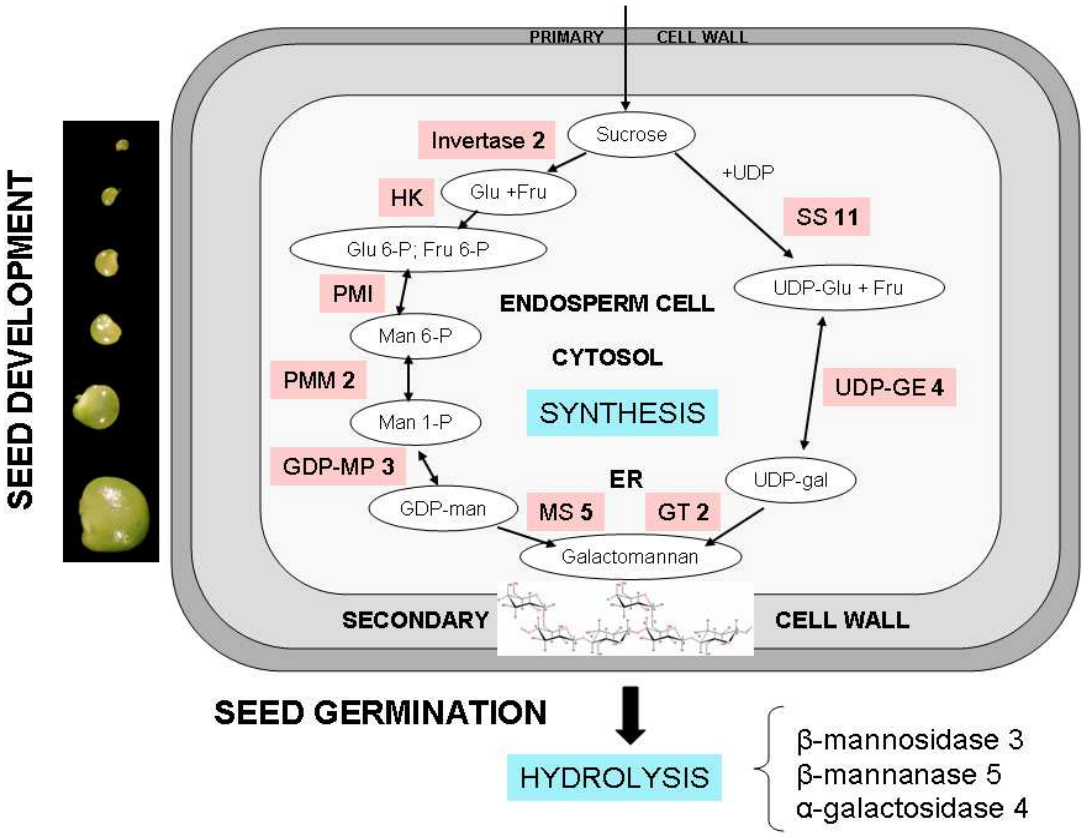


Fig. 2.6: Galactomannan metabolism in guar seeds [114]

The biodegradative hydrolysis of galactomannans during seed development requires the presence of three enzymes α-galactosidase, β-mannanase and β-mannosidase. α - galactosidase hydrolyses the galactose side chain from mannose backbone, β-mannanase cleaves the mannose polymer to oligo mannose and β-mannosidase hydrolyses the oligomannanans. These enzymes are synthesized during the seed development to nourish the developing seedling [114]

2.3.3 Genes involved in galactomannan biosynthesis

Identification of genes involved in the biosynthetic pathway of cell wall polysaccharides has been considered difficult till recent times because of various reasons. Plant cell wall synthesizing enzymes are integral membrane proteins with one or few transmembrane domains. They are quite challenging for biochemical studies as they tend to

be labile, and present in multimeric complexes and are encoded by large gene families whose members may have overlapping function [42].

UDP-galactose epimerase catalyzes the reversible conversion of UDP-D-glucose to UDP-D-galactose. This compound is a precursor for the biosynthesis of many cell wall polymers, including galactomannan. Two cDNA clones encoding two different UDPG epimerases were isolated. [80].

Success in identifying cDNA encoding mannan synthase (ManS) was achieved through the identification of a candidate gene via transcriptional profiling followed by functional expression in a heterologous system such as soybean somatic embryos. Further analysis showed that the gene belongs to cellulose synthase like (*Csl*) gene family [42, 141]. The cellulose synthase-like (*Csl*) genes were first identified in the model plant *Arabidopsis* [136] and rice [75] which are grouped under this family based on sequence homology [32].

The enzyme responsible for the transfer of galactose (Gal) residues to a mannose (Man) on the mannan backbone is a member of the glycosyl transferase (GT) family of proteins called galactomannan galactosyltransferase (GMGT). The putative cDNA encoding galactosyl transferase from fenugreek seed was cloned and expressed in yeast *Pichia pastoris*. The expressed galactosyl transferase showed transfer of galactose to D-manno-oligosaccarides with chain length 5 or more [52]. The mRNA sequence for galactosyl transferase from guar is available in NCBI database with sequence id AJ938067.1 (gmgt1 gene).

The mannan synthase (ManS) and galactomannan galactosyl transferase (GMGT) enzymes are localized in the membrane of Golgi vesicles and are believed to work together very closely to determine the statistical distribution of galactosyl residues along the mannan chain [51].

2.3.4 Genes involved in galactomannan biodegradation

The gene encoding α-galactosidase, an enzyme that hydrolyses the galactose side chain from mannose backbone, was identified using oligo-nucleotide mixed probes based on the terminal amino acid sequence and the sequence of an internal peptide. The nucleotide sequence of the cDNA clone showed that the enzyme is synthesized in the form of a precursor with a 47 amino acid NH_2 terminal extension. This pre-sequence mainly functions to target the protein outside the aleurone cells into the endosperm [115].

2.4 Expressed sequence tags of guar

A database of 16,476 guar seed ESTs was constructed from two cDNA libraries consisting of 8,163 and 8,313 ESTs sequences. Among the seed storage proteins, the most abundant contig represented a conglutin accounting for 3.7% of the total ESTs from both libraries [114].

2.5 Mannose and galactose ratio in galactomannans

Galactomannans with good viscosity are much in demand. This has lead to various biotechnological applications to improve M/G ratio of galactomannans both *in vivo* and *in vitro*. Galactomannans with varying galactose content were prepared by manipulating reaction time, temperature and enzyme/guar gum ratio. Enzymatically modified guar galactomannans with 22-24% galactose contents were found to reproduce the rheological and stabilization properties of locust bean gum [20].

The available information on genes involved in synthesis and biodegradation of galacotmannans lead to transformation studies. Fenugreek seed galactomannan is almost fully substituted by galactose, whereas galactomannan in tobacco seed (*Nicotiana tabaccum*) contains very low level of galactose substitution and exogenously introduced fenugreek

GMGT dominated over the endogenous tobacco GMGT and can operate mutually with the endogenous mannan synthase in tobacco [133].

The *Lotus japonicus* plant has native galactomannan with high galactose substitution (Man/Gal ratio of 1.2-1.3). The transformation of galactomannan galactosyltranferase constructs, resulted in the modification of its galactomannan to one with a lower galactose content at a Man/Gal ratio of 6 [51]. An obvious target of such engineering would be guar which is already used for production of galactomannan. About 30% of the guar transformants produced endosperm with galactomannans where the galactose content was significantly reduced [79].

A ~1.6 kb guar mannan synthase (MS) promoter region has been cloned and characterized by studying the quantitative expression of β-glucuronidase (GUS) directed by MS promoter. GUS expressed specifically in endosperm of transgenic alfalfa [112].

2.6 Applications of guar gum

Guar gum has numerous applications. Fig 2.7 and Table 2.1 shows the applications of guar gum in different industries.

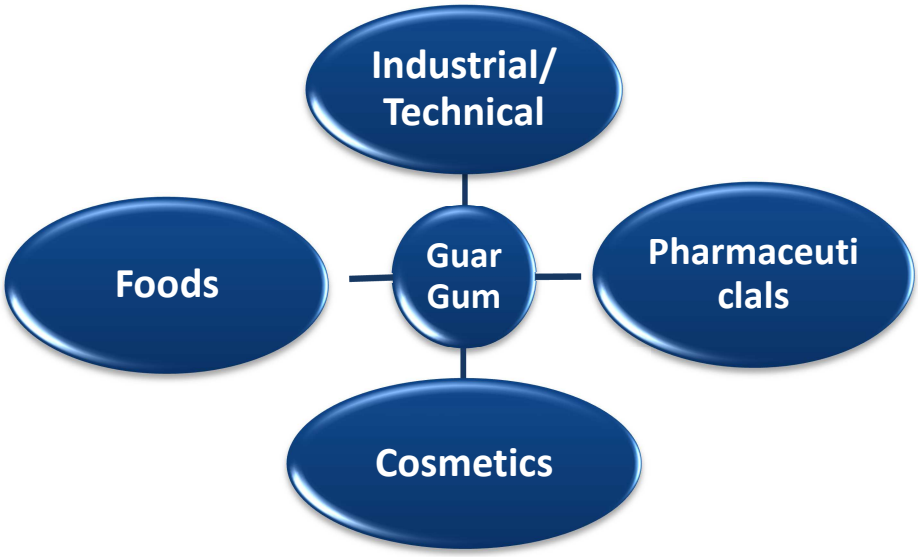


Fig. 2.7. Applications of guar gum

Table 2.1. Applications of guar gum and its modified derivatives in various industries [163]

S. No.	Industry	Uses	Derivatives	Functions
Industrial/ Technical				
1	Oil well drilling	Driling Fluids hydraulic fracturing	Borate cross-linked guar gum, hydroxy alkyl ether derivatives	Control of water loss, viscosity, suspension, turbulence, mobility, friction reduction
2	Textile Printing	Cotton, Rayon silk, wool sizing, carpet printing	Carboxy-methyl guar, hydroxy propyl guar, modified guar gum	Reduces wrap breakage, reduces dusting film forming thickening for dye
3	Paper	Wrapping paper, kraft, photographic paper, filter	Oxidized guar gum, crosslinked guar gum, amino ethyl gum, modified guar gum, guar gum formate	Replaces hemi cellulose, increase strength, fold, pick, pulp hydration, retention of fines, decreases porosity
4	Mining	Concentration of ore, filteration	Aminoethyl guar gum, sulphate of guar gum	Flocculating and settling agent, filter aid
5	Explosive	Stick explosive, blasting slurries	Reticulated guar gum, cyanoethyl ether of guar gum	Water proofing, gelling Agent
6	Water Treatment	Industrial water, drinking water	Food grade guar gum	Coagulant aid (food approved)
7	Tobacco	Reconstitution of fragmental tobacco	Reaction product of carboxymethyl cellulose and guar gum	Binding agent, strengthening agent
8	Coal Mining	Coal suspension, shock impregnation	Borate cross-linked guar gum	Friction reducing suspending agent
9	Fire fighting	Water for fighting fires	Guar gum with ethylene glycol and glycerol	Friction reducing, dispersion and direction control
10	Ceramic	Enamels, electroceramics	Chlorinated guar gum	Fixing, binding thickening agent
11	Photography	Emulsions, gelatine solutions	Borate cross linked guar gum, hydrolysed guar gum	Gelling, hardening agent
12	Synthetic Resins	Polymerization,	Suspension of guar	Thickening, Binding

		suspension, collagen dispersion	gum with CMC	agent
Foods				
13	Frozen foods	Ice creams, Soft serves, frozen cakes	Food grade guar gum with CMC	Water retention, ice crystel inhibitor, stabilizer
14	Bakery	Bread, Cakes, Pastry, Icing	Non-metabolised guar gum	Dough improvement, greater moisture retention, prolonged self life
15	Processed Cheese	Cottage cheese, cream cheese	In combination with other water soluble gums	Increase the yield of curd solids, improves tenderness
16	Dairy Products	Yoghurts, desserts, molasses	In combination with other water soluble gums	Inhibits when separate keeps texture after sterilization
17	Dressing and Sauses	Salad cream, pickles, barbecue relish	In combination with other water soluble gums	Fast, cold dispersible thickening and texturising agent
18	Instant mixes	Pudding sauses, desserts, beverages	In combination with other water soluble gums	Fast, cold dispersible thickening and texturising agent
19	Canned Foods	Pet foods, corned meat, baby foods	In combination with other water soluble gums	Acid resistant thickening and suspending agent
20	Beverages	Cocoa drink, fruit nector, sugarless beverages	In combination with other water soluble gums	Acid resistant thickening and suspending agent
21	Animal Feed	Veterinary preparations, calf milk replacer	In combination with other water soluble gums	Suspending agent, granulating agent
Pharmaceuticals				
22	Pharmaceuticals	• Laxative, slimming aids • Gastric hyper acidity • Diabetic treatment • Cholesterol • Tablets	Food grade guar gum	Bulking agent, bulk forming appetite depressant Synergistic activity with bismuth salt Reduction of urinary glucose loss Stable water soluble suspension

Cosmetic				
23	Cosmetics	Ointment Lotions Hair Shampoos Hair Conditioners	Hydroxypropyl guar (HPG Hydroxypropyl guar (HPG) Food grade guar gum Cationic guar Hydroxypropyl guar (HPG)	Thickening agent gives unctuousness Lubricating, suspending agent Disintegrating and granulating agent Detergent compatible thickener Protective colloid film forming agent

2.7 Cell wall

All green plants are surrounded by a strong layer which is composed of polysaccharides and glycoproteins components. Most abundant biopolymers on earth are the plant cell wall polysaccharides. The cell wall gives shape to the cells and helpful in growth, development and mechanical properties of plant [99]. Plant cell wall has very unique role in human and animals's nutrition as dietary fiber and for industrial uses such as production of renewable biofuel, adaptations to the environmental changes as well as protection against pathogen attacks [128, 132].

2.7.1 General properties of cell wall

Several cell wall biosynthetic activities were demonstrated many years ago in different *in vitro* assays, but it is difficult to demonstrate enzymatic activities directly in heterologous hosts [63]. The two main cell wall polysaccharide fractions are, cellulose and matrix polysaccharides (pectin and hemicelluloses), the former is synthesized at the plasma membrane whereas the latter are made in the Golgi and then exported to the cell wall by exocytosis [129]. Due to the cellulose crystals of several dozen glucan chains form the microfibrillar foundation of plant cell wall [64]. The most important functional property of the plant Golgi apparatus is its ability to synthesize complex matrix polysaccharides of the cell wall. The cell wall matrix polysaccharides are assembled exclusively in the Golgi-cisternae

and transported to the cell surface within Golgi-derived vesicles [99]. Golgi apparatus synthesizes complex matrix polysaccharides of the cell wall (Fig.2.8) [73].

In dicots, 20% of the primary cell wall consists of the polysaccharides xyloglucan. Xyloglucan consists of a β-1,4-glucan backbone with xylosyl, xylosyl-galactosyl or fucosyl–galactosyl–xylosyl residues side chains [28]. The main enzymes responsible for xyloglucan biosynthesis are divided into different families based on donor, acceptor and linkage [167]. Developments in crops and management systems is required for the improvement of lignocellulosic feedstock production [29]. The addition of side chains to xyloglucan is catalyzed by specific transferases such as xylosyltransferase, galactosyltransferase and fucosyltransferase [143]. Glycosyltransferases are implicated with cell wall synthesis. Crystal structure of UDP-glucose specific glycosyltransferases adopted the GT-A fold and possessed DXD motif that co-ordinate with Mg^{+2} [57]. Wall-associated receptor-like kinases (WAKs) is a gene family identified in plants such as *Arabidopsis* and rice which are involved directly in linking the extracellular matrix with intracellular compartments and developmental processes and stress response [82]

2.7.2 Biosynthesis of plant Cell wall polysaccharides

Biosynthesis of plant cell wall polysaccharides is a very complicated process. Both primary and secondary cell walls contain celluloses and hemicelluloses. Cellulose is the most important and normally the most abundant wall component among all the polysaccharides. Cellulose microfibrils are enrooted in a matrix that contains other polysaccharides, glycoproteins and proteins. Cellulose is made at the plasma membrane and deposited directly into the wall. Other matrix components are made in the Golgi and delivered to the wall in secretory vesicles [64]. The successful expression of full length AcTPase with an N-terminal 6xHis tag was achieved in pTrcAc expression vector in soluble form. The expression cassette

was induced in Rosetta2 (DE3) *E. coli* lines. This allowed elucidation of its structure and function [94]. The processes and components involved in assembling functional cell walls are largely unknown. Recent advances had helped us to understand that how cellulose and the non-cellulosic polysaccharides of the plant cell wall are synthesized, assembled, and integrated [23]. There are several transcriptional network is involved in regulation of secondary cell wall synthesis [170].

The Golgi apparatus of eukaryotic cells plays a central role in the processing, sorting, and transport of proteins to intra and extra-cellular compartments. But in plants, it has the additional role of assembling and exporting the non-cellulosic polysaccharides, pectins and hemicelluloses of the cell wall matrix, which are important for plant protection and development. One major part of plant cell walls is a diverse group of polysaccharides, the hemicelluloses. One-third of the wall biomass consists of hemicelluloses which surround heteromannans, xyloglucan, heteroxylans, and mixed-linkage glucan. The fine structure of these polysaccharides and their substitution, varies depending on the plant species and tissue type [47].

The enzymes responsible for elongating glycan chains and forming branches are called glycosyltransferases. They require a specific acceptor molecule and as a substrate, sugars activated by the presence of a nucleotide (nucleotide-sugars) mainly in the form of UDP and GDP-sugars [54, 142]. These enzymes take two forms, multimembrane-spanning (type III) processive glycosyltransferases, and type II glycosyltransferases, the latter possessing a hydrophobic N-terminus, believed to anchor the protein in the membrane, and a C-terminal catalytic domain [64].

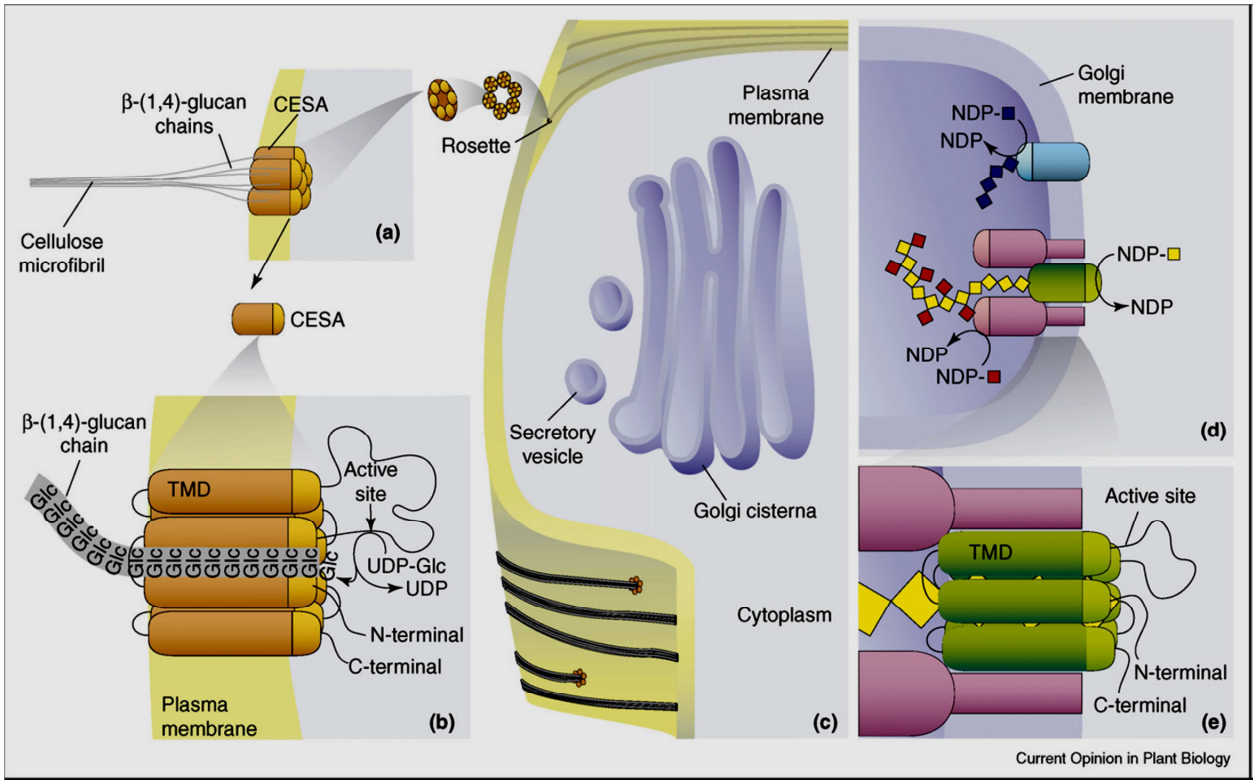


Fig 2.8. Schematic representation of hypotheses regarding cell wall polysaccharide biosynthesis [73].

2.7.2.1 Biosynthesis of cellulose in higher plants

Cellulose is a linear 1,4-β-D-glucan that assembles into paracrystalline microfibrils, each of which contains an estimated 36 parallel polysaccharide chains. Cellulose biosynthesis involves a large multisubunit complex containing at least three different cellulose synthase enzymes and probably other proteins [71]. The cotton fiber cell from the ovule is involved in *in vivo* capacity for the synthesis of cellulose from glucose [24]. These proteins form a complex that appears in plasma membranes as a rosstte structure that is thought to transfer Glucose from cytosolic UDP-Glc to produce multiple extracellular glucan chains that eventually coalesce into a cellular microfibrils.

2.7.2.2 Callose biosynthesis

Callose is a linear homopolymer made up of β-1, 3-linked glucose residue with some β-1,6- branches. It exists in cell walls of a wide variety of higher plants. It is synthesized at the plasma membrane when cells are damaged [25, 38]. Callose plays an important role during a variety of processes in plant development and involved in response to the multiple biotic and abiotic stresses (Fig. 2.9) and callose biosynthesis uses UDP-glucose as a substrate. Several molecular studies and biochemical evidences reported in many plant species indicate that callose is synthesized by a class of enzymes, termed callose synthases.

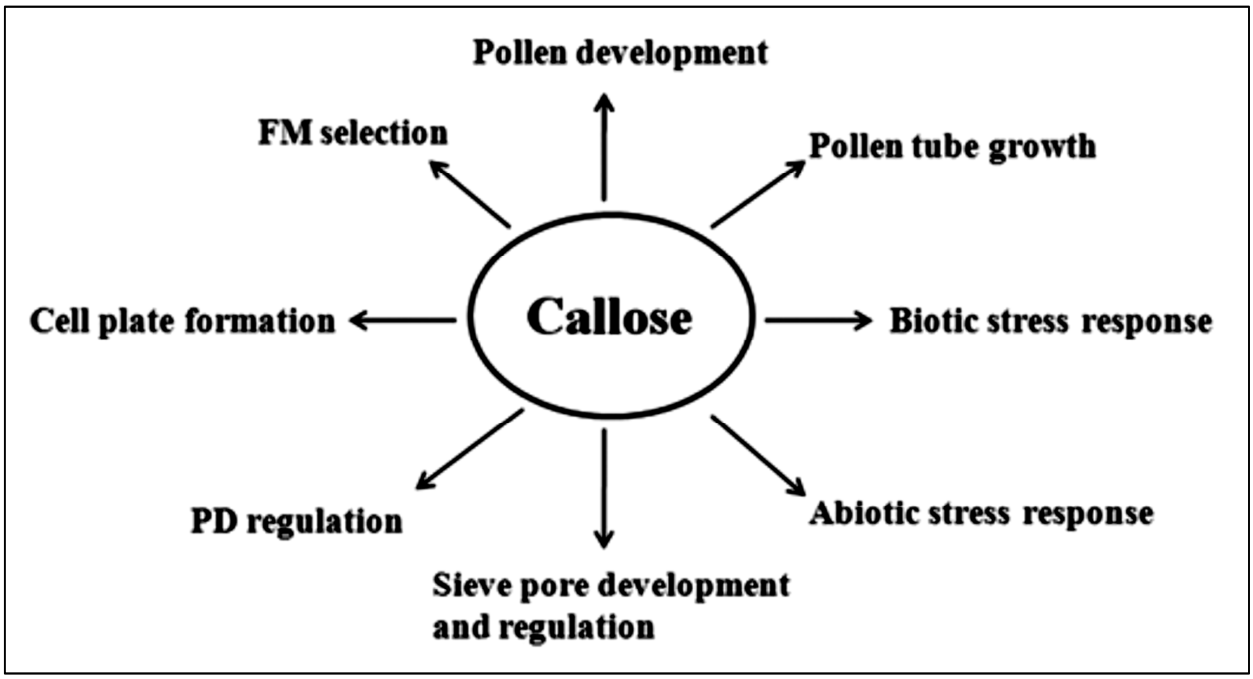


Fig.2.9 Callose is involved in multiple aspects of plant growth and development and response to biotic and abiotic stress [157].

2.7.3 Plant cell wall matrix polysaccharides biosynthesis

Matrix polysaccharides are synthesized in the Golgi and exported to the cell wall by exocytosis. Then these matrix polysaccharides intercalate among cellulose microfilmed brils, which are made at plasmamembrane. Golgi glucan synthesis plays a role in auxin-induced cell

expansion. Recently, certain genes corresponding to glucan synthases have been identified. Several genes from the cellulose synthase-like (*Csl*) family have been found to be involved in the synthesis of various hemicellulosic glycans [141]. Various research groups attempted biochemical purification and substrate labeling to identify polypeptides for cell wall polysaccharides synthases [44].

For the formation of mucin type O-linked glycan UDP-GalNAc:polypeptide alpha-N-acetylgalactosaminyltransferases is required for the catalytic action that is transfer of alpha-N-acetylgalactosamine from UDP-GalNAc to Ser or Thr residues of core proteins [56]. Covalent labeling of proteins with radiolabeled substrate, proved to be relatively promising at least initially, that it allowed identification of a protein of approx. 40kDa that became covalently labeled under GS-I assay condition [44]. The label turned over upon incubation of pre-labeled protein with excess unlabeled substrate, suggesting that it had enzymatic activity. This protein, named reversibly glycosylated protein (RGP), was auto glycosylated , as the purified protein could also be labeled with various UDP-sugars [43]. RGP was specifically localized to the trans- Golgi compartment. The structure of xyloglucan, glucuronoarabino xylan and Galactomannan are given in Fig. 2.10 [40].

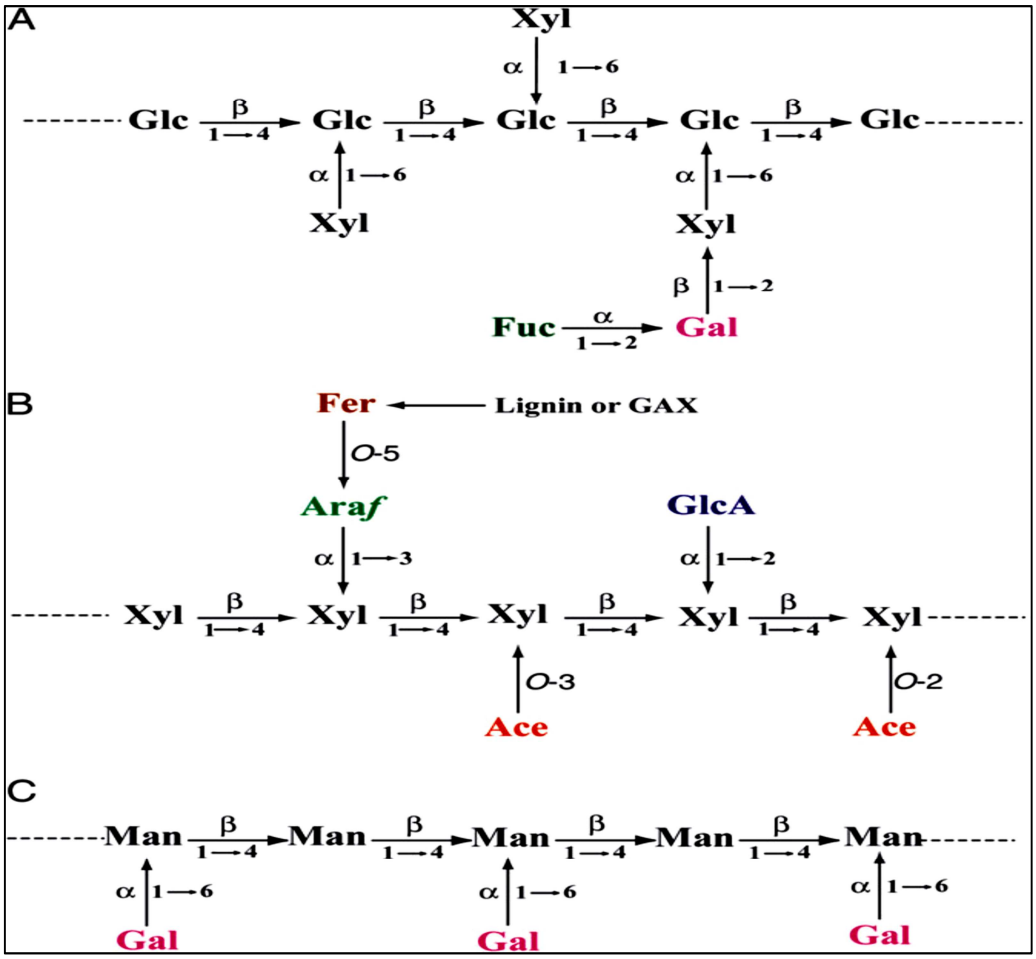


Fig. 2.10. Structure of Xyloglucan (A), Glucuronoarabino xylan (B) and Galactomannan (C) [40].

2.7.3.1 Biosynthesis of Hemicellulose

Hemicellulosic polysaccharides are very complex molecules that are similar to cellulose microfibrils, providing a cross-linked matrix. Hemicelluloses can be divided into four main classes: xyloglucans (XyG), which contain a heavily substituted β-1,4-glucan backbone; (gluco) mannans, containing a variably substituted backbone that includes β-1,4-linked mannose (glucose and mannose) residues; glucuronoarabinoxylans (GAX), containing a substituted β-1,4-linked xylan backbone; and mixedlinkage glucans (MLG), which involve an unsubstituted backbone of glucosyl residues containing both β-1,3- and β-1,4-linkages [149]. Structural similarities between the β-1,4-linked glucan chains of cellulose and the backbones of the various β-linked hemicellulosic polysaccharides led to the prediction that Cellulose Synthase Like (CSL) genes might encode Golgi-localized glycan synthases that are involved in the biosynthesis of these polysaccharides [119] (Fig. 2.11).

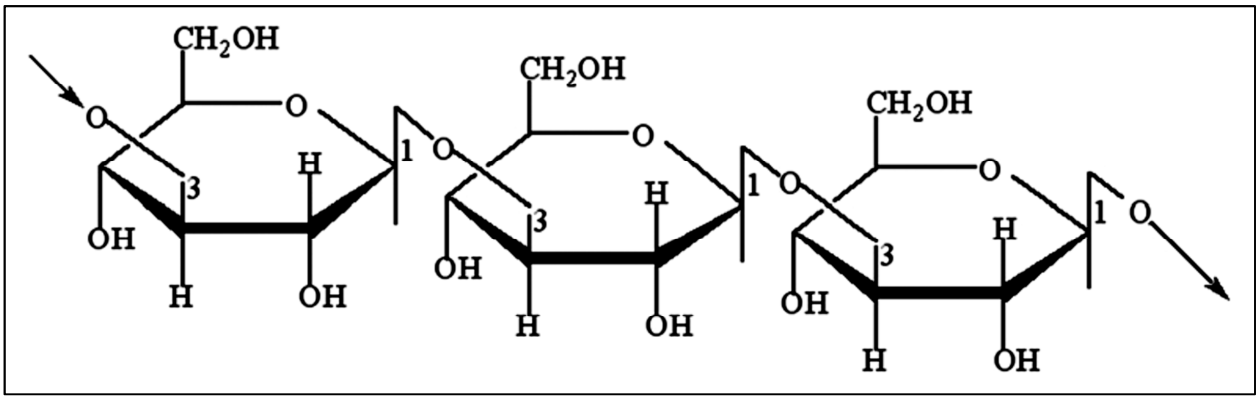


Fig.2.11 Showing fragment of β-1,3- glucan showing how adjacent sugar residues are inverted [119].

2.7.3.2 Mannan biosynthesis

Mannans are also the major hemicellulose in the secondary cell wall of gymnosperms [120]. Heteromannans are synthesized from activated nucleotide sugars. These nucleotide sugars are GDP-mannose, GDP-glucose, and UDP-galactose [100]. The activated nucleotide sugars are then utilized by highly specific, Golgi-localized glycosyltransferases (GTs), which facilitate the formation of the specific linkage between the monomers and thus synthesize the polymer [18].

2.7.3.4 Pectin biosynthesis

The third group of polysaccharides is pectin. The hemicelluloses and pectins constitute the matrix in which cellulose microfibrils are embedded. Structurally and functionally the most complex polysaccharide in plant cell wall is pectin. The structure of pectin constitutes the galactouronic acid rich polysaccharides including homogalactouronan rhamnogalactouronan I and rhamnogalactouronan II (Fig.2.12) [107] .

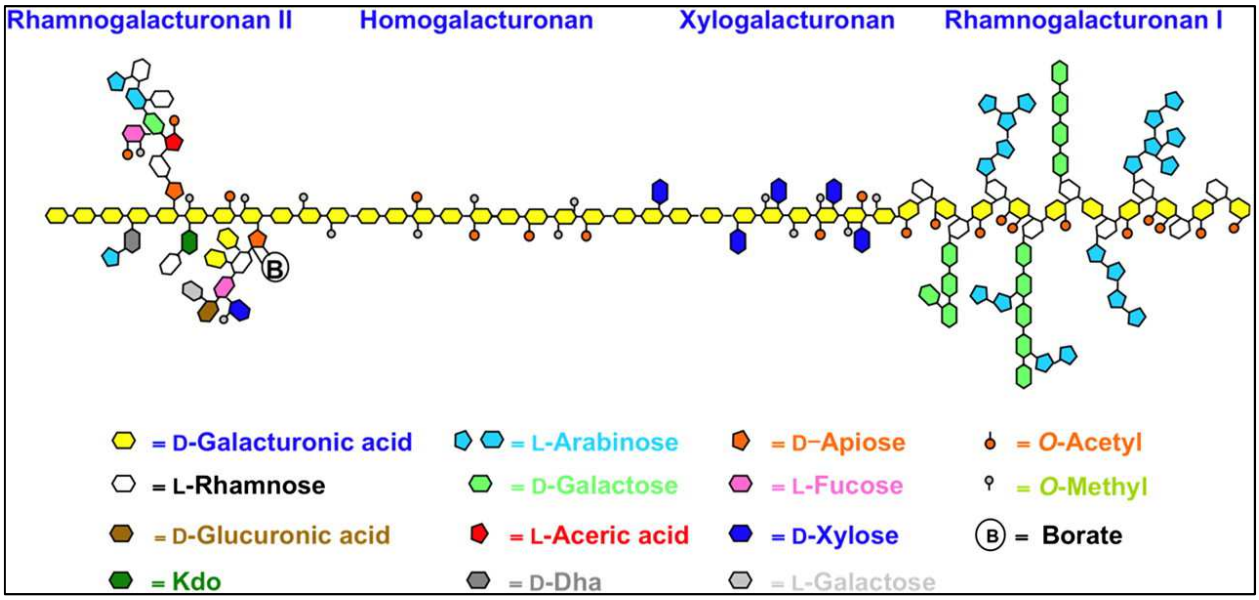


Fig. 2.12 Showing the structure of pectin. Pectin constitutes four different types of polysaccharides, and their structures are shown. Kdo, 3-Deoxy-d-manno-2-octulosonic acid; DHA, 3-deoxy-d-lyxo-2-heptulosaric acid [107].

2.8 Role of glycosyltranferases and sugar-converting enzymes

Golgi glycosyltransferases are required to assemble Golgi mediated complex polysaccharides. Nucleotide-sugar transporters and nucleotide-sugar interconvertion enzymes [135, 137, 158] could interact physically to form complexes within Golgi membranes that would coordinate sugar supply and polymer synthesis [143]

2.9 Reversibly glycosylated polypeptides

The first report of RGPs was detected as a reversibly glycosylatable 40kDa doublet protein in pea membranes with UDP[^{14}C] Glc which is cytosolic in nature but have a tendency to get associated with Golgi membranes, which may be a component of the GS-I system [44]. Subsequently this protein was purified from pea tissue named as RGP1. RGP1, a 41.5kD protein purified from pea has 364 aminoacyl residues which gets localized in trans-Golgi dictyosomal cisternae, as found by immunogold labeling. RGP1 is involved in synthesis of xyloglucan and hemicelluloses. Corn (*Zea mays*) contains a biochemically similar and structurally homologous RGP1, which helps in starch synthesis. The expressed sequence

database also reveals close homologs of pea Rgp1 in *Arabidopsis* and rice (*Oryza sativa*) [43]. Identification of other orthologs in *Arabidopsis thaliana* [37].

Cellulose, a cell wall component of plants, has not yet been isolated and characterized successfully. The cDNAs of cotton (*Gossypium hirsutum*) and rice (Oryza sativa) are homologous of the bacterial celA gene, encoding the catalytic subunit of cellulose synthase. In plant celA gene products, three regions in the sequences were conserved with respect to the proteins encoded by bacterial celA genes. Within these conserved regions four highly conserved subdomains are present, critical for catalysis and binding to the substrate UDP-glucose (UDP-Glc). The overexpressed DNA segment of the cotton celAl gene encodes a polypeptide fragment that spans these domains and binds UDP-Glc, while a similar fragment with deleted domains does not bind to substrate. In developing cotton fibers celAl and celA2 genes expressed at high levels during active secondary wall cellulose synthesis [121].

UDP-Glc:protein transglucosylase (UPTG) (EC 2.4.1.112) is an autocatalytic glycosyl-transferase which helps in starch biosynthesis. In potato, polyclonal antibodies were raised against UPTG to screen a potato swelling stolon tip cDNA expression library. Recombinant UPTG was labeled with UDP-[^{14}C]-Glc and Mn^{2+}, to show its enzyme activity. It was determined that purified as well as recombinant UPTG can be reversibly glycosylated by UDP-Glc, UDP-Xyl or UDP-Gal. RNA hybridization studies and western blot analysis indicates that UPTG mRNA and protein are expressed in all potato tissues. Previous study shows high degree of similarity between UPTG and several plant sequences that encode for proteins localized at the cytoplasmic face of the Golgi and at plasmodesmata [14].

In Arabidopsis thaliana, antibodies were prepared against Reversibly Glycosylated Polypeptide-1 (AtRGP1), to check for expression and intracellular localization of the protein. The concentration of AtRGP1 protein and RNA was found to be highest in roots and suspension-cultured cells, showing that the protein is cytosolic but also peripherally

associated with membranes [37]. Polypeptide assemblies cross-linked by S-S bonds and single polypeptides folded with internal S-S cross-links were detected in particulate membranes and soluble fractions of developing cotyledons of nasturtium (*Tropaeolum majus* L.). When prepared from fruit homogenates, these polypeptides were found to bind reversibly to UDP-Gal (labelled with [^{14}C] Gal or [^{3}H]uridine). Initially, the bound UDP-[^{14}C]Gal could be replaced by adding excess UDP, or exchanged with UDP-Gal, -Glc, -Man or –Xyl [63].

A full-length cDNA encoding RGP was cloned from cotton fiber cells using differential display combined with rapid amplification of the cDNA ends. The gene GhRGP1 encodes for a 359 amino acids long protein having 78-86% identity with RGPs from other plants. Northern blot analysis showed the expression of gene in fiber cells and its abundance during primary cell wall elongation and at the later stage of secondary cell thickening, suggesting that GhRGP1 may be involved in non-cellulosic polysaccharide biosynthesis of the plant cell wall [169]. The complete genome sequence of the Arabidopsis has revealed a total of 40 cellulose synthase (CesA) and cellulose synthase-like (Csl) genes. Recent studies suggests that each CESA polypeptide contains only one catalytic center, and that two or more polypeptides from different genes might be needed to form a functional cellulose synthase complex [39]. In *Arabidopsis thaliana*, characterization of cell wall mutants has led to significant advances in understanding cell wall synthesis and the properties of cell walls [134].

From wheat (*Triticum aestivum*) and rice (*Oryza sativa*) cDNA clones of RGP were selected and sequence comparisons of both showed existence of two classes of proteins, RGP1 and RGP2. Transgenic tobacco (Nicotiana tabacum) plants were generated which overexpress either wheat Rgp1 or Rgp2 and glucosylation assays revealed that that RGP-2 complex was active. The confirmation of RGP1 and RGP2 as high-molecular mass complexes was done by gel filtration experiments. [95]. Isolation of a cDNA clone encoding mannan

synthase (ManS) was reported. This enzyme makes 1, 4-mannan backbone of galactomannan, a hemicellulosic storage polysaccharide in guar seed endosperm walls. [41].

A 41kD protein, SE-WAP41, associated with walls of etiolated maize (*Zea mays*) seedlings was cloned, sequenced and found to be a class 1 reversibly glycosylated polypeptide (C1RGP). Protein gel blot analysis of cell fractions with an antiserum against recombinant SE-WAP41 showed, that it was enriched in the wall fraction. RNA gel blot analysis along the mesocotyl developmental axis and during deetiolation demonstrates that high SE-WAP41 transcript levels correlate spatially and temporally with primary and secondary plasmodesmata (Pd) formation. All four of the Arabidopsis thaliana C1RGP proteins, when fused to green fluorescent protein (GFP) and transiently expressed in tobacco (*Nicotiana tabacum*) epidermal cells, showed fluorescence patterns indicating that they were Golgi and plasmodesmal associated proteins. C1RGPs werfound to be plasmodesmata associated proteins delivered to plasmodesmata via the Golgi apparatus [139].

Reversibly glycosylated polypeptides (RGPs) are implicated in polysaccharide biosynthesis and till date, no direct evidence exists for their involvement in a particular biochemical process. *Arabidopsis thaliana* genome contains five RGP genes out of which RGP1 and RGP2 share the highest sequence identity. Both genes are ubiquitously expressed, but the highest level of expression was observed in actively growing tissues and in mature pollen. RGPs showed cytoplasmic and transient association with Golgi [46]. GhRGP1 gene from cotton has been characterized and was found to be preferentially expressed in fiber cells [162].

Monopartite geminiviruses of the genus Begomovirus have two virion-sense genes, V1 and V2. A protein, SlUPTG1, closely related to a family of plant reversibly glycosylated peptides, was found to interact with V1 protein using yeast two-hybrid system. A SlUPTG1:GFP fusion protein was found to localized to the cell wall and its expression

appeared to be highest in actively dividing tissues [144]. Transient expression for host responces of individual genes encoded by tomato leaf curl virus [145]. The effect of ionic strength of on the activity of StRGP shows incubation of RGP protein at high ionic strength produced a high self-glycosylation level and a high glycosylation reversibility of RGP. Beside of that, incubation at low ionic strength indicates low level of glycosylation. So, low glycosylation reversibility of RGP in *Solanum tuberosum*. Thus at low ionic strength the formation of high molecular weight RGP-containing forms, whereas incubation at high ionic strength produced active RGP with a molecular as expected for the monomer. Another study showed a model for the regulation of the RGP activity and its binding to Golgi membranes as shown in Fig. 2.12 where the sugars linked to oligomeric form of RGP in the Golgi may be transferred to acceptors involved in polysaccharide biosynthesis [34].

Biochemical purification of glucan synthase was unsuccessful because of the labile nature and very low abundance of these enzymes. Several genes from the cellulose synthase-like (Csl) family have been found to be involved in the synthesis of various hemicellulosic glycans [141].

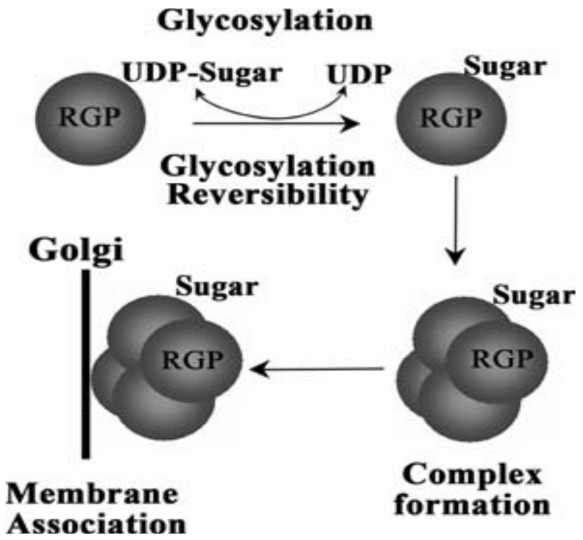


Fig. 2.13 Formation of RGP complexes linked with a decreased ability to glucose to UDP and their association with Golgi. The reaction catalyzed by monomeric soluble RGP is reversible.Monomeric soluble RGP glycosylation activity found due to the formation of high molecular weight RGP containing structures that enhance their affinity for Golgi membrane [34].

In rice, UDP-arabinopyranose mutase (UAM) was found to be involved in interconversion of UDP-arabinopyranose (UDP-Arap) and UDP-arabinofuranose (UDP-Araf). L-Ara, an important constituent of plant cell walls, is found predominantly in the furanose rather than in the thermodynamically more stable pyranose form. Through site-directed mutagenesis studies it was found that in functional UAMs an arginyl residue is reversibly glycosylated with a single glycosyl residue and this residue is required for mutase activity. A DXD motif is required for catalytic activity of this enzyme [84]. These mutases (UAMs) belong to a small gene family encoding the previously named reversibly glycosylated proteins (RGPs) [127].

Except for xylan synthase, known Golgi glycan synthases suggests that the catalytic polypeptide by itself is sufficient for enzyme activity, most likely as a homodimer. The possibility of the involvement of various Csl genes in making more than one product [40]. The role of complex formation is related to rice plant development, in which class 1 Oryza sativa RGP (OsRGP) may involve in an early stage of growing plants, so that higher molecular weight showed an apparent non-reduced form of StRGP protein and the lower radioglycosylation found in reduced form of OsRGP monomer. [35]

2.10 Characterization of proteins using novel techniques

The Opticryst consortium for optimization of protein crystallization developed new crystallization technologies and tools for the crystallization of a number of proteins [59]. For better evaluation of crystals, various techniques used included, DPI (dual polarization interferometry),to detect useful Nucleation; *in situ* DLS (dynamic light scattering), to monitor crystal growth; and UV fluorescence to differentiate protein crystals from salt [58]. Leucoagglutinin protein of *Phaseolus vulgaris* is a homotetrameric legume protein [13]. For the multimerization of PsBGAL α/β barrel containing catalytic site and the other is rich in β-

content are responsible [48]. Thermal denaturation showed that the refolding of phytohemagglutinin at neutral pH is irreversible [12]. *In silico* docking of selective peptide inhibitors with TMPK from *M.tuberculosis* and human was carried out to examine their differential binding at the active site [90]. There are several work has been reported in structural and functional aspect of protein. For the first time structural evidence for the diffusion of aromatic ligands and closure of induced ligand of the hydrophobic channel in heme peroixidase [147].

Several novel genes are functionally characterized by molecular basis like claudin-1/CD81 association, which binds a soluble form of HCVE2 glycoprotein (sE2) [15] and detergent-extracted, purified claudin-1 were antigenic and recognized by specific antibodies. The unique property of biological enzymes is that they can select specific substrate for the selective catalysis of the reactions for a specific function [105]. The most common enzyme of snake venom includes acetylcholinesteases, L-aminoacid oxidases, serine proteinases, metalloproteinases and phospholipases. Protein secondary structure is determined by circular dichroism [67-69]. Protein unfoding can be determined by CD [67]. Thermal stability and resistance to proteolysis make these enzymes more attractive [81]. Hsp90, a ubiquitous molecular chaperone was used for studying substrate binding with systemin. It was found that Hsp90 preferentially binds a locally structured region in a globally unfolded protein, and this binding drives functional changes in the chaperone by lowering a rate-limiting conformational barrier [152]. The analysis of protein composition of tomato spotted wilt virus by PAGE revealed that there are three major and one minor types of structural protein. The three major proteins comprise approximately 98% of the total viral protein and all of them were found to be glycoprotein [106]. One driving force for the complex rearrangements is the intrinsic ATPase activity of Hsp90, has seen with other chaperones[89].

A 55 kDa cruciferin protein from seeds of Moringa oleifera plant was purified and characterized. Crystallographic analysis showed that the crystals belonged to the P6322 hexagonal space group with cell dimensions, a = b = 98.4, c = 274.3 A°. Crystal diffracted to a resolution of 6 Å [3]. Isothermal titration calorimetry is used to directly characterize the thermodynamics of biopolymers binding interactions and the kinetics of enzyme-calayzed reactions [45, 55, 98]. Model-independent differential equation is used for the analysis of ITC [33]. Cellulose-binding domain of exoglucanase Cex from Cellulomonasfimi to insoluble microcrystalline cellulose binding is entropically driven [30].

3. MATERIALS AND METHODS

3.1. PCR amplification and construction of the recombinant plasmid

A full length guar (*C. tetragonoloba* L. Taub) RGP (CtRGP) cDNA was provided by Dr. Dhugga from the previously established EST database available at Du Pont Pioneer, USA). This database of approximately 15, 000 ESTs was constructed from the guar seeds at three developmental stages with approximately 5,000 ESTs derived from each of the cDNA libraries [41]. Fig 3.1 shows the complete sequence of the CtRGP cDNA. It consists of 1,455bp and encodes from the coding region 169-1230, hence length is 1062bp.

GCACGAGGCTCAAACATTGAGACACGTCTCAGTGCAATTCCCCCCACTCGTGCCGAAACCCCCGCCCTCGCACGT
GCCACTACCCCTTTTATATTGCACTCTCCCTCTCTTCTTCTTCCTTCAAATCTCTCTCACACACAGAGACACACA
AACCTCTGTGAAACTACCATGGCATCGGCAACACCATTGTTGAAGGATGAGCTCGACATCGTGATCCCCACAATC
AGGAACCTCGATTTCCTCGAGATGTGGAGGCCGTTCTTTCAGCCCTACCATCTCATAATTGTGCAAGATGGTGAC
CCTTCCAAGACCATTAAGGTCCCTGAGGGATTTGACTATGAGCTCTACAATCGCAATGATATCAATAGGATTCTC
GGTCCCAAGGCCAATTGCATCTCCTTCAAGGACTCTGCTTGTCGCTGCTTTGGTTATATGGTCTCTAAGAAGAAG
TACATCTACACCATTGATGATGATTGCTTCGTTGCTACTGATCCATCAGGACACAAGGTTAATGCACTTGAGCAG
CATATAAAAAACCTGCTCTGCCCATCAACACCCTTCTTTTTCAATACCCTCTATGAACCTTTCAGAGAAGGTGCA
GATTTCGTTCGTGGCTACCCCTTCAGTCTCCGTGAAGGTGTACCAACTGCTGTTTCTCATGGTCTTTGGCTAAAC
ATCCCAGACTACGATGCTCCTACTCAACTTGTGAAACCTCTTGAGAGGAACACAAGGTATGTGGATGCTATCATG
ACCATACCAAAGGGCACTTTGTTTCCCATGTGTGGAATGAACTTGGCATTCGACCGTGATCTTGTTGGACCAGCA
ATGTACTTTGGTCTCATGGGTGATGGTCAACCTATTGGACGCTACGATGACATGTGGGCTGGCTGGTGCTGCAAG
GTAATTTGTGATCACTTGGGATTGGGAATCAAGACTGGTCTACCATATATCTATCACAGCAAGGCTAGTAACCCA
TTTGTTAACCTGAGGAAAGAGTACAAGGGCATATTCTGGCAAGAAGACATTATCCCATTCTTCCAGAACATTGTT
CTTCCAAAAGAAGCTACCACTGTTCAGAAGTGTTACGTTGAGCTAGCCAAGCAAGTCAAGGACAAACTTACCAAA
ATAGATCCTTACTTTGACAAATTAGCAGAGGCCATGGTCACTTGGATTGAAGCTTGGGATGAGCTTAACCCTGCT
GGAGCATCTCTTGCCAACGGCAAAGCATAAATCAGTTTTGGTTAAATTCCTGACAGCCACAGAACATGAAGTGGA
ACTATTTATATACTGTATTTCTTTTTCCTAGTATTTATTTTGATATTGTATTCTACTCTTAAAATTTGCAGCATT
GTTAGCTGTAATTCTATGTAGTAACTTCATGCTTTGAGCTTTAGGAACAATAAAGTACAACGGTCTTTAGGGCGT
TGCTTTCTTTTAAAAAAAAAAAAAAAAAA

Fig. 3.1. Full-length 1,455 bp FASTA sequence of clone code (>lds2c.pk011.p19:fis) for guar RGP (Provided by Dr.Dhugga).

3.1.1. Polymerase chain reaction (PCR)

The selected full-length cDNA fragment of guar RGP was amplified in vitro using the polymerase chain reaction (PCR). PCR was carried out for 20µl reaction volume containing 1U Taq DNA polymerase (Biotools, USA). The reaction master mix is described in Table 3.1. Amplification was carried in a Master cycler gradient programmable thermal cycler (Eppendorf UK). The double-stranded template cDNA was initially denatured at 94℃ for 5

minutes (min.) to produce single stranded cDNA. Each of the 30 amplification cycles began with denaturation of the DNA at 94^0C for 1 min. The temperature was then decreased to 56^0C for 1 min which allowed annealing of the two primers to the 5' and 3' ends of the cDNA. This is followed by a 1.5 min elongation period at 72^0C where the thermostable polymerase synthesized the DNA strands complementary to the cDNA by adding the matching dNTPs to the primers.

Table 3.1: Composition of master mix of PCR reaction

Components	Final concentration
Template	50ng
Forward primer	0.5µl (20pM)
Reverse primer	0.5µl (20pM)
dNTPs	0.5µl (10mM)
MgCl2	0.5µl (50mM)
Buffer	2.5µl (10X)
Taq polymerase	1.0µl
MiliQ water	10.5µl
Total	20.0µl

A final extension carried out at 72^0C for 15 min. and a hold temperature 4^0C at the end. Preparative PCR was carried out using 160µl reaction volume. The coding region of full-length cDNA was amplified by PCR with the primers RGPflFOR1 (5' CGTAGGTCCATATGGCATCGGCAACACCATTG 3') and RGPflREV1 (5' GCAAGTCAGTCGACTGCTTTGCCGTTGGCAAGAG 3'. *NdeI* and *SalI* sites are indicated in bold in primers.

3. 1. 2. PCR purification

The amplified PCR product of CtRGP was purified using Qiagen PCR purification kit (Qiagen, USA). Four volumes of binding buffer B2 (supplied with the kit) added with 10mL isopropanol in PCR sample (50-100 µL) and were mixed properly. Column was centrifuged at

12,000 x g for 1 minute and supernatant was discarded. The column was re-inserted into the collection tube and 650 µL wash buffer W 1 was added with ethanol. Again column was centrifuged at 10,000 x g for 1 min. supernatant was discarded and the column was placed in the same collection tube. The column was centrifuged at maximum speed for 2-3 min. The column was placed into a clean tube of 1.7 mL. A 50 µL elution buffer was added to the center of the column. The column was incubated at room temperature for 1 min. and column was centrifuged at maximum speed for 2 min. Finally the elution tube contained the purified PCR product. The purified DNA was stored at 4℃ for immediate use or at -20℃ for long term storage.

3.1.3 Ethanol precipitation of DNA

Purified DNA was ethanol precipitated by adding $1/10^{th}$ of the sample volume of 3M sodium acetate (pH 5.2) and 3 times sample volume of 100% ethanol. The sample was mixed by inversion and stored at -20^0C for 1-24 hour (hr.) to precipitate DNA. The mixture was centrifuged at 15000 g for 30 min. at room temperature in Eppendorf micro centrifuge. The supernatant was discarded and the pellet was washed twice with 70% ethanol and centrifuged at 15000 g for 5 min. The supernatant was removed and the DNA pellet was air-dried and resuspended in 10 mM Tris pH 8.0.

3.1.4 Measurement of DNA concentration

The quality of the precipitated DNA was checked by gel electrophoresis on 0.8% agarose gel [140]. The quantity of DNA was estimated by measuring absorbance at 260 nm in a spectrophotometer (Varian, USA) and diluted to approximately 100ng/µl. The diluted DNA samples were stored at -20^0C.

3.1.5 Restriction digestion of purified PCR product and pET 29a (+) cloning vector

Purified PCR product and pET 29a (+) (Novagen, UK) cloning vector were digested with restriction endonucleases *Nde*I and *Sal*I (NEB, UK) in two reactions. The reaction mixtures are described in Table 3.2. The components of reaction mixture were mixed gently and incubated at 37^0C for 1 hr.-45min and 3hrs for purified PCR product and pET 29a (+) cloning vector, respectively. The digested purified PCR product and pET 29a (+) cloning vector were analyzed by gel electrophoresis. The DNA of required length was extracted from the gel using low melting agarose by cutting the DNA band and purified using Qiagen gel extraction kit as per the manufacturer's instructions. The extracted DNA was quantified using UV-spectrophotometer.

Table 3.2. Composition of the restriction digestion reaction

Purified PCR product		**pET 29a (+) cloning vector**	
Components	Reaction volume	**components**	Reaction volume
BSA(10X)	3µl	**BSA (10X)**	3µl
Buffer 3	3µl	**Buffer 3**	3µl
PCR amplicon	21 µl	**Plasmid (pET 29a)**	21 µl
NdeI	01 µl	***NdeI***	01 µl
SalI	01 µl	***SalI***	01 µl
Protease /DNase free H_2O	01 µl	**Protease/DNase free H_2O**	01 µl
Total	30 µl	**Total**	30 µl

3. 1. 6. DNA ligation

DNA ligation was carried out for a final volume 20µl by using T4 DNA ligase (NEB, UK). The reaction mixture contained 2µl of 10X ligation buffer, 50 µg (0.025 pmol) purified vector DNA, 1µl DNA ligase and 50µg insert DNA in a molar ration 3:1 (insert: vector). The ligation reaction mixture was incubated at 16^0C for overnight (16 hrs.). The ligated product construct (plasmid) was named as pET 29a.CtRGP **(Fig. 4.7).**

3.1.7. Glycerol stocks preparation

The glycerol stocks of *E.coli* cells containing plasmids od desired sequence were prepared. For 1ml glycerol stock, 550 µl *E.coli* cell culture OD 1.0 was mixed with 450 µl of 50% glycerol (autoclaved). The mixture was vortexed and stored at -80^{0}C.

3.1.8. Transformation into competent E.coli (DH5α) cells with recombinant pET 29a (+) vector

The ligated Plasmid was transformed into *E. coli* DH5α strain competent cells by using heat shock method. Cells were plated on LB agar plate containing kanamycin (50 µg/ml) as a selection resistant marker of the expression vector. Plasmid of transformed colonies was isolated with Invitrogen plasmid isolation kit (Invitrogen, USA). Isolated plasmids of transformed colonies were screened by colony PCR. Cloning was confirmed by restriction digestion by using *Nde*I and *Sal*I restriction enzymes. Restricted plasmids of transformed colonies were run on 0.8% agarose gel.

3.1.9 Purification of Plasmid DNA

A 5 ml of LB medium containing appropriate antibiotics in a sterile 50 ml falcon tube was inoculated with a single colony of the *E.coli*. strain harboring the plasmid of interest. The culture was incubated overnight at 37^{o}C with shaking (200 rpm). The culture was centrifuged at 4500 g for 7 minutes at room temperature (Eppendorf centrifuge 5810 R), the supernatant was discarded and the cell pellet was then processed using a QIAprep Spin Miniprep Kit (Qiagen, USA) as per the manufacturer's instructions. Plasmids were eluted in 30 µl of 10 mM Tris pH 8.0.

3.1.10. Concentration determination and sequencing of plasmid DNA

The quality of the precipitated DNA was checked by gel electrophoresis on 0.8% agarose gel (Sambrooke and Russell 2001). The quantity of DNA was estimated by measuring absorbance at 260 nm in a spectrophotometer (Varian, USA) and diluted to approximately 100 ng/µl. The diluted DNA samples were stored at -20^{0}C.

3.1.11. Sequencing of the DNA construct

The plasmid was recovered and the insert was checked by DNA sequencing. DNA sequencing was commercially done using ABI sequencer by Biolink Company (Applied biosystem) and universal T7 promotor primer for forward 5'-3' and T7 terminator primer for reverse 3'-5' were used for pET 29a (+) plasmid vector. The sequencing results were analysed with nucleotide BLAST (Basic sequence alignment search tool) (http:// www.ncbi.nlm.nih.gov/BLAST) and clustal W [155] and was then compared with the original cDNA sequence. The nucleotide sequence of the construct was then translated into its protein sequence using the Translation Tool in SwissProt (http://www.expasy.org/too).

3.2. Recombinant protein over-expression and cell lysis

Plasmid was isolated from the *E.coli* DH5α cells containing the recombinant pET 29a (+) plasmid having an insert of CtRGP cDNA. The pET 29a.CtRGP expression construct was transformed into *E.coli* BL21 (DE3). Cells were grown at 37℃ in Leuria broth (LB) containing 50 µg/ml kanamycin until A_{600} nm $\sim 0.6 - 0.8$. Following induction with 0.4mM isopropyl 1-thio-β-D-galactopyranoside (IPTG, Gold Sigma, USA), the cultures were grown at 18℃, 25℃ and 37℃ for overnight, 8 hours and 6 hours respectively. The best yield was obtained at 18℃ and this temperature was used for large scale production of recombinant CtRGP protein.

3.3. Purification

The RGP protein was produced as a fusion protein with the 6X His-tag at the C-terminal end. The *E.coli* cells were harvested by centrifugation (5, 000 X g; 10 min) at 4℃ and resuspended in lysis buffer (25 mM Tris, pH 8.0 and 100 mM NaCl). The resuspended cells were subjected to sonication. After sonication and centrifugation (10, 000 X g for 60min) at 4℃, the supernatant was loaded onto a Ni^{+2}-nitrotriacetic acid column previously equilibrated with lysis buffer. The column was washed with 25volumes of binding buffer (25 mM Tris pH 8.0 and 250mM NaCl) and 50mL and 40mL wash buffer I and II [20mM and 40mM Imidazole pH=8.0] respectively. His-tagged protein was eluted using a gradient of imidazole ranging from 20mM to 250mM, at pH 8.0. The purification was confirmed by analysis on 12% SDS-PAGE gel using Coomassie blue and silver staining. The RGP protein was concentrated to 9.5-12mg/ml with a protein concentrator Amicon Ultra 15mL (Millipore, UK). Bradford assay was used to determine the protein concentration [17] with Bradford reagent (Sigma). Protein concentration was also determined using molar extinction coefficient for CtRGP calculated using ProtParam [61].

3.3.1 Western blot analysis

For western blot analysis, the cell lysates were prepared after harvesting them in lysis buffer (25 mM Tris pH 8.0 and 100mM NaCl. A 20 µg of total protein samples was separated in polyacrylamide gel with SDS (5% stacking gel, 12% separation gel),using the protocol [93], in a minigel system (Mini-Protean II Dual Slab Gel System from Bio-rad, Mississauga, ON, Canada), 100 volts was applied and an electrophoresis buffer (50mM Tris; 384mM glycine; 0.1% SDS) at pH 8.3 was used. The protein band was transferred to a nylon membrane and immunoblotting was carried out according to the standard protocol by the method of [156]. The analyzed proteins were transferred to nylon membrane. The blots were

then blocked with TBST buffer (20 mM Tris-Cl, pH 7.5, 150 mM sodium chloride, 0.05% Tween-20) containing 5% skim milk powder. They were then washed with TBST buffer (without milk powder) and incubated overnight at 4ºC with TBST buffer containing appropriate amounts of Primary antibody (1:10,000) as per Dhugga and coworker [43]. RGP protein detection was done by Pea anti-RGP purified antibody (gifted by Dr. Dhugga USA) polyclonal plant specific primary antibodies with a negative control of α-galactosidase from green coffee beans purified (20 µg) (Sigma G8507-5UN). The nylon blots were washed and incubated with anti-rabbit secondary antibody (1:20,000) conjugated with horse radish peroxidase (HRP). Specific protein was detected by the Chemilluminensce detection system.

3.3.2. Determination of protein concentration

Protein was quantified using Bradford's assay reagent (Sigma). Five microlitre of protein sample was mixed with 795µl of buffer (purification buffer) and 200µl of Bradford's reagent was added to it and incubated for 2 min. The protein concentration was determined using Perkin Elmer Lambda 25 UV-visible spectrophotometer by using the absorbance reading at 595nm. Standard curve of bovine serum albumin (BSA) was generated every time the assay was performed.

3.3.3 SDS-PAGE analysis

SDS–PAGE was performed using a gel system (Atto). All protein gels contained 12% polyacrylamide. The recipe for 12% resolving gel is shown in Table 3.3 and 5% stacking gel is shown in Table 3.4.

The gel assembly was prepared using glass plates and 20 mL of resolving gel was poured between the plates and allowed to set for 20 minutes. Once set, unpolymerized acrylamide was washed off using Milli Q water. 8.0 mL of stacking gel was poured over resolving gel, 14 well-comb was inserted and the gel was allowed to set for 5 minutes.

Samples to load on the gel were prepared by adding 6X Laemmli buffer (30% 2-mercaptoethanol, 12% SDS, 10% glycerol, 0.1% bromophenol blue, 440 mM Tris pH=8.0) to a final concentration of 1X, then boiling at 95^{o}C for 5 min. Gels were run at 200 V for 90 minutes in 1x SDS buffer (25mM Tris, 192 mM glycine and 0.1% SDS) and 5 µL pre-stained pink protein standards (Bio-Rad, USA) was run along with samples in each gel.

3.3.4 Coomassie Blue staining

Gels were stained with Coomassie blue R250 stain (40% Methanol, 10% acetic acid, 0.1% Coomassie blue R250 (Fisher) for 30 minutes and destained with solution (40% Methanol, 10% acetic acid) until a colorless gels with bright blue bands appeared.

3.3.5 Silver staining

The PAGE gels was first put in fixing solution (50% methanol and 10% acetic acid) for 35 minutes then 5% methanol was used for incubation the gel for 15 minutes with changing the solution once in between. After fixing, the gel was washed thrice with milliQ water for 3 minutes each. The gel was then sensitized with 0.02% sodium thiosulphate for 2 minutes. Following sensitization the gel was washed three times with milli Q water for 20s each and then stained with staining solution (Silver nitrates 0.2% (Sigma), and 35 % formaldehyde solution for 30 min. Before developing, the gel was washed three times with milli Q water for 60s each. The bands were developed with developing solution (Sodium carbonate (Na_2CO_3) 3%, 37% formaldehyde solution and Sodium thiosulphate ($Na_2S_2O_3.5H_2O$) 2%) within 10 minutes. The development was stopped with stop solution (Na_2-EDTA) 1.4% once the bands were clearly visible. The gel was then washed once with 20% ethanol before drying the gels.

Table 3.3. Composition of 12% resolving gel for SDS-PAGE analysis.

Component	Volume
Distilled water (dH_2O)	6.3 mL
30% Acrylamide solution	8.0 mL
1.5 M Tris pH=8.0	5.2 mL
10% SDS	200 µL
1% APS	200 µL
TEMED	60 µL
Total	Approx. 20 mL

Table 3.4. Composition of 5% stacking gel for SDS-PAGE analysis

Components	Volume
Distilled water (dH2O)	5.5 mL
30% Acrylamide solution	1.5 mL
1.0 M Tris pH=6.8	1.0 mL
10% SDS	80 µL
1% APS	80 µL
TEMED	30 µL
Total	Approx. 10mL

3.2. Biophysical characterization of RGP of guar

3.2.1. Analytical Size Exclusion Chromatography

Analytical size exclusion chromatography (SEC) was done by ÄKTA Protein Purification System (GE Healthcare) using Superdex^{TM} 200, HiLoad^{TM} 16/60 prep grade column. The column was equilibrated with a buffer containing 25 mM Tris-HCl, pH 8.0 and 100 mM NaCl. The eluted purified protein was buffer exchanged into 25 mM Tris, pH 8.0 and

100 mM NaCl and then concentrated upto approximately 200 µl (8 mg/ml) using a 10 kDa cutoff Amicon Ultra-15 centrifugal concentrator (Merck, USA) and injected into the column at a flow rate of 0.5 ml/min. The molecular weight standards used were from Sigma – Aldrich, Ve/Vo obtained for each standard was plotted against the log of the molecular weight to generate a standard curve as per the manufacturer's instructions (Ve is elution volume and Vo is void volume). The collected fractions were evaluated by 12% SDS PAGE electrophoresis.

3.2.2. MALDI-TOF MS/MS and Peptide mass fingerprinting analysis

Molecular mass determination and identification of purified CtRGP protein (4 mg/ml) was used for Matrix-assisted laser desorption /ionization (MALDI-TOF MS/MS) and peptide mass fingerprinting (PMF). This analysis was done at Sandor Proteomics Ltd., Hyderabad. MALDI-TOF/TOF MS Bruker Daltonics ULTRAFLEX III instrument was used for the analysis. The sample given in I D gel protein band was diced to small pieces and placed in new Eppendorf tubes, was trypsin digested and the peptides obtained were mixed with cyano-4-hydroxycinnamic acid (HCCA) matrix in 1:1 ratio and the resulting 2µl was spotted onto the MALDI plate. After air drying the sample, it was analyzed on the MALDI TOF/TOF ULTRAFLEX III instrument and further analysis was done with FLEX analysis software for obtaining the peptide mass fingerprint. The masses obtained from the peptide mass fingerprint were submitted for Mascot search in "CONCERNED" database for identification of the protein.

3.2.3. Circular Dichroism

Circular Dichroism (CD) studies of CtRGP protein was done by Applied Photophysics (Chirascan^{TM} UK) CD spectroscopy equipment having a Peltier temperature control. The protein sample was prepared at a concentration of 0.5 mg/ml in 10 mM phosphate buffer, pH 7.4. All CD measurements were taken in the spectral range of 190-260 nm at 20℃, using an

optical path length 0.05 mm with scanning speed 100nm/min, a band width of 1nm and response time of 1 sec. The final spectrum was result of 20 spectra which were accumulated, averaged and corrected from the baseline for buffer solution contribution and normalized to circular ellipticity (θ) [7]. Helical parameters were calculated by using K2D2 online server.

3.2.3.1 Thermal denaturation

Thermal denaturation analysis of the purified CtRGP protein was done by recording the circular ellipticity (θ) values at a fixed wavelength of 222 nm while the sample heated from 10℃ to 90℃ [7] and the thermal denaturation curve was plotted according to Greenfield [67, 70] The denaturation curve was normalized to apparent fraction folded, according to the equation $[\theta] = [(\varepsilon_F - \varepsilon_U)\alpha] + \varepsilon_U$, where [θ] was the ellipticity at any temperature, ε_F is the ellipticity when the protein is fully folded, and ε_U is the ellipticity when the protein is totally unfolded, as previously described [67, 70]. The melting temperature of CtRGP protein was calculating multiple thermal melting program from Global 3^{TM} analysis software.

3.2.3.2. Chemical denaturation

Chemical denaturation analysis of CtRGP protein was done by the method described by Greenfield and coworker [68]. The protein samples (0.5 mg/ml in 10 mM phosphate buffer, pH 7.4) were incubated with increasing concentrations of denaturant guanidium hydrochloride (0-6 M) and urea (0, 2, 4, 6 and 8 M) (Sigma) at 4℃ for 16 hrs. The circular ellipticity (θ) values were recorded in the spectral range of 190-260 nm wavelengths at 20℃ and the spectra were analyzed by using K2D2 online server [122].

3.2.4. Intrinsic Fluorescence spectroscopy

The tryptophan intrinsic fluorescence spectroscopy of purified CtRGP protein was determined with the help of Varian Cary Eclipse Fluorescence Spectrophotometer. The

CtRGP protein concentration was 0.35 mg/ml in 25mM Tris pH 8.0 and 100mM NaCl buffer. Buffer solutions of pH values 4, 5, 6, 7 and 8 M were prepared by the addition of 0.1 M NaOH or 0.1 M HCl using a TIMPLE-85 pH meter and the effect of ionic strength on protein fluorescence in water was observed by addition of varying amounts (0.1M to 1.0M) of NaCl. Fluorescence spectra was recorded at 300-400nm.

3.2.5 Isothermal Titration Calorimetric binding Assay

Isothermal titration calorimetric (ITC) binding assay of purified CtRGP protein with its substrate UDP-glucose and two inhibitors, diisothiocyanostilbenedisulphonic acid (DIDS) and dithiobisnitrobenzoic acid (DTNB) using iTC200 Microcal calorimeter (GE Healthcare) by following the method of Wiseman and coworker [161]. The solutions of the CtRGP, UDP-glucose and each of the two inhibitors were prepared in a buffer solution containing 25mMTris, pH 8.0 and 100mM NaCl. The sample cell was filled with 225µl of 300 µM CtRGP protein solution and the syringe of the calorimeter was loaded with 60 µl of 0.5 mM UDP-glucose (Abcam, USA) as a ligand. In the second and third experiments, the test cell was filled as above but syringe was loaded with 60 µl of 0.5 mM DIDS (Sigma, USA) and 0.5 mM DTNB (Sigma, USA) as ligands, respectively. The reference cell was filled with protein free buffer solution containing 25mMTris, pH 8.0 and 100mM NaCl. Twenty injections of each protein-ligand mixture (2µl per injection) were used for each binding assay. The time interval between two injections was set at 100 sec. All titrations were performed at 20℃ temperature and 700 rpm stirring rate. The enthalpy of each binding assay was recorded. The data so obtained were fitted via non-linear least squares minimization method to determine binding stoichiometry (n), binding dissociation constant (K_d) and change in enthalpy of binding (ΔH). The data analysis was done using Origin software version 7.0. The changes in free energy (ΔG) and entropy (ΔS) were calculated by using following equations

$$\Delta G = RTlnKd \ldots\ldots\ldots\ldots\ldots\ldots \text{ (I)}$$

$$\Delta G = \Delta H - T\Delta S \text{.....(II)}$$

The value of the binding dissociation constant (K_d) was used to compare the relative binding of substrate and inhibitors with the purified CtRGP protein.

3.3. Crystallization trials of purified reversibly glycosylated polypeptide of guar

Crystallization trials were carried out both manual and commercial crystallization screens (Hampton Research Inc.Aliso Viejo, CA). Homogeneously purified protein was obtained from IMAC. Dialysis buffer (25mM pH=8.0 and 100mM NaCl) was used for dialysis. Dialyzed purified protein was concentrated upto 10mg/ml in an Amicon Ultra concentrator (Merk, USA). Crystallization screens (Crystal Screen I & II, PEG/Ion I & II, Index, Salt and Crystal screen cryo) were procured from Hampton Research, Germany. For optimization reagents were made by highest purity ACS grade chemicals. To remove the insoluble impurity or particles solutions were filtered through 0.22µ filter. The prepared reagents were maintained 4℃. PEG solutions were prepared by overnight stirring and stored in light protected bottles.

3.4. *In silico* analysis of CtRGP Protein sequence

The nucleotide and amino acid sequences of CtRGP protein have not been submitted to NCBI database till now. The full length 1455 bp sequence of clone code >lds2c.pk011.p19:fis for guar RGP was obtained from DuPont Pioneer Hi breed USA. The complete sequence of the CtRGP cDNA has been given below Fig.3.2. It consists of 1,455 bp and encodes an open reading frame of 353 amino acid residues using ORF finder tool of NCBI.

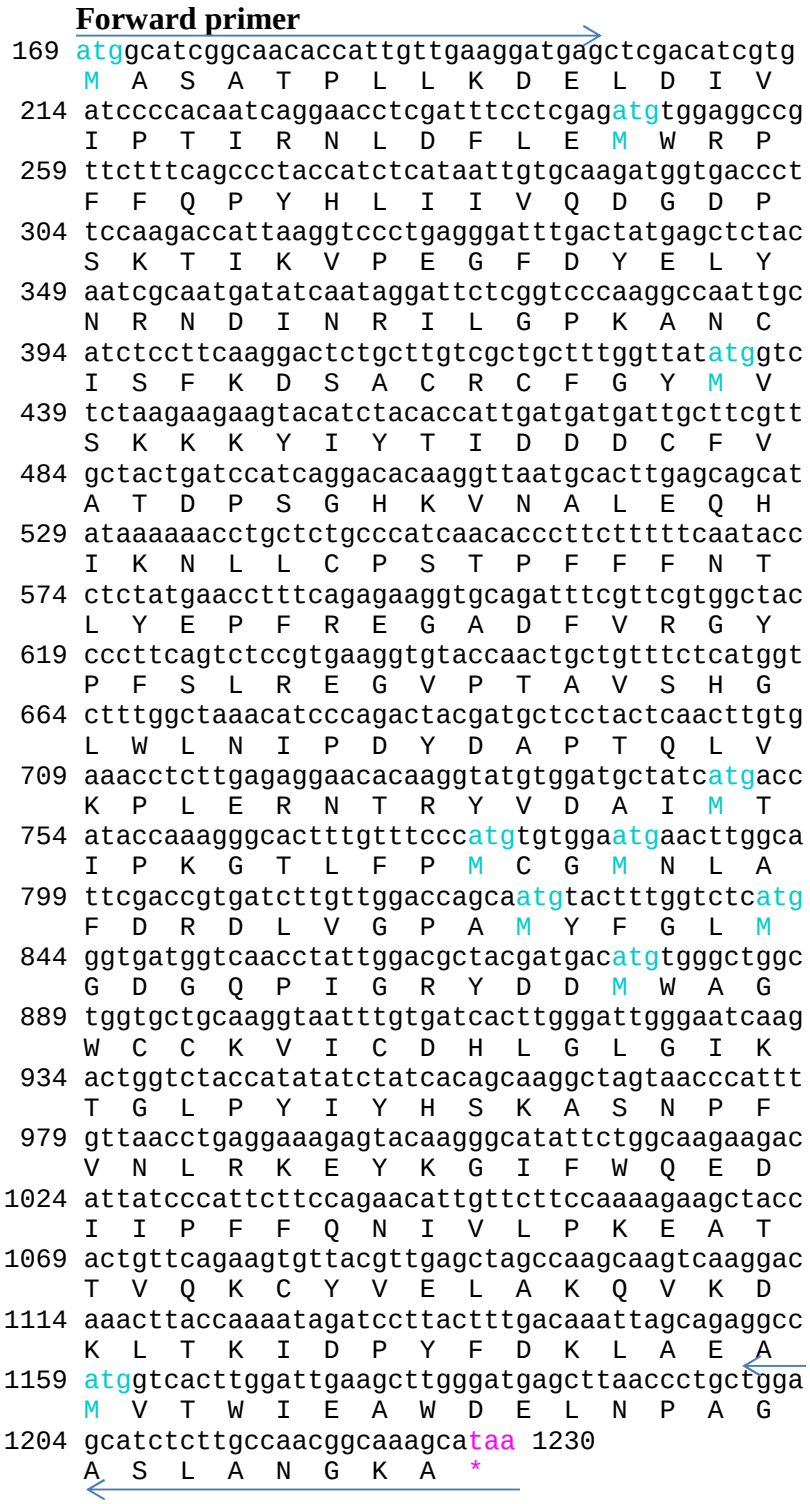


Fig. 3.2. The nucleotide sequence of CtRGP encodes an open reading frame of 353 amino acid residues using ORF finder tool of NCBI in one letter code. The position of forward and reverse primer sequences are indicated by the arrow headed lines above the nucleotide sequence.

The above protein sequence was used for sequence phylogenetic analysis, molecular modeling and Molecular docking.

3.4.1. Sequence and phylogenetic analysis

The ExPASy's ProtParam server (http:// expasy.Org/cgi-bin/Protparam) was used for physico-chemical characterization and calculation of theoretical isoelectric point (pI), molecular weight, total number of positive and negative residues, extinction coefficient [65], instability index [72], aliphatic index [8] and grand average hydropathy (GRAVY) [8]. The secondary structure prediction was done by the PSIPRED (Psi-BLAST based secondary structure prediction) server. Secondary structure of this protein of ~42 kDa molecular weight was predicted using the FASTA sequences in GOR IV [60] and SOPMA [62]. The functional annotation and domain finding was determined by InterProscan tool. The phylogenetic tree was constructed by using MEGA 6.0 SERVER as described by Tamura and coworker [154]. Hydropathy plot was generated by Kyte-Doolitle hydropathy plot server [92].

3.4.2. Molecular modeling

The crystal structure of RGP protein has not been determined, so the substrate specificity of this CtRGP protein to its substrate UDP-glucose was investigated by computer modeling using the threader tool by Scratch Protein predictor tool (Cheng et.al.2005). The 3Dpro of SCRATCH server [26] (http://www.igp.uci.edu) was used to predict the three dimensional (3D) structure of CtRGP protein because CtRGP protein sequence have not good structural templates for homology modeling. Here tertiary structure prediction methods are evaluated by the Critical Assessment of Structure Prediction (CASP).

Refinement of predicted CtRGP model by 3Drefine tool [11] generated five models. The model having energy was evaluated for structural analysis by SAVES server. Refined models were evaluated for their stereo-chemical quality by PROCHECK [97]. The model having minimum number of residues in the disallowed region was selected and further validated by VERIFY-3D to check the compatibility of 3D model[16, 101]. ERRAT plot was also used for the structural validation which shows the value of error for each amino acid [27].

The error indicated by ERRAT plot was rectified by loop refinement tool by changing the loop conformation of the model. This process was continued until each amino acid in the protein reached below 95% cut –off value in ERRAT plot. Again stereo-chemical quality of model was evaluated by PROCHECK. The final 3D model of the CtRGP protein was then evaluated for potential errors by ProSA-web server [160].

PDBsum is a web-based database providing a largely pictorial summary of the key information on each macromolecular structure deposited at the Protein data bank (PDB) was used for CtRGP. It includes images of the structure, annotated plots of each protein chain's secondary structure, detailed structural analyses generated by the PROMOTIF PROGRAM [96].

3.4.3. Molecular docking

Molecular docking of CtRGP protein with substrate UDP-glucose, and two inhibitors DIDS and DTNB was carried out using Autodock 4.2.0. [109] All the docking files were prepared using MGLTool 1.5.4 (http:// mgltools.scripps.edu/) by properly assigning gasteiger charges. The linear molecules of UDP-glucose, DIDS and DTNB were first generated using a chemical drawing tool Avogardro 1.1.0-win 32 (http://avagadro.openmolecules.net/) [74]. PDB 2Q3E and 4CLZ are the coordinates of the substrate UDP-glucose and inhibitor DIDS, respectively. Grid maps were generated using Autogrid having the grid box dimensions 60 x 60 x 60 Å with the grid spacing 1.000Å. Each job comprised of 50 independent runs. Molecular docking simulations were performed using Lamarckian genetic algorithm and a maximum of 2.5 million energy evaluations were performed with a maximum of 250,000 generations. Results were sorted according to their energy profile. The interaction between protein and ligands was visualized by PyMoL software [36]. The validation of docking results was done by docking the native substrate UDP-glucose into the protein coordinates independently.

4. RESULTS

4.1 PCR amplification and construction of the recombinant plasmid

The present work was started with a cDNA construct containing a full length sequence of guar CtRGP gene in cloning vector pBlueScript SK (+). This cDNA construct was obtained from a guar cDNA library, the construction of which has already been reported by Dhugga and coworker [41]. The 1062 bp cDNA sequence from the above construct was amplified by using RGPf1FOR1 and RGPf1REV1, forward and reverses primers, respectively. The resulting amplicon, after its verification by gel electrophoresis on 1% agarose gel (Fig. 4.1), was inserted in the expression vector pET 29a (+) using *Nde*l and *Sal*l restriction sites for subcloning. The recombinant plasmid was introduced into *E. coli* DH5α cells by transformation. The kanamycin resistant (Km^R) transformants were selected (Fig. 4.2 & 4.3) and the recombinant plasmid construction was confirmed by colony PCR and restriction analysis with *Nde*l and *Sal*l restriction enzymes (Fig. 4.4 & 4.5) and DNA sequencing (Fig. 4.6 & 4.7). The recombinant plasmid was named as pET29a (+).CtRGP (Fig. 4.8) and used in the further experiments for obtaining CtRGP protein.

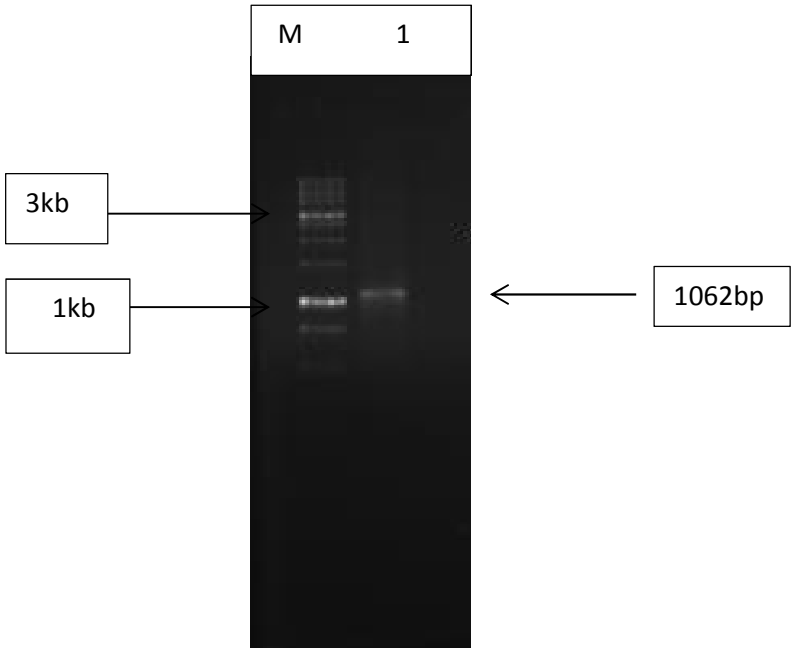


Fig. 4.1: Amplification of *Ct*RGP gene on agarose gel. Lane M: 1 kb DNA ladder (Biochem life science) and lane 2: 1062bp *Ct*RGP gene.

(a) **(b)**

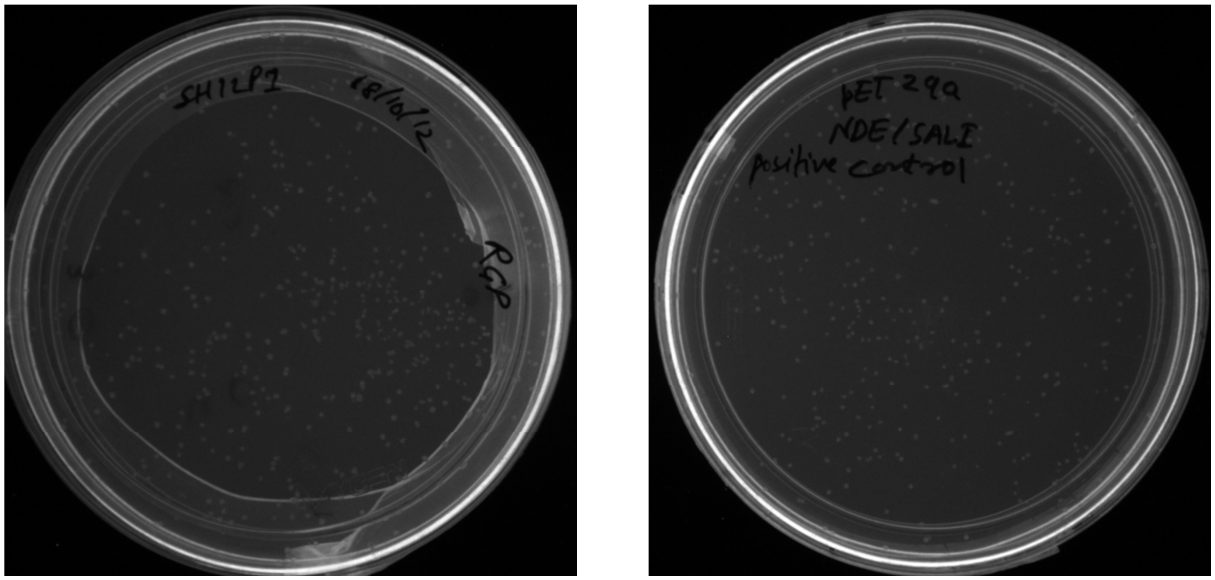


Fig. 4.2: Transformation results of CtRGP gene in *E.coli* DH5α (a) Transform colonies of CtRGP and (b) Positive control.

(a) **(b)**

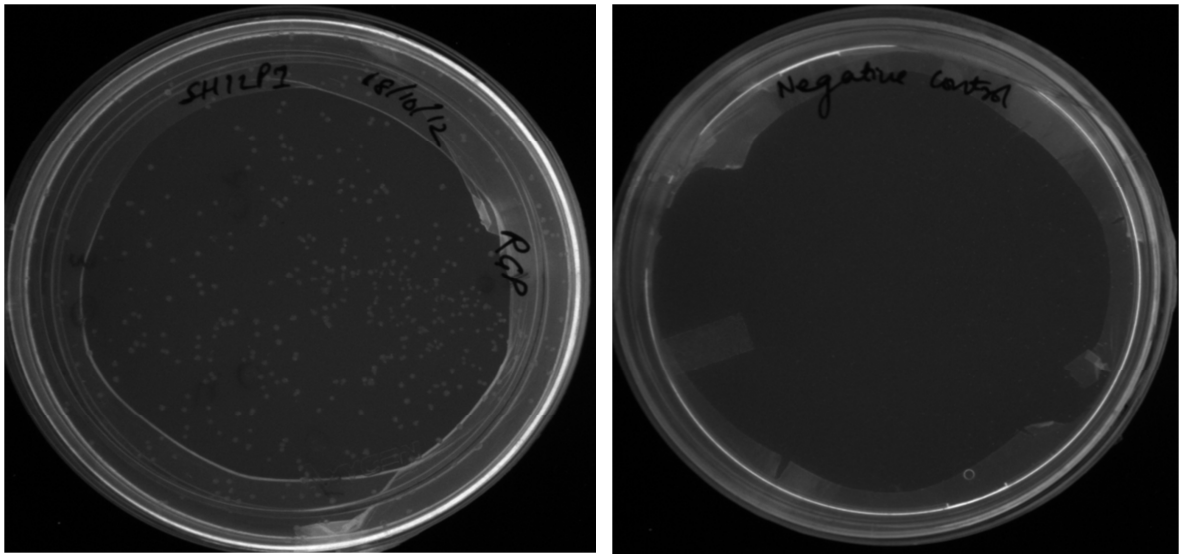


Fig. 4.3: Transformation results of CtRGP gene in *E.coli* DH5α (a) Transform colonies of CtRGP and (b) Negative control.

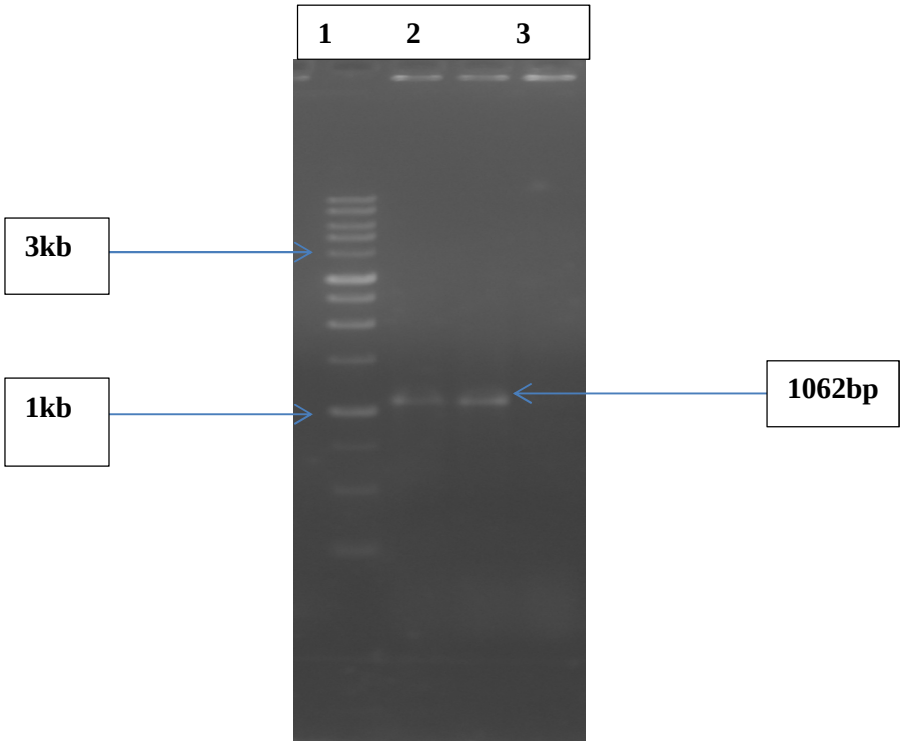


Fig. 4.4: Showing sub-cloning confirmation of CtRGP gene in pET 29a (+) vector by colony PCR. Lane 1: 1kb DNA marker, Lane 2: amplified 1062bp CtRGP gene of transform colony I and Lane 3: amplified 1062bp CtRGP gene of transform colony II.

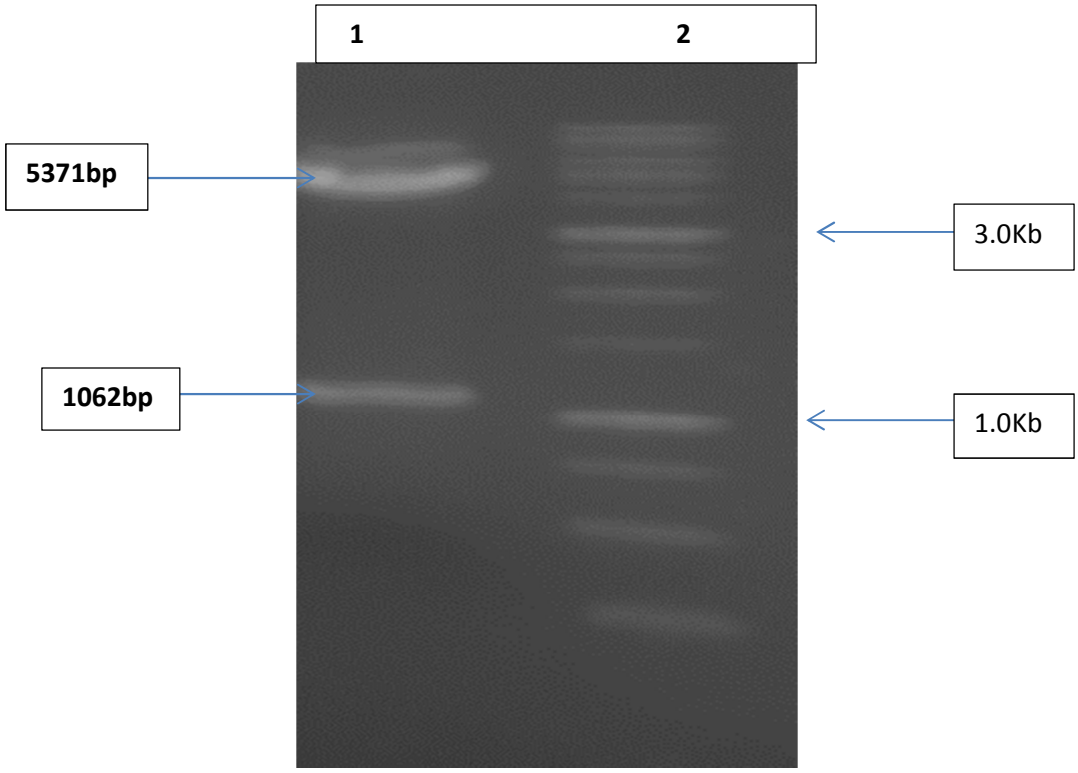


Fig. 4.5: Cloning confirmation by restriction digestion. Lane 1: Restricted fragments 5371bp (pET29a vector) and 1062bp for CtRGP gene, and Lane 2: 1Kb DNA marker (Biochem life science).

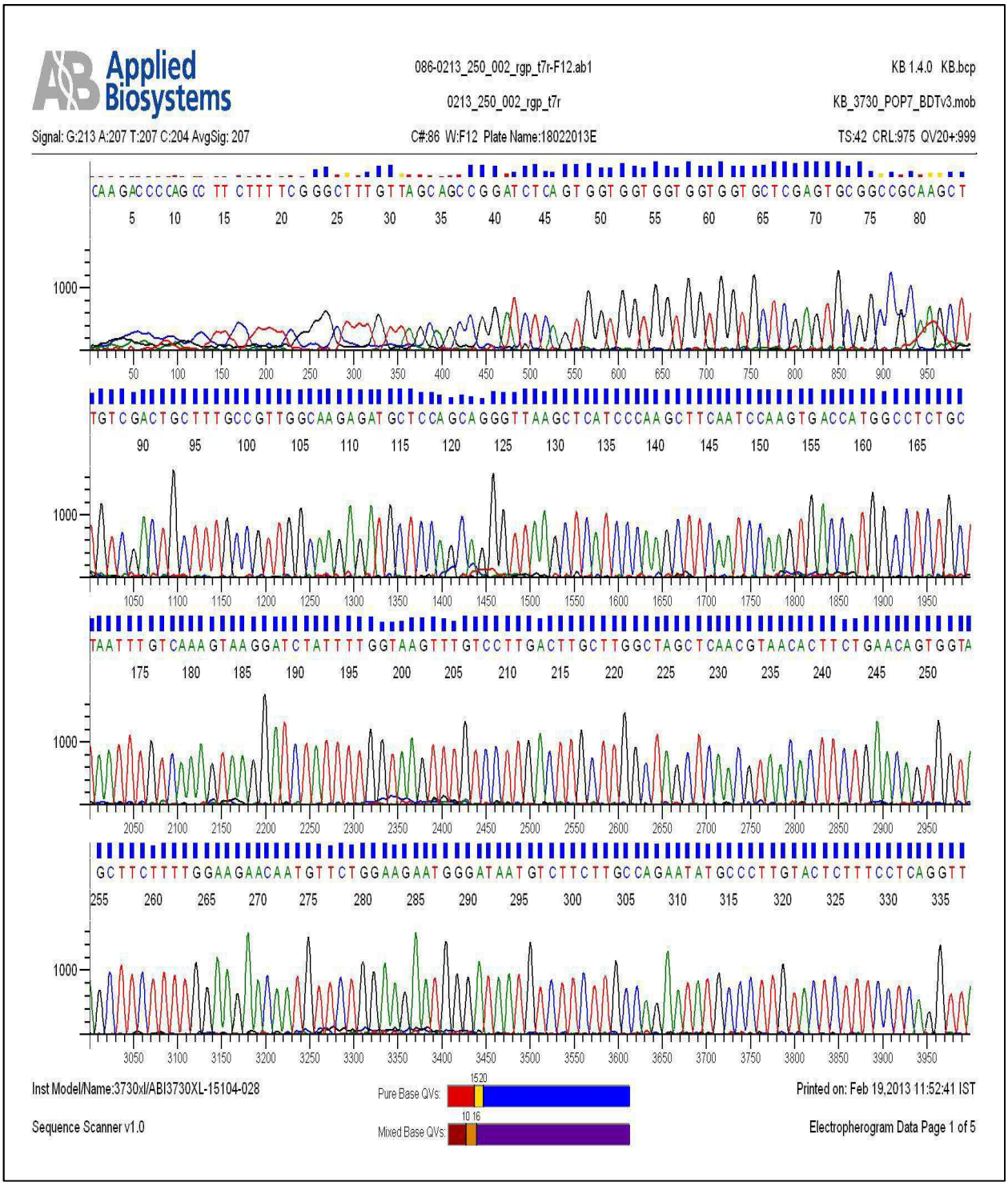


Fig. 4.6: Chromatogram showing DNA sequencing of CtRGP cloned gene by using T7 forward primer.

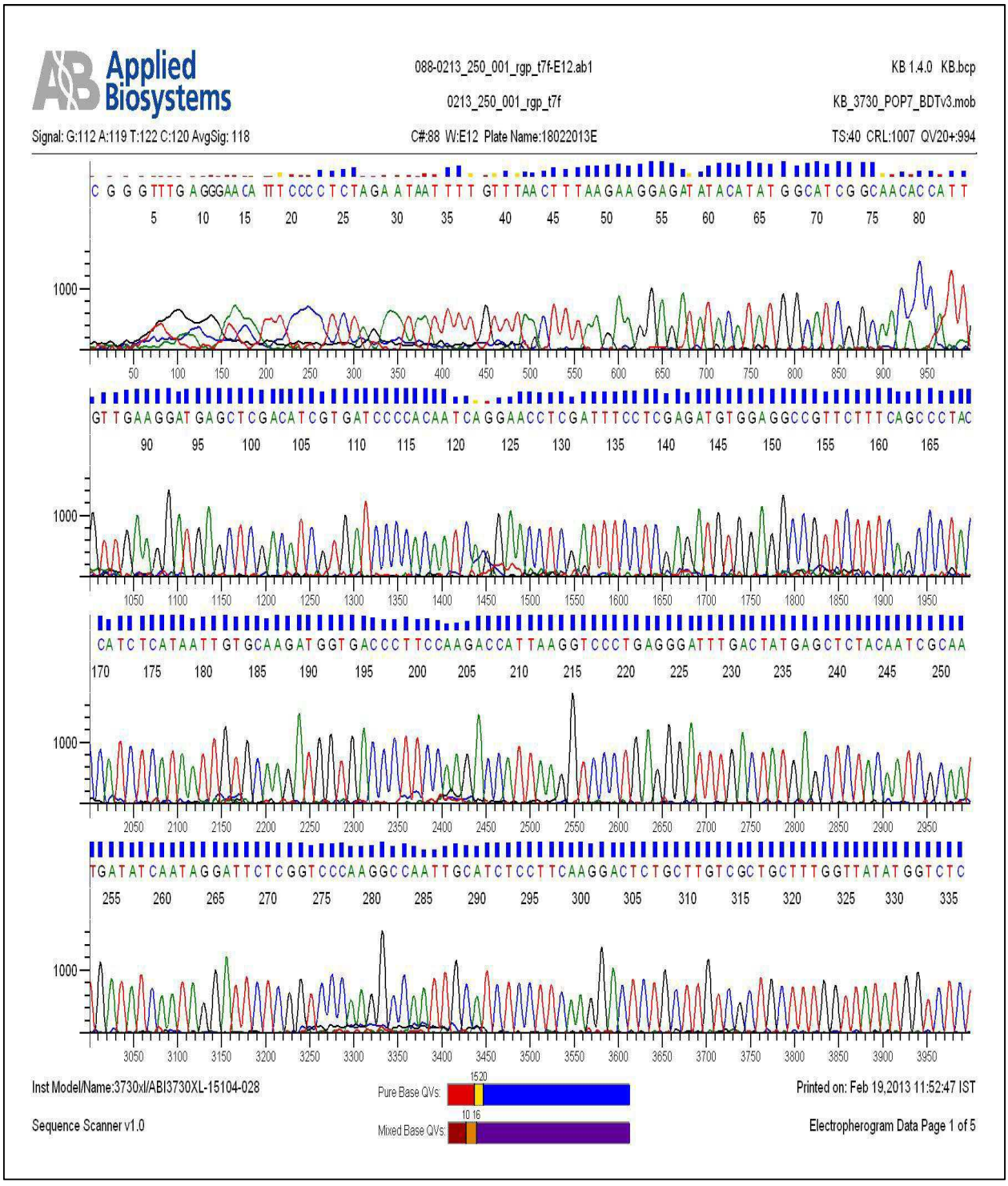


Fig. 4.7: Chromatogram showing DNA sequencing of CtRGP gene by using T7 reverse primer.

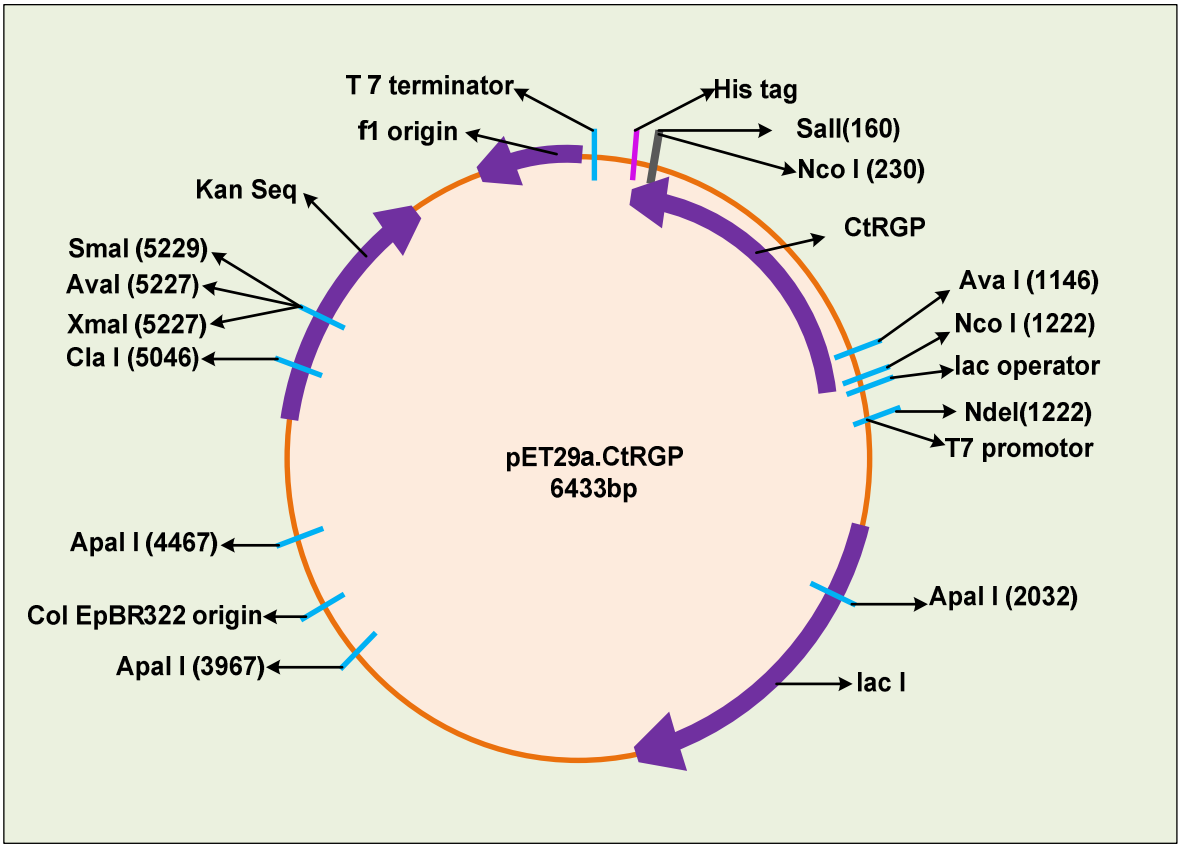


Fig. 4.8: Clone construct of CtRGP gene in pET29a(+) cloning vector named as pET 29a (+).CtRGP .

4.2 Production and Purification of Recombinant CtRGP Protein

A plasmid preparation from the *E. coli* DH5α cells containing the recombinant plasmid pET29a (+).CtRGP was used to transform the expression host *E. coli* BL21 (DE3) cells using the kanamycin resistance marker of the recombinant plasmid. One of the kanamycin resistant transformant, after confirmation of the presence of recombinant plasmid in it, was used for overproduction of CtRGP protein. In this *E. coli* strain, CtRGP protein is produced as a fusion protein with 6X His-tag at the C-terminal end. Optimal expression of CtRGP-His-tag fusion protein was achieved at 18^{o}C temperature and 14-16 hours incubation after induction (Fig. 4.8). Immobilized metal affinity chromatography (IMAC) of the cell lysate yielded sufficiently pure CtRGP protein. A single band of CtRGP protein corresponding to 42 kDa was observed by SDS-PAGE in the elution fractions (Fig. 4.9). The purity of the purified protein was checked by silver staining and the protein was found to be more than 95% pure

(Fig. 4.9). The purified protein band from SDS-PAGE was subjected to Western blotting and the blotted band showed agglutination with pea anti-RGP purified antibody confirming that the purified protein to be an RGP protein (Fig. 4.10).

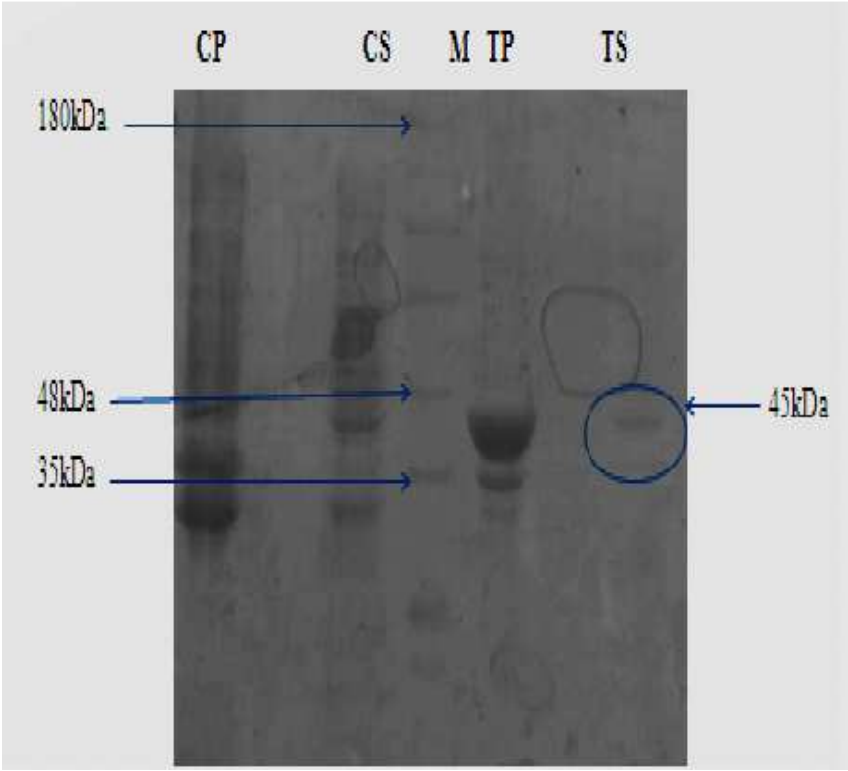


Fig. 4.9: SDS-PAGE analysis of small scale recombinant RGP in *E.coli.* strain BL21(DE3) showing: uninduced bacterial pellet (CP), uninduced supernatant (CS) , pre-stained protein size marker (M), induced bacterial pellet (TP) and induced supernatant (TS).

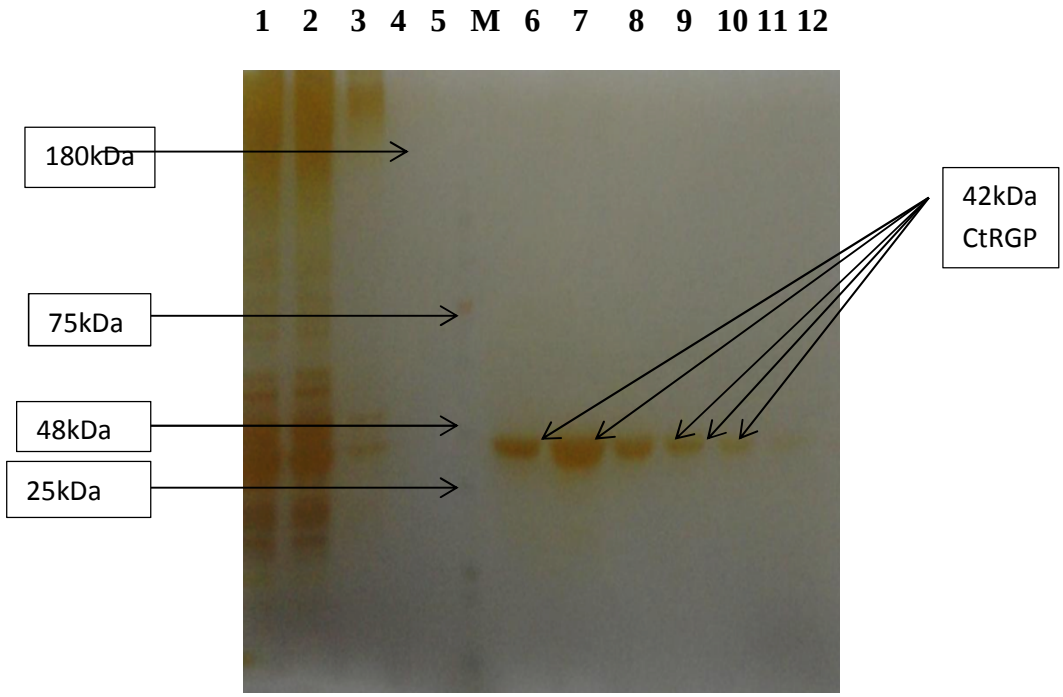


Fig. 4.10 : Silver staining of purified recombinant RGP from guar. M=Pre-stained protein marker 1=Pellet; 2=Supernatant; 3=Supernatant FLT; 4=Wash I FLT (50mL); 5=Wash II FLT-1(10mL); 6=Elution-1 (7mL); 7= E-2 (5mL); 8=E-3 (5mL); 9=E-4 (5mL); 10=E-5 (5mL); 11=E-6 (10ml) and 12=E-7 (13mL).

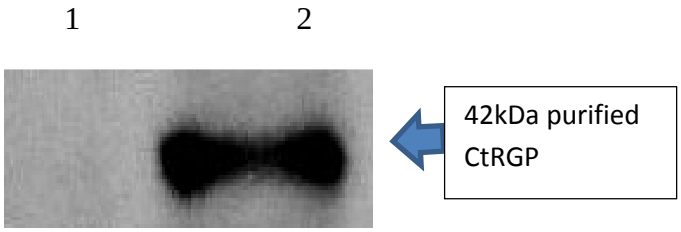


Fig 4.11: Western blot result for the confirmation of purified CtRGP. Lane 1= negative control (alpha-galactosidase from coffee plant); 2= 20µg of purified CtRGP.

4.3 Oligomeric State of CtRGP Protein

The oligomeric state of the purified CtRGP protein was determined by analytical size exclusion chromatography. When the purified protein (8mg/ml) was injected into the Superdex™ 200 column, a single sharp peak was obtained between 55 to 60.3 ml elution volumes (Fig. 4.12). By comparison with molecular mass standards, this peak was found to correspond to an estimated molecular mass of 180 kDa (Fig. 4.13). As the molecular mass of monomeric form of CtRGP protein has been found to be 42 kDa (Fig. 4.14), it can be concluded from this result that the purified CtRGP protein is in tetrameric form.

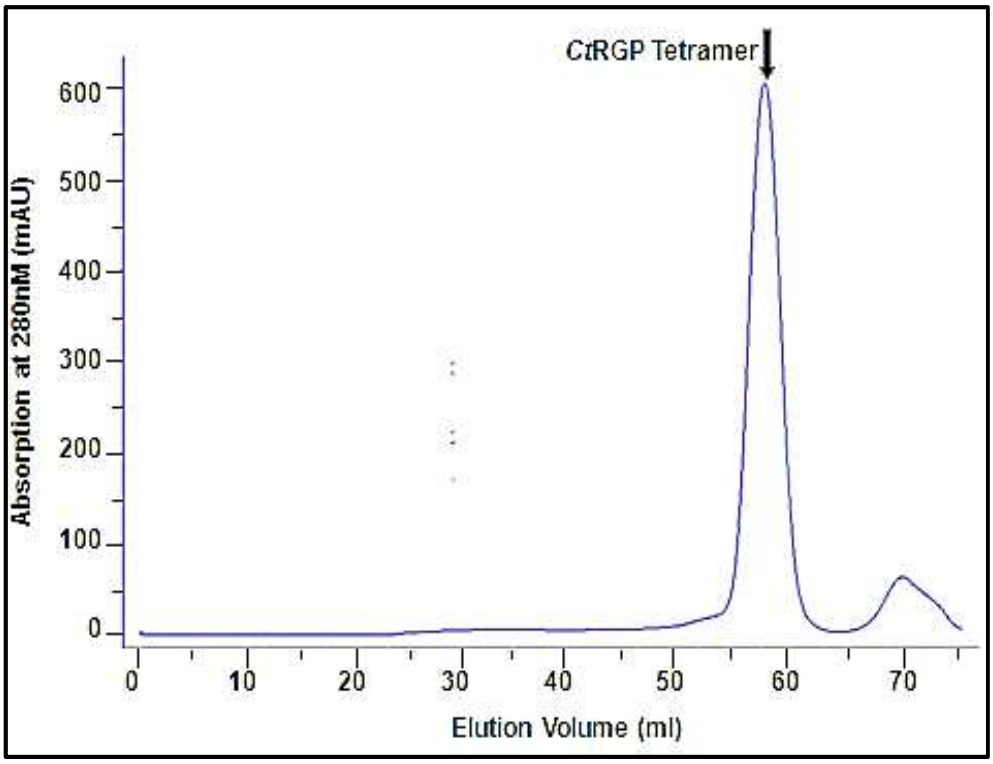


Fig. 4.12 : Analytical size-exclusion chromatography result of CtRGP in a Superdex-200 High load™ 16/60 prep grade column equilibrated with a buffer containing 25mM Tris HCl pH=8.0 and 100mM NaCl.

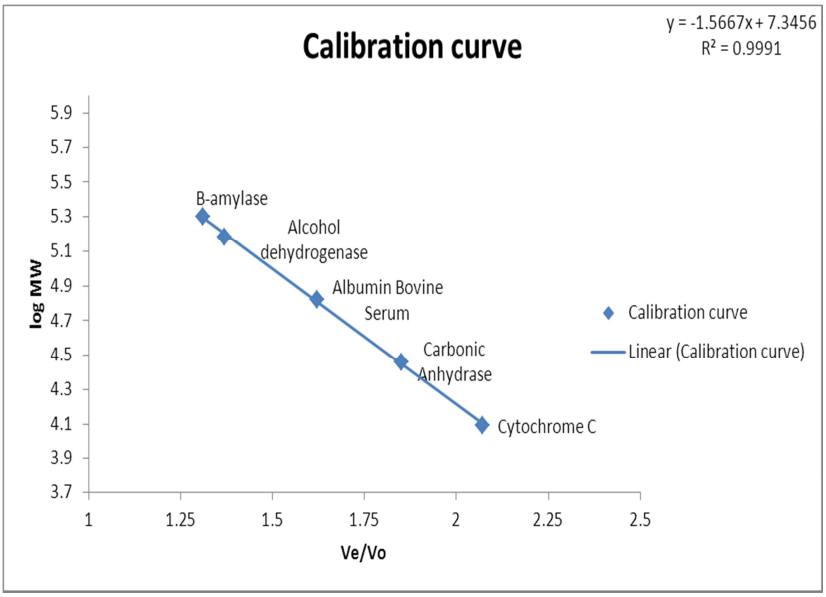


Fig. 4.13: Calibration curve for analytical gel filtration of purified CtRGP protein.

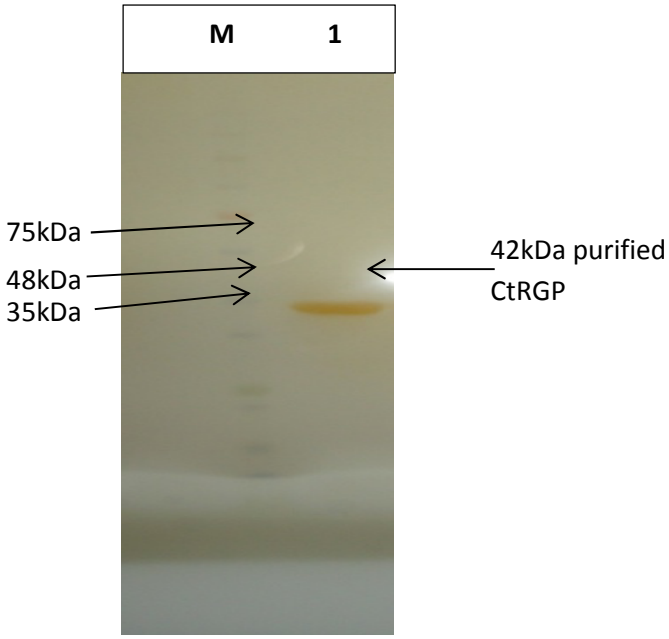


Fig. 4.14: Silver stained 12% SDS-PAGE gel after analytical gel filtration. M= Pre stained protein marker and Lane 1= Purified CtRGP protein

4.4 MALDI-TOF MS/MS and Peptide mass fingerprint analysis

The purified CtRGP protein solution (4mg/ml) was subjected to MALDI-TOF mass spectroscopy and peptide mass fingerprinting (PMF). The MALDI-TOF analysis showed the molecular mass of the CtRGP protein to be 42, 123.21 Da (Fig. 4.15). For PMF analysis, the CtRGP protein was digested with trypsin and the resulting digested fragments were sequenced. These sequences were used to fish out the homologous protein sequences from the available with the help of Mascot search engine "CONCERNED" database for identification of the protein. A total 20 hits were obtained (Table 4.1). Most of these proteins are the enzymes involved in the plant cell wall synthesis. The query sequence showed maximum homology score of 117 for UDP-arabinopyranose mutase 1-like protein of *Glycine max.* This protein had nine regions of homology with the query protein. The first 20 amino acids of UDP- arabinopyranose mutase 1-like protein were identical to the first 20 amino acids of CtRGP protein (Table 4.2).

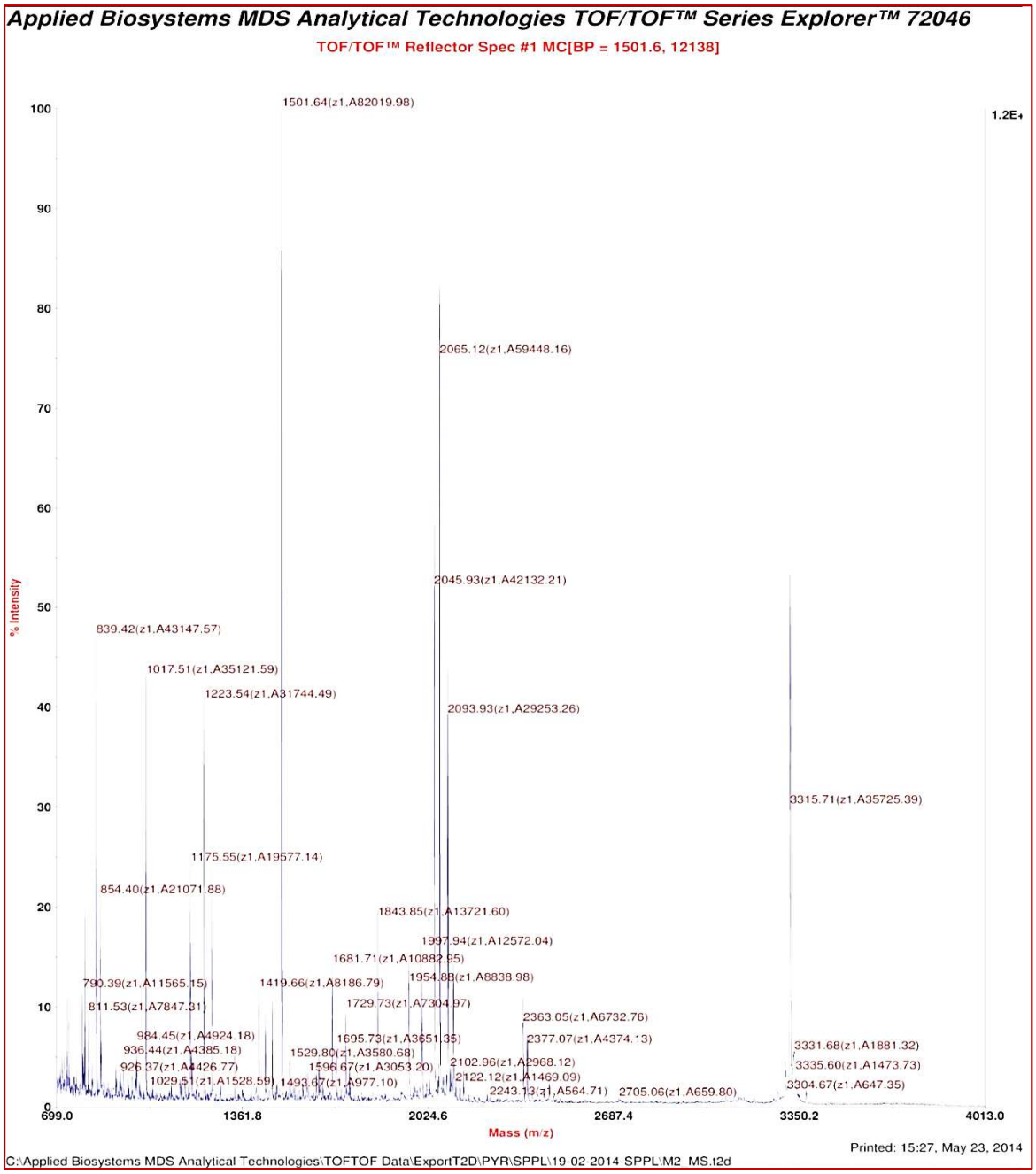


Fig. 4.15: Isotopic distribution mass spectrum of recombinant purified CtRGP protein. Deconvolution using the multi-protonated peaks confirmed the correct molecular mass of recombinant CtRGP (monoisotopic mass = 42132.21Da).

Table 4.1: Peptide mass fingerprint results show 20 hits of the matching sequences of trypsin digested purified CtRGP protein product.

S.no.	Accession	Mass	Score	Description
01	gi 356568813	40655	117	Predicted:UDP- arabinopyranosemutase1-like protein[Glycine max]
02	gi 356526536	40697	97	Predicted: alpha-1,4-glucan –prptein synthase[UDP-forming] 1-like protein [Glycine max]
03	gi 255648166	40613	97	Unknown [Glycine max]
04	gi 449438723	41143	97	Predicted: alpha-1,4-glucan-protein synthase [UDP-forming]2-like protein [Cucumis sativus]
05	gi 593332010	40737	80	Hypothetical protein PHAVU_008G028900g [phaseolus vulgaris]
06	gi 470136886	40931	79	Predicted: alpha-1,4- glucan-protein synthase [UDP-forming]2-like [Fragaria vesca subsp.vesca]
07	gi 567896626	41951	79	Hypothetical protein CICLE_v10020723mg [Citrus clementiana]
08	gi 568847778	41947	79	Predicted: UDP-arabinopyranose mutase 3-like [Citrus sinensis]
09	gi 502101859	40584	77	Predicted: UDP-arabinopyranosemutase 2-like [Cicer arietinum]
10	gi 225922209	53701	74	nucleocapsid protein [Influenza A virus (A/bar-headed goose/Qinghai/F/2007(H5N1)
11	gi 590631415	117811	67	E3 ubiquitin-protein ligase UPL7isoform 6 [Theobroma cacao]
12	gi 590631407	127639	65	E3 ubiquitin-protein ligase UPL7isoform 4 [Theobroma cacao]
13	gi 590631405	130557	64	E3 ubiquitin-protein ligase UPL7isoform 3 partial [Theobroma cacao]
14	gi 590631403	130113	64	E3 ubiquitin-protein ligase UPL7isoform 2 [Theobroma cacao]
15	gi 489711715	38571	64	Glucosyltransferase [Lactobacillus delbrueckii]
16	gi 590631400	132506	64	Ubiquitin-proteinligase 7 isoform 1[Theobromacacao
17	gi 18077708	41337	63	reversibly glycosylated polypeptide [Gossypium hirsutum]
18	gi 110559491	41333	63	GRP-like protein 2 [Gossypium hirsutum]
19	gi 168065267	113077	63	predicted protein [Phycomitrella patens]
20	gi 587860138	41966	62	Reversibly glycosylated polypeptide [Mous notabilis]

Table 4.2: Peptide mass fingerprint result shows the maximum 20 sequences of CtRGP protein sequence similarity with the predicted: UDP-arabinopyranose mutase1-like protein [Glycine max].

Observed	Mass.(expt.)	Mass.(calc)	Ppm	Start	End	Miss	Peptide
839.4024	838.3951	838.4007	-6.66	73	79	0	K.ANCISFK.D
1017.4970	1016.4897	1016.5403	-49.77	266	274	0	K.ASNPFBNLR.K
1251.6935	1250.6862	1250.6693	13.6	305	314	1	K.CYIELAKQVK.E
1501.6066	1500.5993	1500.6885	-59.42	51	62	0	K.VPEGFDYELYNR.N
1729.6812	1728.6739	1728.7786	-60.55	199	213	0	K.GTLFPMCGMNLAFDR.D
1843.8082	1842.8009	1842.9152	-62.00	48	62	1	K.TIKVPEGFDYELYNR.N
2065.0757	2064.0684	2064.1830	-55.52	1	20	0	M.ASATPLLKDELDIVIPTIR.M
2113.8713	2112.8640	2112.9864	-57.94	51	67	1	K.VPEGFDYELYNRNDINR.I
3315.6509	3314.6436	3314.7229	-26.03	156	185	0	R.EGVPTAVSHGLWLNIPDYDAPTQLVKPLER.N

4.5 Circular Dichroism (CD) spectroscopy of CtRGP Protein

The CD analysis of CtRGP protein was carried out under native and denaturing conditions to determine the stability of this protein. The protein solution was prepared in phosphate buffer as this buffer causes less noise as compared to other buffers. The SDS-PAGE analysis showed that the protein was quite stable at 4℃ temperature in phosphate buffer even upto a period of one week (Fig. 4.16). The CD spectrum of the native CtRGP protein exhibited negative peaks at 208 nm (-15548.7 [θ]) and 222 nm (-15827.9 [θ]) (Fig. 4.17). These peaks are typical of α-helix rich proteins. The secondary structure content prediction from the spectrum obtained for native CtRGP protein using K2D2 server resulted in 32.45% α-helices and 12.92% β-sheets. This result confirmed the correct folding conformation of purified CtRGP protein.Thermal analysis of CtRGP protein was done at 222 nm in the temperature range of 10-90^{o}C. The calculated melting temperature (T_m) for CtRGP protein was found to be 68.9 ^{o}C (Fig. 4.18). The CtRGP protein was treated with different concentrations of denaturants, GdHCl and urea at pH 7.4. Spectral measurements after 16 hours showed increase in molar ellipticity of the spectra at 4M and 6M concentrations of GdHCl and urea, respectively (Fig. 4.19 & 4.20).

The above results of CD analysis indicate that the CtRGP protein is stable upto 68.9 ^{o}C temperature and resistant to sufficiently high concentrations of denaturants, GdHCl (4M) and urea (6M).

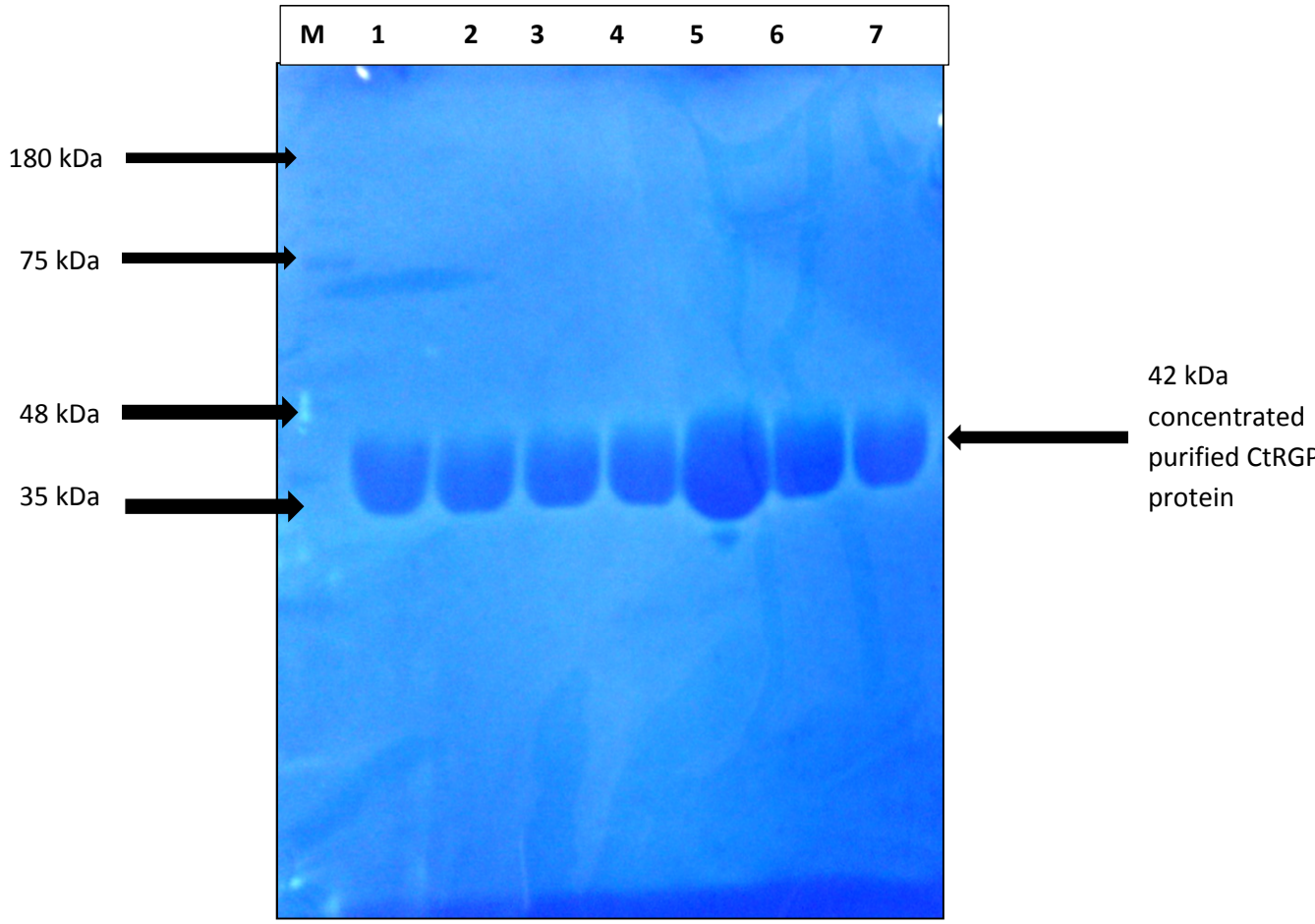


Fig. 4.16: Coomassie blue stained SDS-PAGE gel for the stability test of the purified CtRGP at 4℃ upto first day to seventh days. M =Pre-stained protein marker, Lane 1=1st day, Lane 2=2nd day, Lane3 =3rd day, Lane 4=4th day, Lane 5=5th day, Lane 6= 6th day and Lane 7=7th day

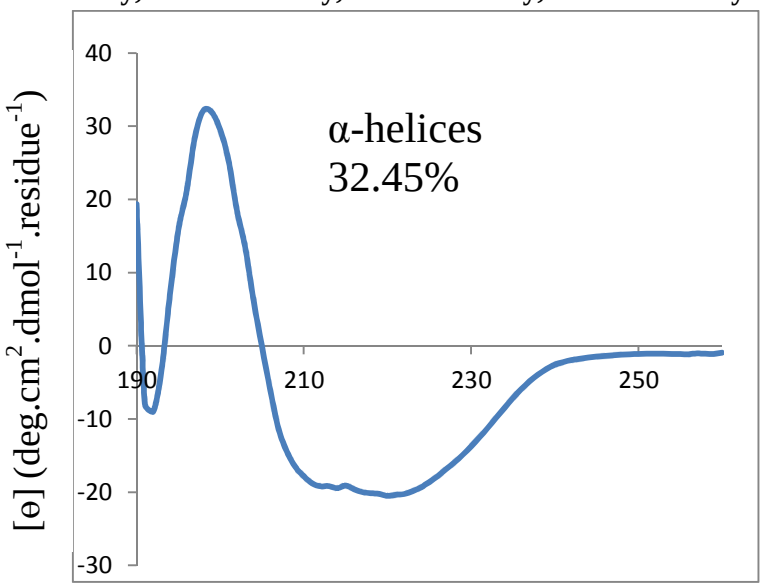


Fig. 4.17: Far-UV circular dichroism spectra of recombinant CtRGP protein. Experiments were performed at 20℃ with 200µl of protein at 0.5 mg/ml in 10mM of phosphate buffer at pH 7.4.

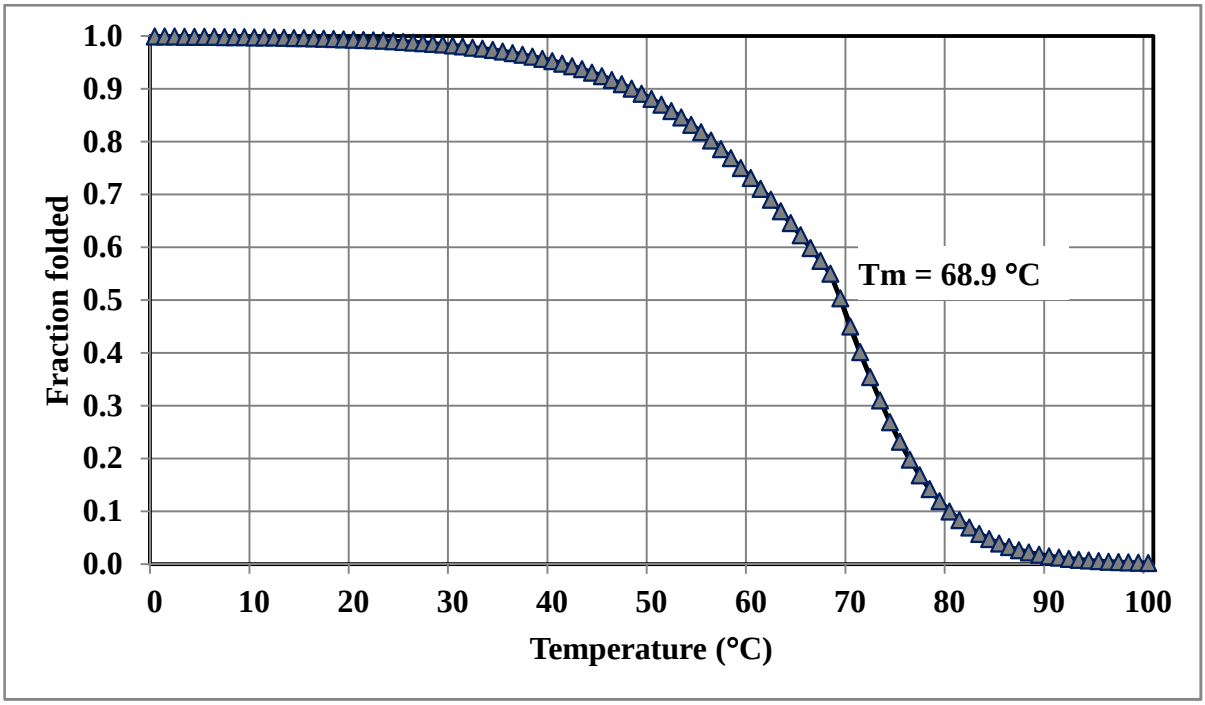


Fig. 4.18: Thermal denaturation curve of CtRGP protein obtained by circular dichroism spectra at a wavelength of 222 nm while heating from to 10^0C to 90°C.

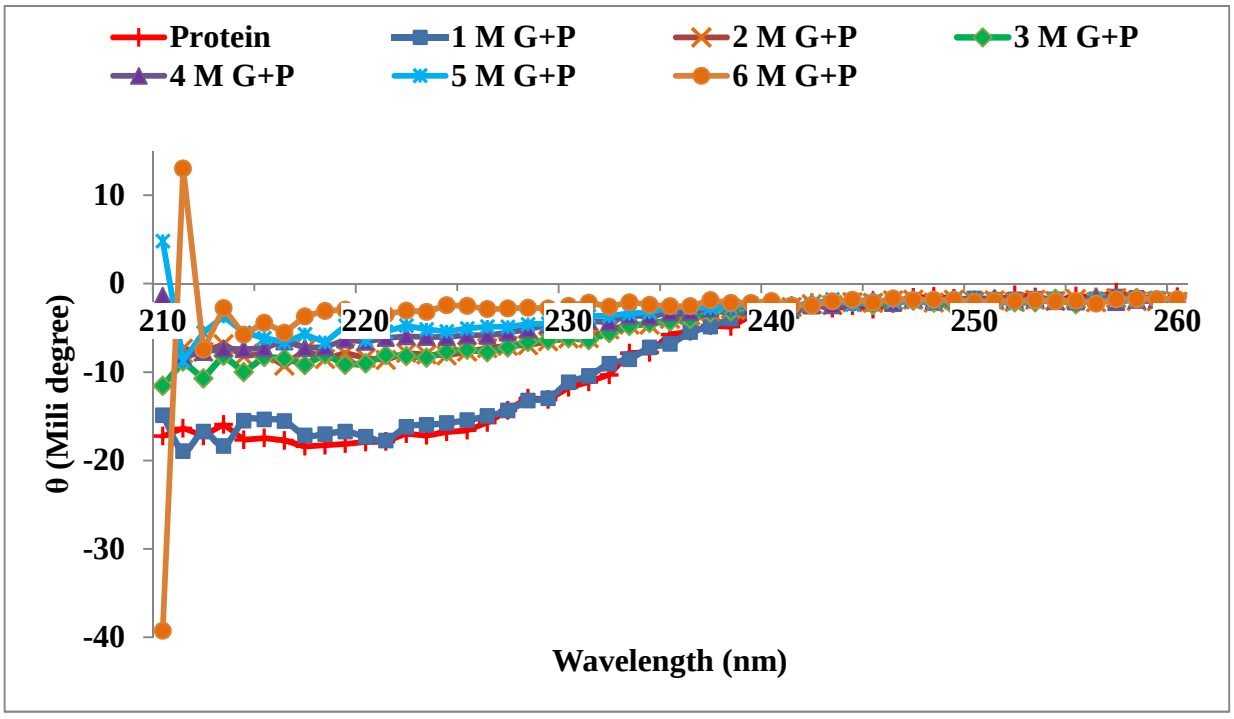


Fig. 4.19: Chemical denaturation of CtRGP protein using 0-6M GdHCl. Results shows maximum denaturation occurs at 4M of GdHCl.

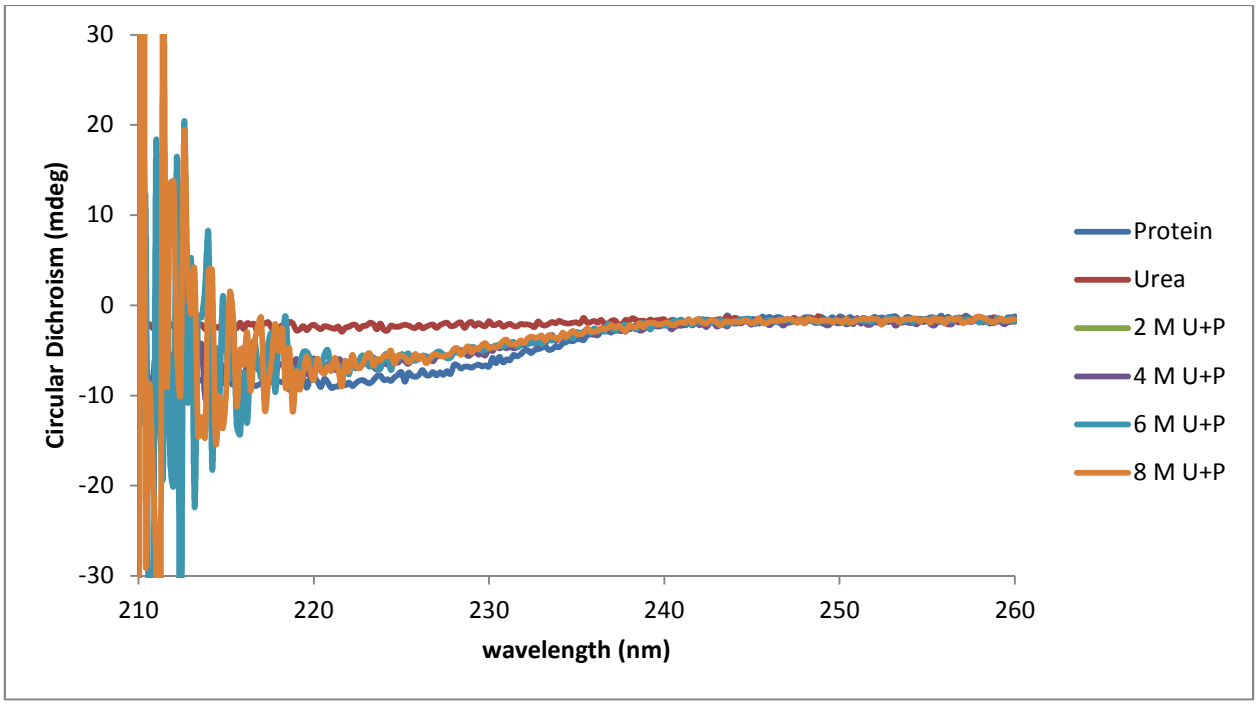


Fig. 4.20: Chemical denaturation study of CtRGP protein using concentration of Urea at 2M, 4M, 6M and 8M. CD Spectra showed final denaturation occurs at 4M concentration of Urea.

4.6 Intrinsic Tryptophan Fluorescence Spectroscopy

The fluorescence spectra of CtRGP protein at different pH values of 25 mM Tris and at different molar concentrations of NaCl showing changes in fluorescence intensity have been shown in Fig. 4.21 &. 4.22, Emission maximum for the CtRGP protein was observed at pH 8.0. A decrease in fluorescence intensity was found at pH 4, 5, 6 and 7. No significant changes in fluorescence intesity of CtRGP protein were observed with the increase in salt concentration. There was also no change in the emission maximum of the protein.

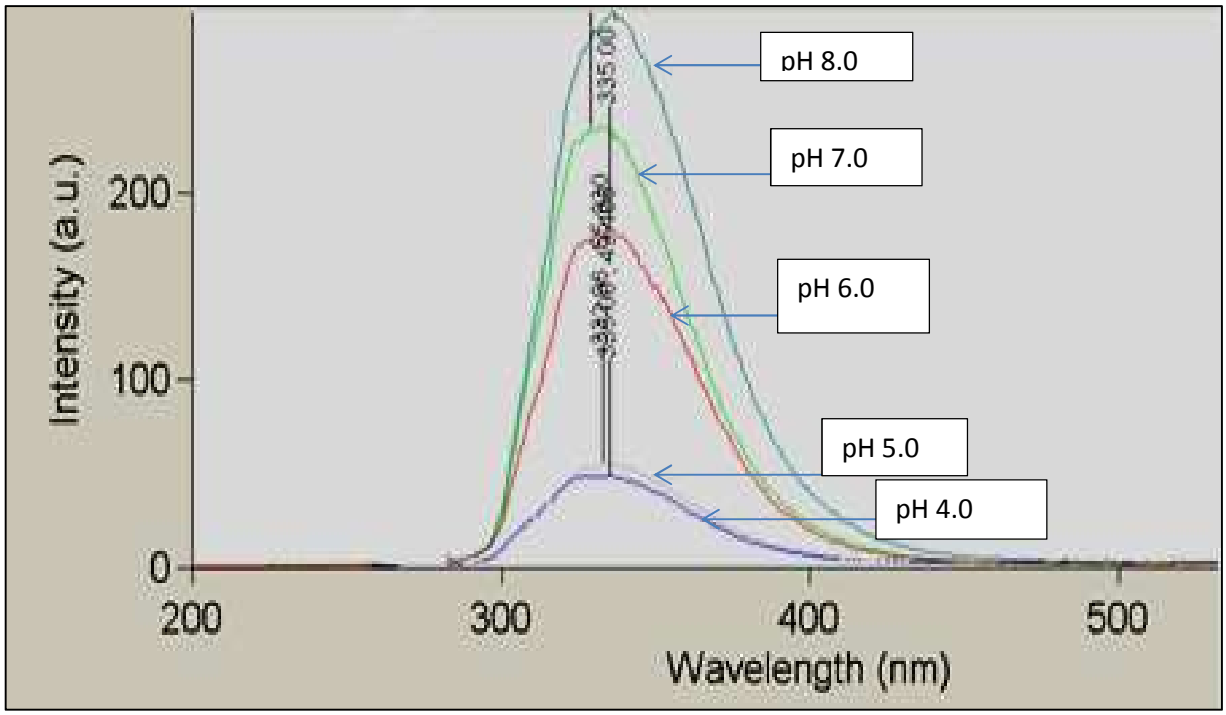


Fig. 4.21: Fluorescence spectra showing the effect of pH on fluorescence intensity at excitation 295 nm and emission at 343nm in CtRGP protein at different pH buffer from 4.0 to 8.0. The protein concentration was 0.35 mg/ml, and solution pH was varied by addition of 0.1 M NaOH or 0.1 M HCl

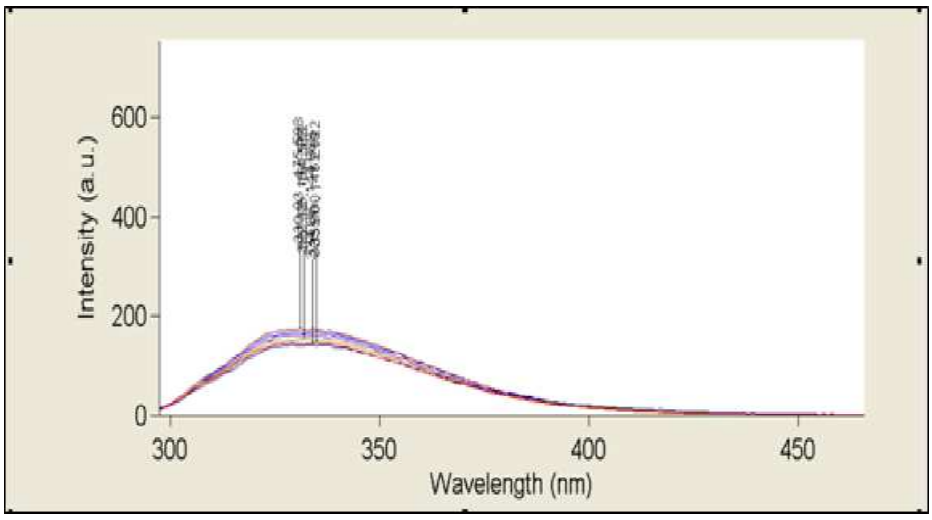


Fig. 4.22: Effect of ionic strength on the fluorescence intensity on the fluorescence spectra of CtRGP protein in water on excitation at 295nm and emission at 343nm. The protein concentration was 0.35mg/ml., and solution ionic strength was adjusted by addition of Sodium chloride.

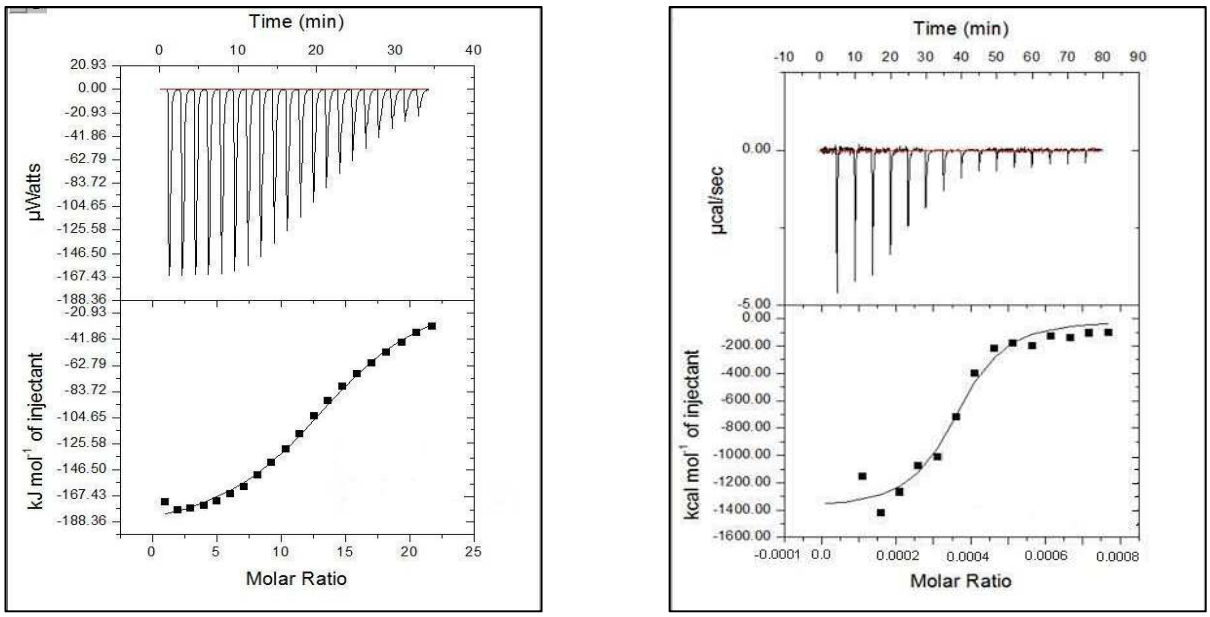
4.7 Determination of CtRGP protein-ligands interaction energetics by ITC

The results of representative titrations of purified CtRGP protein with its substrate UDP-glucose and two inhibitors DIDS and DTNB at 293K have been presented in Fig. 4.23. The upper panels in the figure show that each titration exhibited a monotonic decrease in exothermic heat of binding with each successive injection until saturation. The lower panels indicate a typical sigmoid binding isothermal curve in each titration. A nonlinear least square fit of the ITC data to the identical site model (lower panels in Fig. 4.23 a, b & c) show that the purified CtRGP protein has only one type of site for its substrate and inhibitors. The figure indicates that the binding of the CtRGP protein with UDP-glucose, DIDS or DTNB is an exothermic reaction and the binding isotherms can be fitted to the same model.

Thermodynamic values for the binding of ligands to the purified CtRGP protein have been presented in the Table 4.3. The value of ΔG for binding of UDP- glucose to CtRGP is -9.52 kcal/mol. This binding is enthalpically favoured ($\Delta H = -19.9\, kcal/mol$) but is entropically unfavourable ($T\Delta S = -10.8\, kcal/mol$). ΔG for binding of inhibitor DIDS is -3.10 kcal/mol and is also enthalpically favoured ($\Delta H = -16.5\, kcal/mol$) but entropically unfavourable ($T\Delta S$ =-13.5 kcal/mol). The ΔG for binding of inhibitor DTNB is both enthalpically (ΔH= -10.2 kcal/mol) and entropically ($T\Delta S$= -8.5 kcal/mol) favoured. The value of dissociation constant (K_d) for the inhibitor DTNB is about 7 and 22 folds more than the K_d values for the substrate UDP-glucose and inhibitor DIDS, respectively.

The above results of protein- ligands energetics indicate that the CtRGP protein has a strong, moderate and weak binding affinities with inhibitor DIDS, substrate UDP-glucose and inhibitor DTNB, respectively.

(a) (b)



(c)

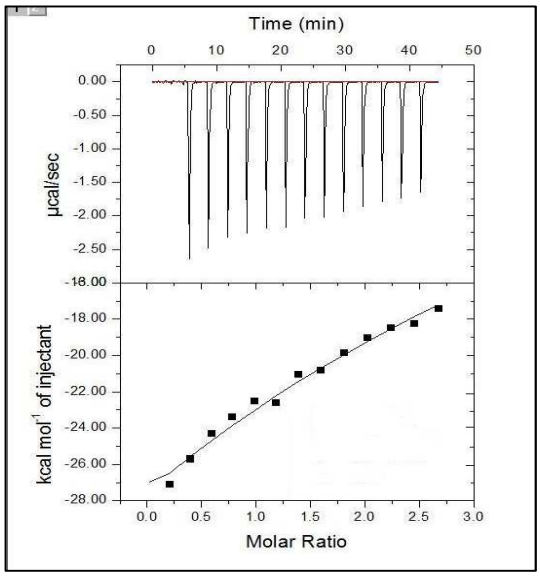


Fig. 4.23: Isothermal calorimetric titration of CtRGP with nucleotide sugar substate UDP-glucose (a) and two inhibitors DIDS (b) and DTNB (c). Upper panels correspond to titration kinetics. Lower panels show the integrated binding isotherms. Molar ratio refers to protein monomer.

Table 4.3: Thermodynamic values for the binding of ligands to purified CtRGP protein.

S. No	Ligand	K_d(μM)	ΔG (kcal/mol)	ΔH(kcal/mol)	TΔS (kcal/mol)	n
1	UDP-glucose	5.06±0.2	-9.52±0.2	-19.9±0.2	-10.8	0.93±0.01
2	DIDS	1.66±0.4	-3.10±0.4	-16.5±0.2	-13.5	0.95±0.02
3	DTNB	36.4±0.4	-2.1±0.2	-10.2±0.6	-8.5	0.85±0.04

4.8. Crystallization trials of guar RGP protein

The crystal hit was observed in Crystal screen cryo in which buffer condition was 0.075M HEPES sodium pH 7.5, precipitant was 1.125 M Lithium sulfate monohydrate and 25% glycerol. The crystal quality was optimized by manual screening around the initial crystal hit (Appendix (i) to (ii). The single crystal of the protein was used for the X-ray diffraction experiment. Crystal was mounted in cryo loop and flash-cooled by direct immersion in liquid nitrogen prior to X-ray diffraction analysis. Diffracting pattern of crystal could not be found due to very small size of crystal.

(a) (b)

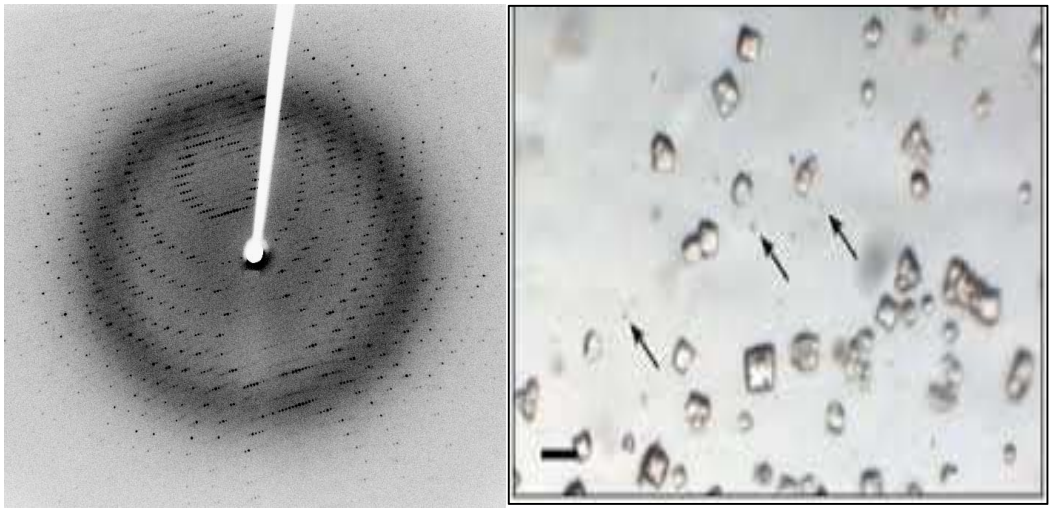


Fig. 4.24. Showing (a) Diffraction pattern of CtRGP protein and (b) Crystals of CtRGP protein. Crystallization condition: Buffer 0.075M HEPES sodium at pH 7.5, precipitant 1.125 M Lithium sulfate monohydrate and 25% glycerol at 4℃.

4.9. Sequence and phylogenetic analysis for CtRGP protein

Nucleotide sequence of CtRGP encodes for 353 amino acids (Fig. 3.2). Physiochemical properties of CtRGP protein were analysed by ProtParam which revealed the predicted isoelectric point (pI), molecular mas, instability index and aliphatic index were found to be 5.75, 41kDa, 29.33 and 83.99 respectively. The predicted value of Grand average of hydropathicity (GRAVY) index of CtRGP is -0.169. Hydropathy profile analysis showed that CtRGP protein is a hydrophilic protein because the peaks obtained in Kyte-Doolittle plot is less than 1.6. Hydropathy profile also revealed that CtRGP lacks a trans-membrane region, suggesting that CtRGP is a soluble protein which functions in the cytoplasm (Fig. 4.25). Nucleotide sequence of CtRGP protein showed maximum sequence similarity with predicted *Glycine max* UDP-arabinopyranose mutase- 1like protein (gi|571549519) is 91%.

Secondary structure prediction by PSIPRED showed (Fig. 4.26) helical content is high in the CtRGP protein, followed by extended strands. Secondary structure prediction by SOPMA and GOR IV also showed that helical content is high in CtRGP protein. Interproscan result indicates that this protein belongs to the member of RGP superfamily having DXD motifs (Fig. 4.27). Visual and computer-aided iPSORT analysis of protein sequence predicted that the protein lacks a signal sequence that might direct CtRGP to a membrane compartment within the cell. Phylogenetic tree analysis indicates that all the members presented in this phylogenetic tree having same molecular (an intramolecular transferase or glycosyltransferase activity) as well as biological functions (cell wall synthesis) (Fig. 4.28). Thus it can be implicated that CtRGP protein may also be involved in intramolecular transferase activity or glycosyltransferase activity and help in cell wall synthesis.

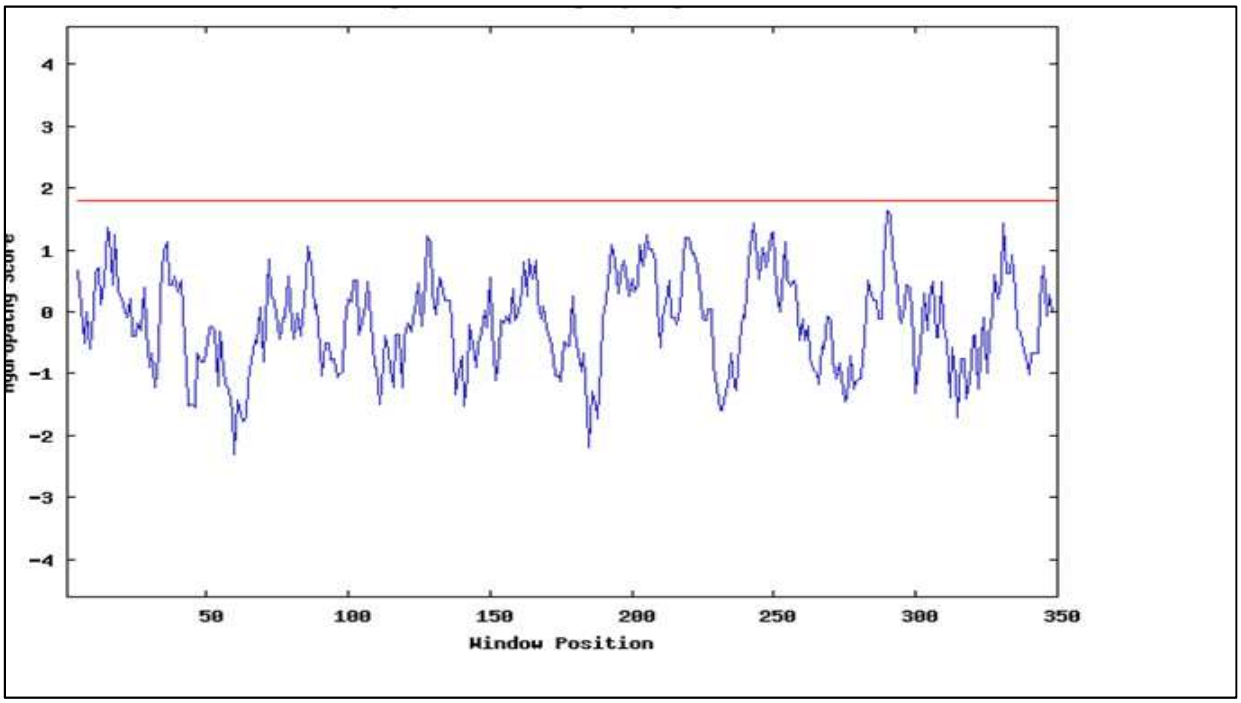


Fig. 4.25: Kyte & Doolittle hydropathy plot for CtRGP protein is showing CtRGP protein is a hydrophilic protein because the peaks obtained is less than 1.6.

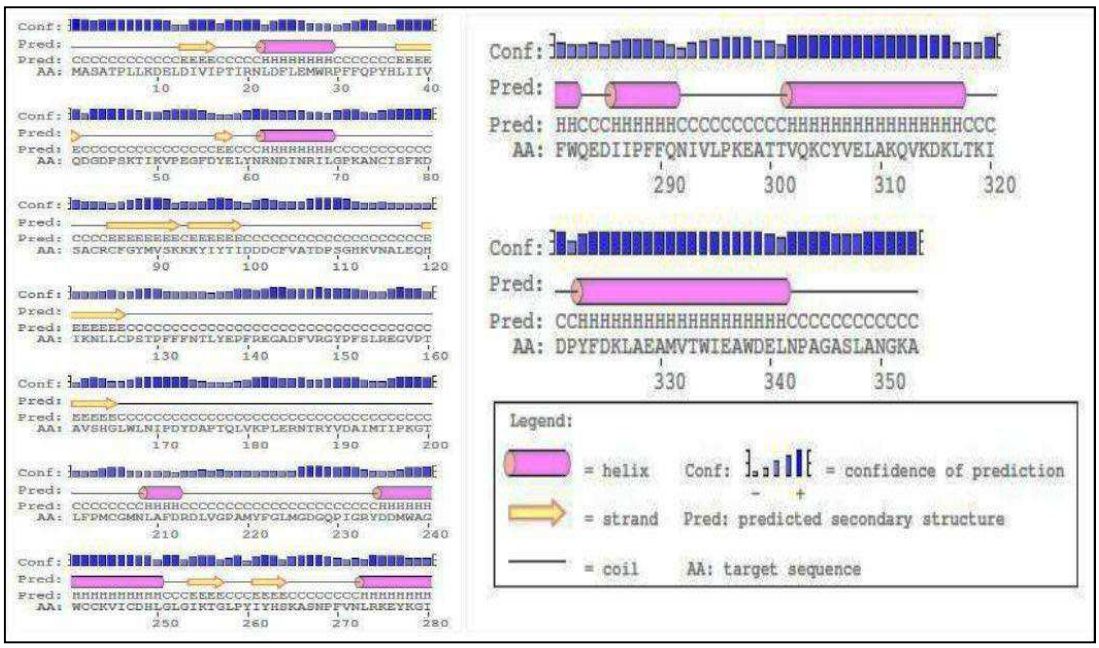


Fig. 4.26: PSIPRED result shows 87 helices (24.47%), 45 strands (12.47%) and 221 coils (62.60%) present at various positions in the CtRGP . The result revealed that random coils dominated among secondary structure element followed by helices and strands.

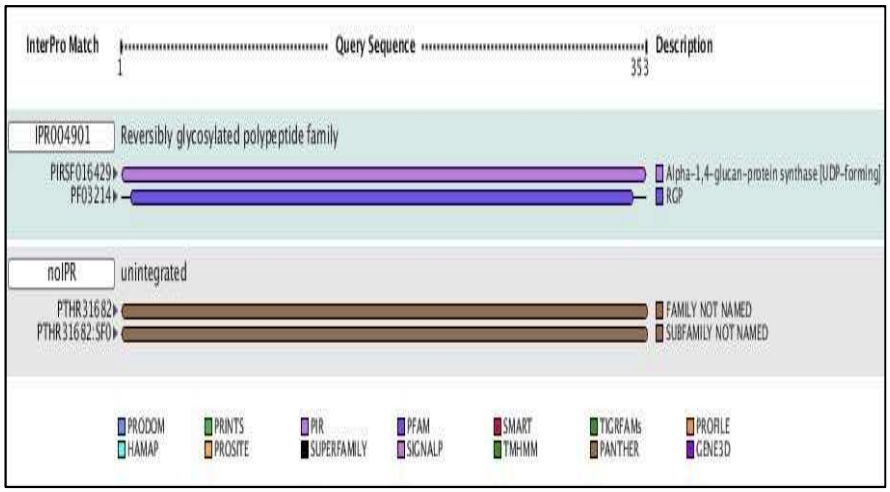


Fig. 4.27: InterProScan result showing reversibly glycosylated polypeptide from guar is member of RGP superfamily.

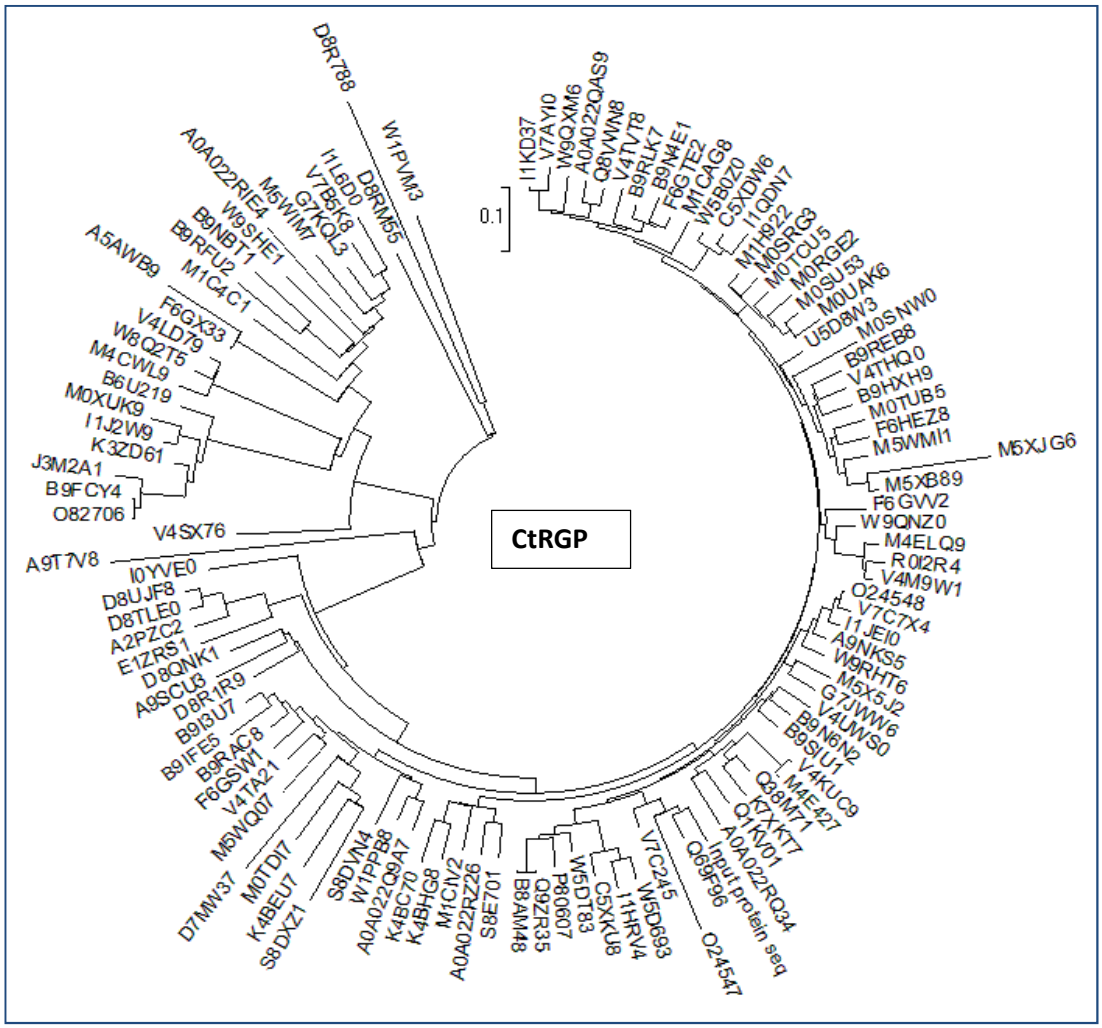


Figure 4.28: The phylogenetic tree constructed for CtRGP protein with molecular evolutionary genetics analysis tool MEGA 6.0 (Tamura *et.al*. 2013). All members present in this phylogenetic tree having the same biological function: cell wall synthesis and molecular function: involved in glycosyltransferase activity or intramolecular transferases.

4.10. Molecular Modeling of CtRGP Protein

The protein blast of CtRGP results showed in Fig. 4.29. Homology modeling of CtRGP was not possible because the matching templates have very high E-value and less query coverage as shown in the Table 4.4. So, the 3D structure was generated by Scratch Protein Predictor tool (3Dpro) and visualized. . The generated 3D model was then refined by i3Drefine. The five models generated and subjected to SAVES server for structure analysis. Refined model structure have been shown in Fig. 4.30 (a) and evaluated for their stereo-chemical quality by PROCHECK and result shows 86.4% residues of predicted structure of CtRGP protein in allowed region (Fig. 4.31(a). The validation of modeled structure of CtRGP protein was done by using server ProSA web. This analysis showed that the protein folding energy is negative and z-score value (-9.25) is in the range of native conformation Fig. 4.31(b). Hence the overall model quality of CtRGP protein is good. The three dimentional structure of CtRGP showed total 10 cys. residues are present out of ten five are involved in disulfide linkage (Fig. 4.30 b). The conformational stability determined by PDBsum indicated that CtRGP has 14 α - helices (Fig. 4.32 and Table 4.5).

Three structure prediction revealed that stability of CtRGP is may be due to the disulfide linkage and conformational stability due to α - helices.

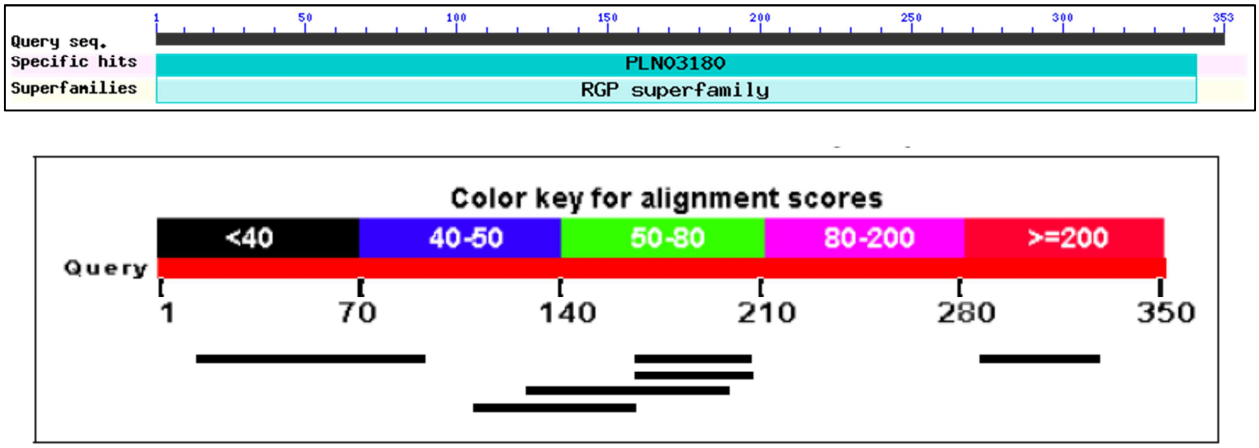


Fig 4.29 Protein Blast result showing the distribution of six blast hits on the Query Sequence CtRGP.

Table 4.4 Protein blast of CtRGP protein sequences

S.No	Discription	Total Score	Query Level	E-Value	Identity	Accession
1	Chain B, Cryoem Structure Of Cytoplasmic Polyhedrosis Virus [Bombyx mori cypovirus 1	31.2	11%	1.5	40%	3IZ3_B
2	Chain A, 3.88 Angstrom Structure Of Cytoplasmic Polyhedrosis Virus By Cryo-Electron Microscopy [Bombyx mori cypovirus 1]	31.2	11%	1.9	40%	3CNF_A
3	Chain A, Crystal Structure Of A Putative Dihydrodipicolinate Synthetase From Pseudomonas Aeruginosa [Pseudomonas aeruginosa]	30.0	20%	2.8	30%	3NA8_A
4	Chain A, Molecular Basis Of Vancomycin Resistance Transfer In Staphylococcus Aureus [Staphylococcus aureus]	28.9	15%	5.9	30%	4HT4_A
5	Chain A, Molecular Basis Of Vancomycin Resistance Transfer In Staphylococcus Aureus [Staphylococcus aureus]	26.9	11%	8.2	33%	2DN9_A
6	Chain A, Human Fibrillarin [Homo sapiens]	28.1	22%	9.6	25%	2IPX_A

(a) (b)

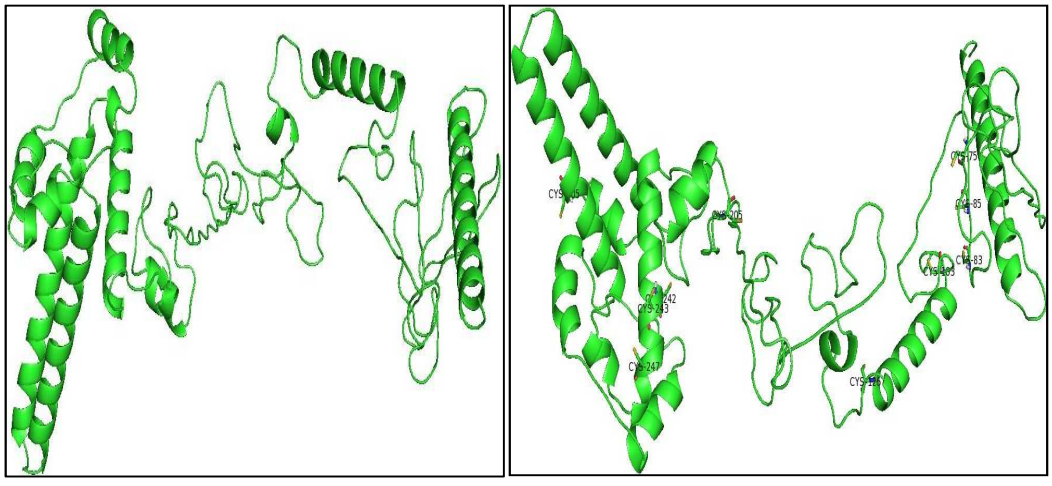


Fig. 4.30: (a) Predicted 3D-modeled structure of CtRGP by Scratch Protein Predictor tool. (b) Ten cysteine residues are located in the positions 75, 83, 85, 103, 126, 205, 242, 243, 247 and 305 in which out of ten five are involved in bonding states Cys 75, 83, 103, 242 and 305. Viewed by PyMol

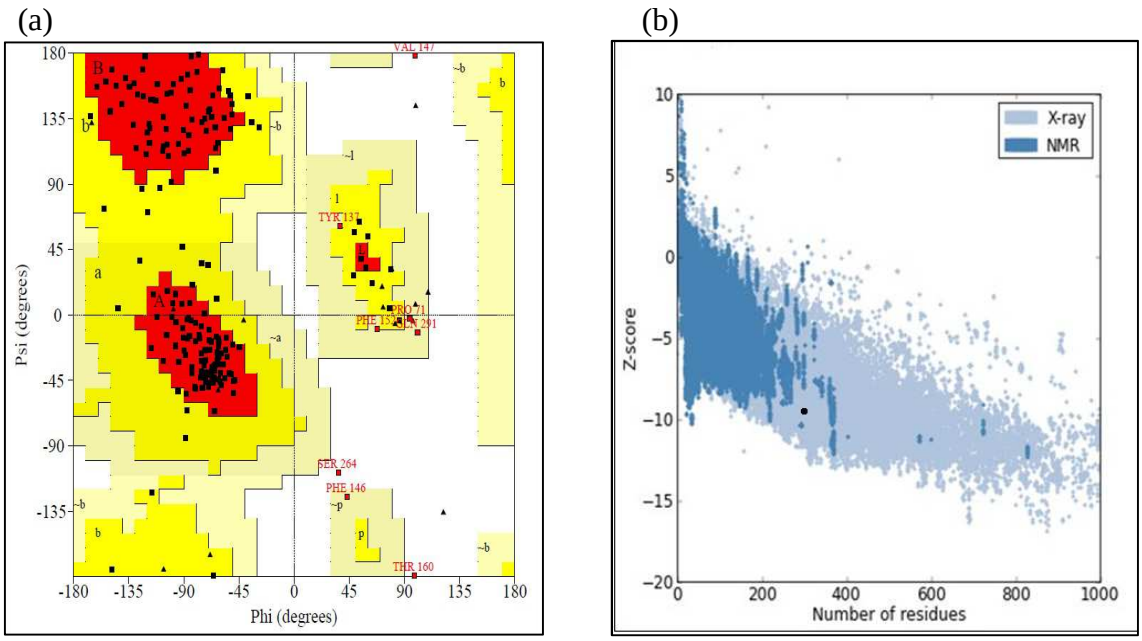


Fig. 4.31: (a) Ramachandran plot of predicted model of CtRGP: This figure is generated by PROCHECK. The red regions in the graph indicate the most allowed or preferred regions covers 86.4%, whereas the yellow regions represent allowed regions11.3%, glycine is represented by triangles 23 and other residues are represented by squares 22 shows predicted model is good quality and (b) Result of predicted model by PROSAWEB.

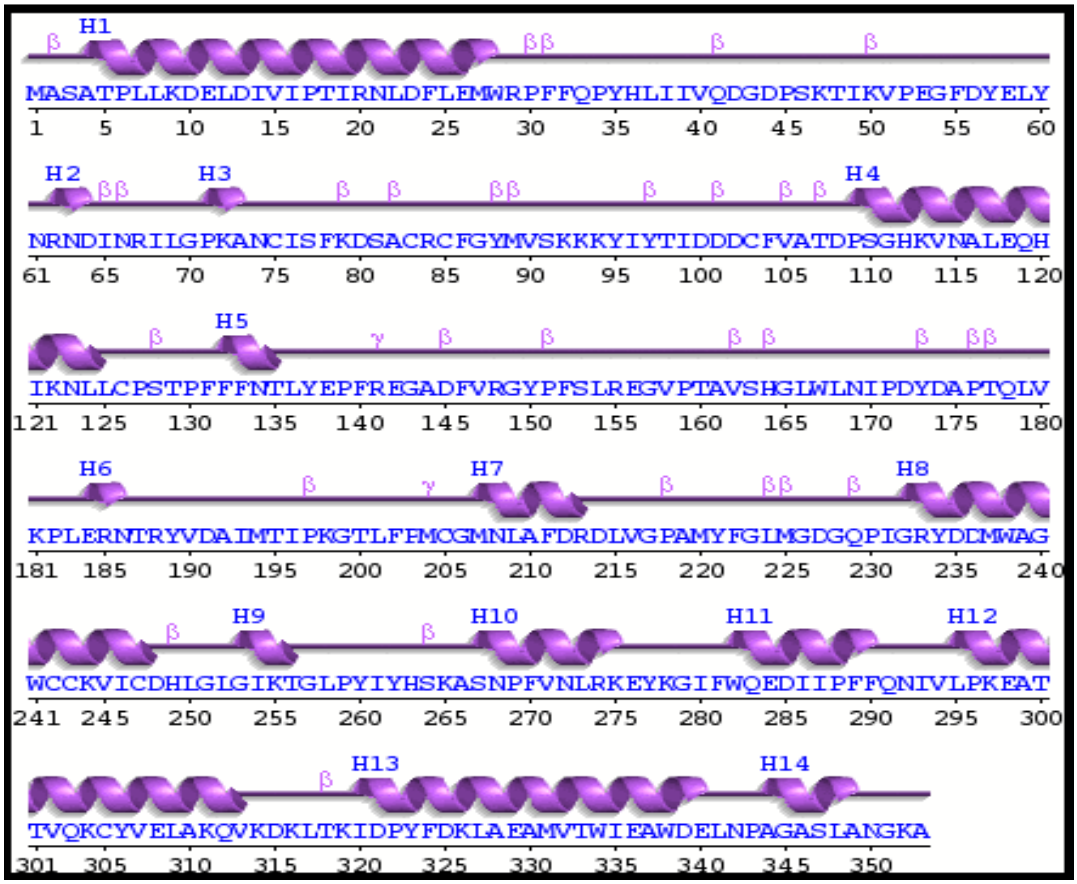


Fig. 4.32: Overview of the secondary structure of CtRGP protein represented as wiring diagram prepared by PDB Sum server [96].

Table 4.5: Details of alpha helices present in secondary structure of CtRGP protein

Helical form	Start	End	No .of residue	Sequences
α1	Ala4	Trp28	25	ATPLLKDELDIVIPTIRNLDFLEMW
α2	Arg 62	Asp 64	03	RND
α3	Pro 71	Ala 73	03	PKA
α4	Pro109	Leu 125	17	PSGHKVNALEQHIKNLL
α5	Phe 132	Thr 135	04	FFNT
α6	Glu 184	Asn 186	03	ERN
α7	Met 207	Arg 213	07	MNIAFDR
α8	Gly 232	Asp 248	17	GRYDDMWAGWCCKVICD
α9	Gly 253	Thr 256	04	GIKT
α10	Ser 267	Lys 275	09	SNPFVNLRK
α11	Trp 282	Phe 290	09	WQEDIIPFF
α12	Leu 295	Val 313	19	LPKEATTVQKCYVELAKQV
α13	Ile320	Glu 340	21	IDPYFDKLAEAMVIWIEAWDE
α14	Ala 344	Ala 349	6	AGASIA

4.11. Molecular docking of substrates at the active sites

The active site of the CtRGP protein is located in the deep grid groove. All the grid information are same for the substrate and two inhibitors (Table 4.6.). The binding sites showed as a surface view in the Fig.4.33. The active site is defined by a strictly conserved catalytic domain. In case of substrate UDP-glucose DXD catalytic domain is present. Where, Arg - 141 position present with H-bond 2.7Å at the catalytic site and apart from that other important residues Ser-163 (3.6 Å), Thr-177 (2.6 Å), Leu-168 (1.8 Å), Asn-169 (2.7 Å) and Lys 198(2.9 Å) are presented with molecular surface view (Fig. 4.35). The docking score for the UDP-glucose is -5.02.

Table 4.6: Grid information generated by autodock 4.2.0 and common for all three ligands.

Grid Map	Atom Type	Minimum Energy (kcal/mol)	Maximum energy (kcal/mol)
1	A	-0.96	3.00e+05
2	C	-1.08	3.00e+05
3	OA	-1.73	2.52e+05
4	N	-1.07	3.00e+05
5	P	-1.18	3.01e+05
6	HD	-0.72	2.00e+05
7	e	-30.78	2.55e+01

(a) (b)

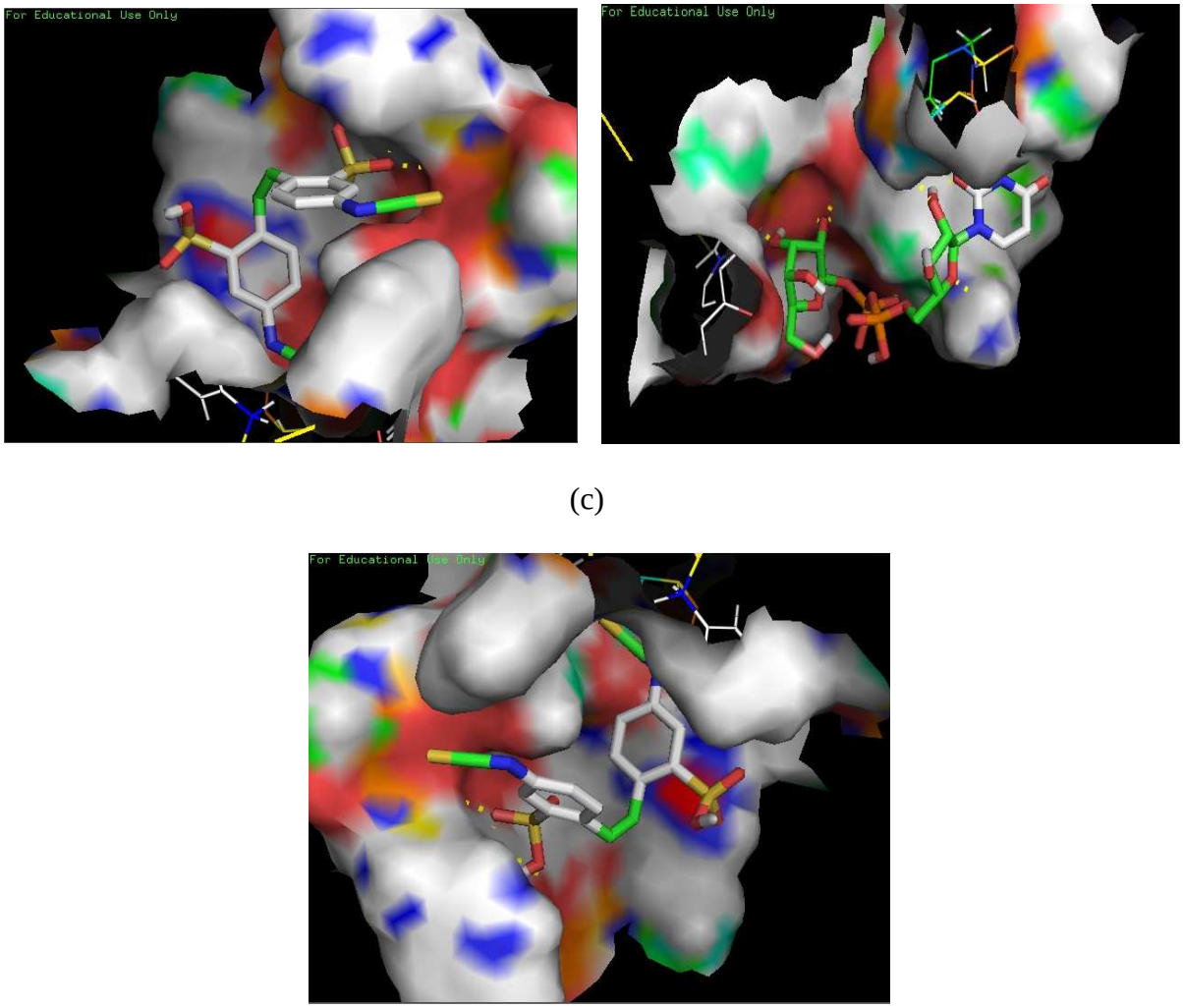


Fig. 4.33: Showing docking result of CtRGP protein interactions with (a) UDP-glucose, (b) inhibitor DIDS and (c) inhibitor DTNB having single binding site.

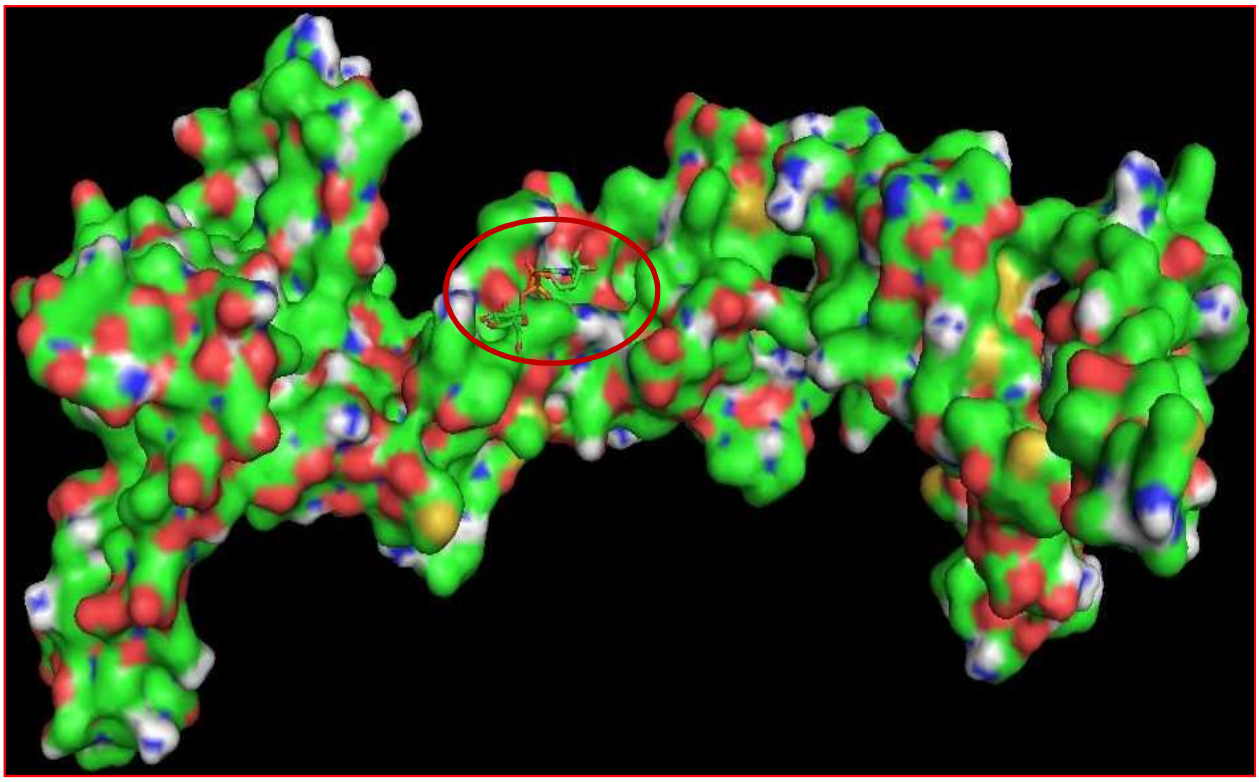


Fig. 4.34: Docking result showing molecular surface for CtRGP (3D-structure from Scratch Protein Predictor) with UDP –glucose from 2Q3E interaction using Autodock 4.2.0 and viewed by PymoL.

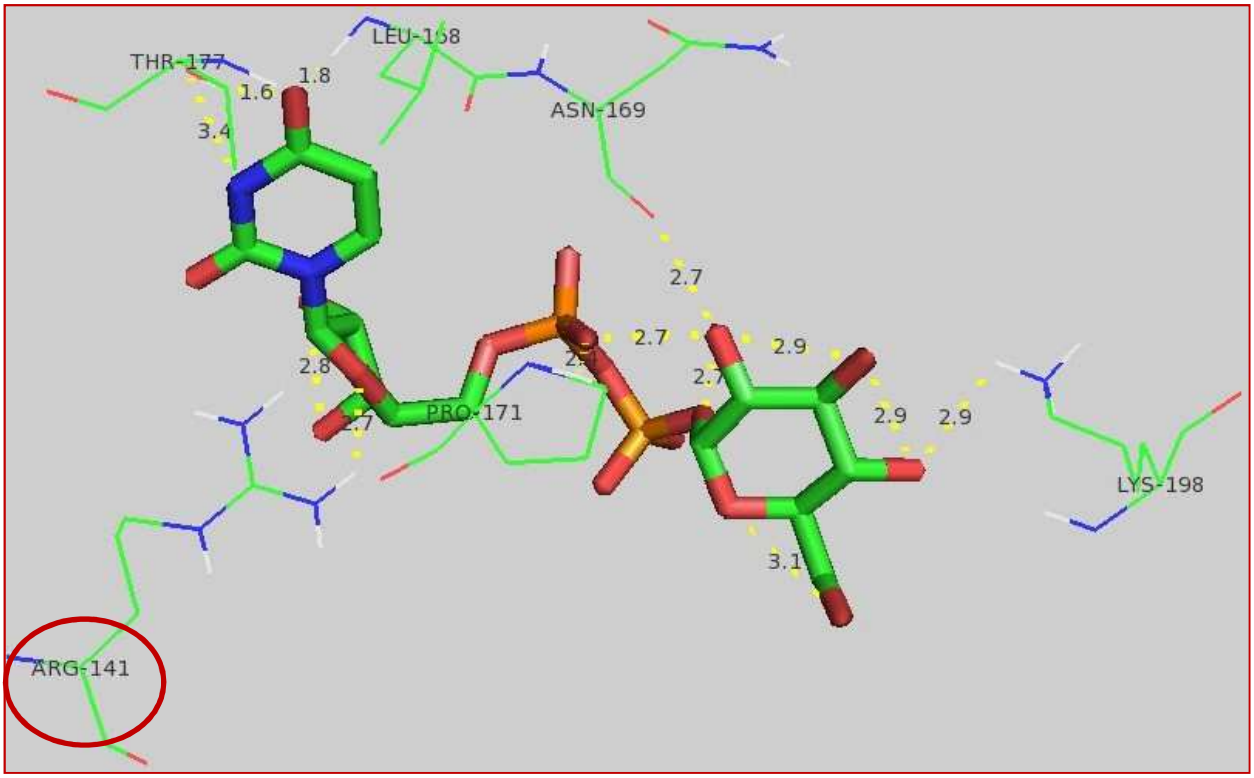


Fig. 4.35: Showing docking of CtRGP protein with UDP-glucose interaction in which Arg-141, Lys-198, Thr-177 Leu-168 and Asn-169 residues are involved by using Autodock 4.2.0 and viewed by Pymol

5. DISCUSSION

The biophysical characterization of guar RGP has not been done so far because of the challenge in isolating the protein in pure, stable and functional form. In this study, cDNa for the guar RGP has been sub- cloned and expressed in *E.coli*. The over-expression of the guar CDNA has been achieved in *E.coli* BL21 (DE3) expression host. Efficient purification was achieved using 6X His-tag. A single band of pure protein corresponding to 42kDa was observed in elution fractions by SDS-PAGE which was confirmed to be CtRGP by western blotting using pea anti-RGP antibody. RGP found to be The size exclusion chromatography showed that the protein exists as a tetramer.

Through MALDI-TOF analysis the molecular mass of CtRGP protein was found to be 42,123.21 Da that corresponded to the result as shown from ProtParam server as well as SDS-PAGE. The molecular mass of RGP has been reported to be quite similar to the molecular mass value of RGP obtained in the other plant species [14, 37, 44, 168]. Hence this protein may be having some role in glycosyltransferase activity.

Conformational stability of CtRGP was examined using CD. Two large peaks of negative ellipticity were obtained which centered at 208 and 222nm, a characteristic of protein with high fraction of α-helices. Thermal stability of CtRGP was also examined by CD, revealing the increase stability even at 68.9℃ corresponding to higher α-helical content. Spectral measurements indicate the protein to be resistant to sufficiently high concentrations of denaturants that is GdHCl (4M) and urea (6M).

Unfolding of proteins causes changes in wavelength and fluorescence intensity due to presence of tryptophan residues in them. Seven tryptophan residues are present in CtRGP protein. Unfolding of CtRGP due to change in pH of the buffer causes decrease only in fluorescence intensity, indicating incomplete protein denaturation within this pH range (pH4 to pH7). Change in ionic strength showed no significant difference in fluorescence intensity.

From the sequence analysis of CtRGP protein, the predicted pI revealed the protein to be acidic. The protein is predicted to be highly thermostable because of high aliphatic index of 29.33. The negative value of GRAVY -0.169 implicates CtRGP to be hydrophilic in nature and the hydropathy plot showed that it lacks transmembrane regions, thereby suggesting that CtRGP is a soluble protein. Ramachandran plot analysis of the energy minimized 3D model showed 86.4% residues were present in most allowed or preferred region and 11.3% in allowed region. ProSA analysis showed that protein folding energy of the modeled structure is in agreement favoring the validation of the model. Docking results indicated that main residues involved in the interaction are Arg, Ser, and Thr. For the glycosyl transferase activity Ser, Thr ,and Arg is required for the interaction. So, this protein may involved in glycosyl transferase activity and may play role in cell wall synthesis.

This is the first report of biophysical characterization of RGP from any plant species. It is expected that this knowledge would be helpful in obtaining the crystal structure of this important protein. This understanding would also be helpful in knowing the exact biophysical function of this novel plant protein.

6. REFERENCES

1. Ahlawat, A., Dhingra, H. R. and Dhankar, J. S. In vitro regeneration of wild species of guar (*Cyamopsis serrata* and *Cyamopsis senegalensis*). *Journal Of Krishi Vigyan* 1(2):48-55 (2013).

2. Ahlawat, A., Pahuja, S. K. and Dhingra, H. R. Overcoming interspecific hybridization barriers in Cyamopsis species. *International Journal of Biotechnology and Bioengineering Research* 4(3):181-190 (2013).

3. Akrem, A., Yousef, N., Begum, A., Negm, A., Meyer, A., Perbandt, M., Buck, F. and Betzel, C. Preliminary crystallographic analysis of a cruciferin protein from seeds of *Moringa oleifera*. *The Protein Journal* 33(3):253-257 (2014).

4. Anderson, E. Endosperm mucilages of legumes. *Industrial & Engineering Chemistry* 41(12):2887-2890 (1949).

5. Arora, P., Dilbaghi, N. and Chaudhury, A. Detection of double stranded RNA in phytopathogenic Macrophomina phaseolina causing charcoal rot in *Cyamopsis tetragonoloba*. *Molecular Biology Reports* 39(3):3047-3054 (2012).

6. Arora, R. N. and Pahuja, S. K. Mutagenesis in guar [*Cyamopsis tetragonoloba* (L.) Taub.]. *Plant Mutation Reports* 2(1):7-9 (2008).

7. As Takeda, A., Z Freitas, F., J Magro, A., E Bernardes, N., Ah Fernandes, C., D Goncalves, R., Celia Bertolini, M. and Rm Fontes, M. Biophysical characterization of the recombinant Importin- α from *Neurospora crassa*. *Protein and Peptide Letters* 20(1):8-16 (2013).

8. Atsushi, I. Thermostability and aliphatic index of globular proteins. *Journal of Biochemistry* 88(6):1895-1898 (1980).

9. Bauer, H. G., Asp, N.-G., Dahlqvist, A., Fredlund, P. E., Nyman, M. and A. Dahlqvist, R. Effect of two kinds of pectin and guar gum on 1, 2-dimethylhydrazine initiation of colon tumors and on fecal β-glucuronidase activity in the rat. *Cancer Research* 41(6):2518-2523 (1981).

10. Beckwith, R. Depending on guar for shale oil and gas development. *Journal of Petroleum Technology* 44:50 (2012).

11. Bhattacharya, D. and Cheng, J. 3Drefine: Consistent protein structure refinement by optimizing hydrogen bonding network and atomic level energy minimization. *Proteins: Structure, Function, and Bioinformatics* 81(1):119-131 (2012).

12. Biswas, S. and Kayastha, A. M. Thermal stability of *Phaseolus vulgaris* leucoagglutinin: a differential scanning calorimetry study. *Journal of Biochemistry and Molecular Biology* 35(5):472-475 (2002).

13. Biswas, S. and Kayastha, A. M. Unfolding and refolding of Leucoagglutinin (PHA-L), an oligomeric lectin from kidney beans (*Phaseolus vulgaris*). *Biochimica et Biophysica Acta (BBA)-General Subjects* 1674(1):40-49 (2004).

14. Bocca, S. N., Kissen, R., Rojas-Beltran, J. A., Noel, F., Gebhardt, C., Moreno, S., du Jardin, P. and Tandecarz, J. S. Molecular cloning and characterization of the enzyme UDP-glucose: protein transglucosylase from potato. *Plant Physiology and Biochemistry* 37(11):809-819 (1999).

15. Bonander, N., Jamshad, M., Oberthur, D., Clare, M., Barwell, J., Hu, K., Farquhar, M. J., Stamataki, Z., Harris, H. J., Dierks, K., Betzel, C., McKeating, J. A. and Bill, R. M. Production, purification and characterization of recombinant, full-length human Claudin-1. *PLoS One* 8(5):e64517 (2013).

16. Bowie, J. U., Luthy, R. and Eisenberg, D. A method to identify protein sequences that fold into a known three-dimensional structure. *Science* 253(5016):164-170 (1991).

17. Bradford, M. M. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Analytical Biochemistry* 72(1):248-254 (1976).

18. Breton, C., Snajdrova, L., Jeanneau, C., Koca, J. and Imberty, A. Structures and mechanisms of glycosyltransferases. *Glycobiology* 16(2):29R-37R (2006).

19. Buckeridge, M. S., Pessoa dos Santos, H. and Tiné, M. A. S. Mobilisation of storage cell wall polysaccharides in seeds. *Plant Physiology and Biochemistry* 38(1-2):141-156 (2000).

20. Bulpin, P. V., Gidley, M. J., Jeffcoat, R. and Underwood, D. R. Development of a biotechnological process for the modification of galactomannan polymers with plant α-galactosidase. *Carbohydrate Polymers* 12:155-168 (1990).

21. Butt, M. S., Shahzadi, N., Sharif, M. K. and Nasir, M. Guar gum: a miracle therapy for hypercholesterolemia, hyperglycemia and obesity. *Critical Reviews in Food Science and Nutrition* 47(4):389-396 (2007).

22. Carpita, N., Tierney, M. and Campbell, M. Molecular biology of the plant cell wall: searching for the genes that define structure, architecture and dynamics. *Plant Molecular Biology* 47(1-2): 1-5(2001).

23. Carpita, N. C. Progress in the biological synthesis of the plant cell wall: new ideas for improving biomass for bioenergy. *Current Opinion in Biotechnology* 23(3):330-337 (2011).

24. Carpita, N. C. and Delmer, D. P. Protection of cellulose synthesis in detached cotton fibers by polyethylene glycol. *Plant Physiology* 66(5):911-916 (1980).

25. Chen, X. Y. and Kim, J. Y. Callose synthesis in higher plants. *Plant Signaling & Behavior* 4(6):489-492 (2009).

26. Cheng, J., Randall, A. Z., Sweredoski, M. J. and Baldi, P. SCRATCH: a protein structure and structural feature prediction server. *Nucleic Acids Research* 33(2):W72-W76 (2005).

27. Colovos, C. and Yeates, T. O. Verification of protein structures: patterns of nonbonded atomic interactions. *Protein Science* 2(9):1511-1519 (1993).

28. Cosgrove, D. J., Li, L. C., Cho, H.-T., Hoffmann-Benning, S., Moore, R. C. and Blecker, D. The growing world of expansins. *Plant and Cell Physiology* 43(12):1436-1444 (2002).

29. Coulman, B., Dalai, A., Heaton, E., Lee, C. P., Lefsrud, M., Levin, D., Lemaux, P. G., Neale, D., Shoemaker, S. P. and Singh, J. Developments in crops and management systems to improve lignocellulosic feedstock production. *Biofuels, Bioproducts and Biorefining* 7(5):582-601 (2013).

30. Creagh, A. L., Ong, E., Jervis, E., Kilburn, D. G. and Haynes, C. A. Binding of the cellulose-binding domain of exoglucanase Cex from Cellulomonas fimi to insoluble microcrystalline cellulose is entropically driven. *Proceedings of the National Academy of Sciences* 93(22):12229-12234 (1996).

31. Cunha, P. L. R., Paula, R. and Feitosa, J. Purification of guar gum for biological applications. *International Journal of Biological Macromolecules* 41(3):324-331 (2007).

32. Cutler, S. Cellulose synthesis: cloning in silico. *Current Biology* 7(2):R108-R111 (1997).

33. Dam, T. K., Torres, M., Brewer, C. F. and Casadevall, A. Isothermal titration calorimetry reveals differential binding thermodynamics of variable region-identical antibodies differing in constant region for a univalent ligand. *Journal of Biological Chemistry* 283(46):31366-31370 (2008).

34. De Pino, V. n., Boran, M., Norambuena, L., Gonzalez, M., Reyes, F., Orellana, A. and Moreno, S. Complex formation regulates the glycosylation of the reversibly glycosylated polypeptide. *Planta* 226(2):335-345 (2007).

35. De Pino, V. n., Busjle, C. M. and Moreno, S. Oligomerization of the reversibly glycosylated polypeptide: its role during rice plant development and in the regulation of self-glycosylation. *Protoplasma* 250(1):111-119 (2012).

36. DeLano, W. L. The PyMOL molecular graphics system. (2002).

37. Delgado, I. J., Wang, Z., de Rocher, A., Keegstra, K. and Raikhel, N. V. Cloning and characterization of AtRGP1 A reversibly autoglycosylated Arabidopsis protein implicated in cell wall biosynthesis. *Plant Physiology* 116(4):1339-1350 (1998).

38. Delmer, D. P. Cellulose biosynthesis. *Annual Review of Plant Physiology* 38(1):259-290 (1987).

39. Dhugga, K. S. Building the wall: genes and enzyme complexes for polysaccharide synthases. *Current Opinion in Plant Biology* 4(6):488-493 (2001).

40. Dhugga, K. S. Biosynthesis of non-cellulosic polysaccharides of plant cell walls. *Phytochemistry* 74:8-19 (2012).

41. Dhugga, K. S., Barreiro, R., Whitten, B., Stecca, K., Hazebroek, J., Randhawa, G. S., Dolan, M., Kinney, A. J., Tomes, D., Nichols, S. and Anderson, P. Guar seed -mannan synthase is a member of the cellulose synthase super gene family. *Science* 303(5656):363-366 (2004).

42. Dhugga, K. S., Barreiro, R., Whitten, B., Stecca, K., Hazebroek, J., Randhawa, G. S., Dolan, M., Kinney, A. J., Tomes, D., Nichols, S. and Anderson, P. Guar seed beta-mannan synthase is a member of the cellulose synthase super gene family. *Science* 303(5656):363-366 (2004).

43. Dhugga, K. S., Tiwari, S. C. and Ray, P. M. A reversibly glycosylated polypeptide (RGP1) possibly involved in plant cell wall synthesis: purification, gene cloning, and trans-Golgi localization. *Proceedings of the National Academy of Sciences* 94(14):7679-7684 (1997).

44. Dhugga, K. S., Ulvskov, P., Gallagher, S. R. and Ray, P. M. Plant polypeptides reversibly glycosylated by UDP-glucose. Possible components of Golgi beta-glucan synthase in pea cells. *Journal of Biological Chemistry* 266(32):21977-21984 (1991).

45. Doyle, M. L. Characterization of binding interactions by isothermal titration calorimetry. *Current Opinion in Biotechnology* 8(1):31-35 (1997).

46. Drakakaki, G., Zabotina, O., Delgado, I., Robert, S. p., Keegstra, K. and Raikhel, N. Arabidopsis reversibly glycosylated polypeptides 1 and 2 are essential for pollen development. *Plant Physiology* 142(4):1480-1492 (2006).

47. Driouich, A., Follet-Gueye, M. L., Bernard, S., Kousar, S., Chevalier, L., VicrÃe-Gibouin, M. t. and Lerouxel, O. Golgi-mediated synthesis and secretion of matrix polysaccharides of the primary cell wall of higher plants. *Frontiers in Plant Science* 3:79 (2012).

48. Dwevedi, A., Dubey, V. K., Jagannadham, M. V. and Kayastha, A. M. Insights into pH-induced conformational transition of beta-galactosidase from *Pisum sativum*

leading to its multimerization. *Applied Biochemistry and Biotechnology* 162(8):2294-2312 (2010).

49. Edwards, M., Bulpin, P. V., Dea, I. C. M. and Reid, J. S. G. Biosynthesis of legume-seed galactomannans in vitro. *Planta* 178(1):41-51 (1989).

50. Edwards, M., Scott, C., Gidley, M. J. and Reid, J. S. G. Control of mannose/galactose ratio during galactomannan formation in developing legume seeds. *Planta* 187(1):67-74 (1992).

51. Edwards, M. E., Choo, T. S., Dickson, C. A., Scott, C., Gidley, M. J. and Reid, J. S. G. The seeds of *Lotus japonicus* lines transformed with sense, antisense, and sense/antisense galactomannan galactosyl transferase constructs have structurally altered galactomannans in their endosperm cell walls. *Plant Physiology* 134(3):1153-1162 (2004).

52. Edwards, M. E., Dickson, C. A., Chengappa, S., Sidebottom, C., Gidley, M. J. and Reid, J. S. Molecular characterisation of a membrane-bound galactosyltransferase of plant cell wall matrix polysaccharide biosynthesis. *The Plant Journal* 19(6):691-697 (1999).

53. Edwards, M. E., Dickson, C. A., Chengappa, S., Sidebottom, C., Gidley, M. J. and Reid, J. S. Molecular characterisation of a membrane bound galactosyltransferase of plant cell wall matrix polysaccharide biosynthesis. *The Plant Journal* 19(6):691-697 (1999).

54. Feingold, D. S. and Avigad, G. Sugar nucleotide transformations in plants. *The Biochemistry of Plants: a Comprehensive Treatise (USA)* (1980).

55. Freyer, M. W. and Lewis, E. A. Isothermal titration calorimetry: experimental design, data analysis, and probing macromolecule/ligand binding and kinetic interactions. *Methods in Cell Biology* 84:79-113 (2008).

56. Fritz, T. A., Hurley, J. H., Trinh, L.-B., Shiloach, J. and Tabak, L. A. The beginnings of mucin biosynthesis: The crystal structure of UDP-GalNAc: polypeptide α-N-acetylgalactosaminyltransferase-T1. *Proceedings of the National Academy of Sciences* 101(43):15307-15312 (2004).

57. Fulton, Z., McAlister, A., Wilce, M. C. J., Brammananth, R., Zaker-Tabrizi, L., Perugini, M. A., Bottomley, S. P., Coppel, R. L., Crellin, P. K. and Rossjohn, J. Crystal structure of a UDP-glucose-specific glycosyltransferase from a Mycobacterium species. *Journal of Biological Chemistry* 283(41):27881-27890 (2008).

58. Garcia-Caballero, A., Gavira, J. A., Pineda-Molina, E., Chayen, N. E., Govada, L., Khurshid, S., Saridakis, E., Boudjemline, A., Swann, M. J. and Shaw Stewart, P. Optimization of protein crystallization: The OptiCryst project. *Crystal Growth & Design* 11(6):2112-2121 (2011).

59. Garcia-Caballero, A., Gavira, J. A., Pineda-Molina, E., Chayen, N. E., Govada, L., Khurshid, S., Saridakis, E., Boudjemline, A., Swann, M. J., Shaw Stewart, P., Briggs, R. A., Kolek, S. A., Oberthuer, D., Dierks, K., Betzel, C., Santana, M., Hobbs, J. R., Thaw, P., Savill, T. J., Mesters, J. R., Hilgenfeld, R., Bonander, N. and Bill, R. M. Optimization of protein crystallization: The OptiCryst project. *Crystal Growth & Design* 11(6):2112-2121 (2011).

60. Garnier, J., Gibrat, J. F., Robson, B. and Doolittle, R. F. GOR secondary structure prediction method version IV. *Methods in Enzymology* 266:540-553 (1996).

61. Gasteiger, E., Gattiker, A., Hoogland, C., Ivanyi, I., Appel, R. D. and Bairoch, A. ExPASy: the proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Research* 31(13):3784-3788 (2003).

62. Geourjon, C. and Deleage, G. SOPMA: significant improvements in protein secondary structure prediction by consensus prediction from multiple alignments. *Computer Applications in the Biosciences* 11(6):681-684 (1995).

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

64. Gibeaut, D. M. and Carpita, N. C. Biosynthesis of plant cell wall polysaccharides. *The FASEB Journal* 8(12):904-915 (1994).

65. Gill, S. C. and Von Hippel, P. H. Calculation of protein extinction coefficients from amino acid sequence data. *Analytical Biochemistry* 182(2):319-326 (1989).

66. Goyal, M. and Sharma, S. K. Traditional wisdom and value addition prospects of arid foods of desert region of North West India. *Indian Journal of Traditional Knowledge* 8(4):581-585 (2009).

67. Greenfield, N. J. Analysis of circular dichroism data. *Methods in Enzymology* 383:282 (2004).

68. Greenfield, N. J. Determination of the folding of proteins as a function of denaturants, osmolytes or ligands using circular dichroism. *Nature Protocols* 1(6):2733-2741 (2007).

69. Greenfield, N. J. Using circular dichroism spectra to estimate protein secondary structure. *Nature Protocols* 1(6):2876-2890 (2007).

70. Greenfield, N. J. and Fasman, G. D. Computed circular dichroism spectra for the evaluation of protein conformation. *Biochemistry* 8(10):4108-4116 (1969).

71. Guerriero, G., Fugelstad, J. and Bulone, V. What do we really know about cellulose biosynthesis in higher plants? *Journal of Integrative Plant Biology* 52(2):161-175 (2010).

72. Guruprasad, K., Reddy, B. V. B. and Pandit, M. W. Correlation between stability of a protein and its dipeptide composition: a novel approach for predicting *in vivo* stability of a protein from its primary sequence. *Protein Engineering* 4(2):155-161 (1990).

73. Handford, M. Biosynthesis of plant cell walls. Cien. Inv. Agr.(In English) 33 (3): 149-166. *Ciencia e InvestigaciÃ³n Agraria* 33(3):149-166 (2006).

74. Hanwell, M. D., Curtis, D. E., Lonie, D. C., Vandermeersch, T., Zurek, E. and Hutchison, G. R. Avogadro: An advanced semantic chemical editor, visualization, and analysis platform. *Journal of Cheminformatics* 4:17 (2012).

75. Hazen, S. P., Scott-Craig, J. S. and Walton, J. D. Cellulose synthase-like genes of rice. *Plant Physiology* 128(2):336-340 (2002).

76. Huang, Y., Yu, H. and Xiao, C. pH-sensitive cationic guar gum/poly (acrylic acid) polyelectrolyte hydrogels: Swelling and in vitro drug release. *Carbohydrate Polymers* 69(4):774-783 (2007).

77. Hymowitz, T. The trans-domestication concept as applied to guar. *Economic Botany* 26(1):49-60 (1972).

78. Hymowitz, T. and Matlock, R. S. Guar in the United States. *Oklahoma Agricultural Experiment Station Technical Bulletin* 611:1-34 (1963).

79. Joersbo, M., Marcussen, J. and Brunstedt, J. *In vivo* modification of the cell wall polysaccharide galactomannan of guar transformed with α -galactosidase gene cloned from senna. *Molecular Breeding* 7(3):211-219 (2001).

80. Joersbo, M., Pedersen, S. G., Nielsen, J. E., Marcussen, J. and Brunstedt, J. Isolation and expression of two cDNA clones encoding UDP-galactose epimerase expressed in developing seeds of the endospermous legume guar. *Plant Science* 149:147–154 (1999).

81. Kang, T. S., Georgieva, D., Genov, N., Murakami, M. T., Sinha, M., Kumar, R. P., Kaur, P., Kumar, S., Dey, S. and Sharma, S. Enzymatic toxins from snake venom: structural characterization and mechanism of catalysis. *FEBS Journal* 278(23):4544-4576 (2011).

82. Kaur, R., Singh, K. and Singh, J. A root-specific wall-associated kinase gene, HvWAK1, regulates root growth and is highly divergent in barley and other cereals. *Functional & Integrative Genomics* 13(2):167-177 (2013).

83. Konishi, T., Miyazaki, Y., Yamakawa, S., Iwai, H., Satoh, S. and Ishii, T. Purification and biochemical characterization of recombinant rice UDP-arabinopyranose mutase generated in insect cells. *Bioscience, Biotechnology, and Biochemistry* 74(1):191-194 (2010).

84. Konishi, T., Ohnishi-Kameyama, M., Funane, K., Miyazaki, Y., Konishi, T. and Ishii, T. An arginyl residue in rice UDP-arabinopyranose mutase is required for catalytic activity and autoglycosylation. *Carbohydrate Research* 345(6):787-791 (2010).

85. Konishi, T., Takeda, T., Miyazaki, Y., Ohnishi-Kameyama, M., Hayashi, T., O'Neill, M. A. and Ishii, T. A plant mutase that interconverts UDP-arabinofuranose and UDP-arabinopyranose. *Glycobiology* 17(3):345-354 (2007).

86. Kotadiya, R., Patel, V. and Patel, H. Guar gum: A better polysaccharide for colonic drug delivery. *Pharmaceutical Information Latest Reviews* 6(2) (2008).

87. Kotadiya, R., Patel, V. and Patel, H. Guar gum: A better polysaccharide for colonic drug delivery. *Pharmaceutical Reviews* 6 (2010).

88. Krishnaiah, Y. S. R., Satyanarayana, V., Dinesh Kumar, B. and Karthikeyan, R. S. In vitro drug release studies on guar gum-based colon targeted oral drug delivery systems of 5-fluorouracil. *European Journal of Pharmaceutical Sciences* 16(3):185-192 (2002).

89. Krukenberg, K. A., Southworth, D. R., Street, T. O. and Agard, D. A. pH-dependent conformational changes in bacterial Hsp90 reveal a Grp94-like conformation at pH 6 that is highly active in suppression of citrate synthase aggregation. *Journal of Molecular Biology* 390(2):278-291 (2009).

90. Kumar, M., Sharma, S., Srinivasan, A., Singh, T. P. and Kaur, P. Structure-based in-silico rational design of a selective peptide inhibitor for thymidine monophosphate kinase of mycobacterium tuberculosis. *Journal of Molecular Modeling* 17(5):1173-1182 (2011).

91. Kuravadi, N. A., Verma. S., Pareek, S., Gahlot, P., Kumari, S., Tanwar, U. K., Bhatele P., Choudhary, M., Gill, K. S., Pruthi, V., Tripathi, S. K., Dhugga, K. S. and Randhawa, G. S. Guar: An Industrial Crop from Marginal Farms. In *Agricultural Sustainability: Progress and Prospects in Crop Research*, editey by G S Bhullar and N K Bhullar (*Elsevier Science Amestrdam)* 47–60 (2013).

92. Kyte, J. and Doolittle, R. F. A simple method for displaying the hydropathic character of a protein. *Journal of Molecular Biology* 157(1):105-132 (1982).

93. Laemmli, U. K. and Favre, M. Maturation of the head of bacteriophage T4: I. DNA packaging events. *Journal of Molecular Biology* 80(4):575-599 (1973).

94. Lamb-Palmer, N. D., Singh, M., Dalton, J. P. and Singh, J. Prokaryotic expression and purification of soluble maize Ac transposase. *Molecular Biotechnology* 54(2):685-691 (2013).

95. Langeveld, S. M. J., Vennik, M., Kottenhagen, M., Van Wijk, R., Buijk, A., Kijne, J. W. and de Pater, S. Glucosylation activity and complex formation of two classes of reversibly glycosylated polypeptides. *Plant Physiology* 129(1):278-289 (2002).

96. Laskowski, R. A. PDBsum: summaries and analyses of PDB structures. *Nucleic Acids Research* 29(1):221-222 (2001).

97. Laskowski, R. A., MacArthur, M. W., Moss, D. S. and Thornton, J. M. PROCHECK: a program to check the stereochemical quality of protein structures. *Journal of Applied Crystallography* 26(2):283-291 (1993).

98. Leavitt, S. and Freire, E. Direct measurement of protein binding energetics by isothermal titration calorimetry. *Current Opinion in Structural Biology* 11(5):560-566 (2001).

99. Lerouxel, O., Cavalier, D. M., Liepman, A. H. and Keegstra, K. Biosynthesis of plant cell wall polysaccharides-a complex process. *Current Opinion in Plant Biology* 9(6):621-630 (2006).

100. Liepman, A. H., Wilkerson, C. G. and Keegstra, K. Expression of cellulose synthase-like (Csl) genes in insect cells reveals that CslA family members encode mannan synthases. *Proceedings of the National Academy of Sciences* 102(6):2221-2226 (2005).

101. Liithy, R., Bowie, J. U. and Eisenberg, D. Assessment of protein models with three-dimensional profiles. *Nature* 356(6364):83-85 (1992).

102. Mahammad, S., Prud'homme, R. K., Roberts, G. W. and Khan, S. A. Kinetics of enzymatic depolymerization of guar galactomannan. *Biomacromolecules* 7(9):2583-2590 (2006).

103. Mathiyazhagan, S., Pahuja, S. K. and Ahlawat, A. Regeneration in cultivated (*Cyamopsis tetragonoloba* L.) and wild species (*C. Serrata*) of Guar. *Legume Research-An International Journal* 36(2):180-187 (2013).

104. Mestechkina, N. M., Egorov, A. V. and Shcherbukhin, V. D. Synthesis of galactomannan sulfates. *Applied Biochemistry and Microbiology* 42(3):326-330 (2006).

105. Miller, D. W. and Agard, D. A. Enzyme specificity under dynamic control: A normal mode analysis of α-lytic protease. *Journal of Molecular Biology* 286(1):267-278 (1999).

106. Mohamed, N. A., Randles, J. W. and Francki, R. I. B. Protein composition of tomato spotted wilt virus. *Virology* 56(1):12-21 (1973).

107. Mohnen, D. Pectin structure and biosynthesis. *Current Opinion in Plant Biology* 11(3):266-277 (2008).

108. Moreno, S., Cardini, C. E. and Tandecarz, J. S. alpha- Glucan synthesis on a protein primer, uridine diphosphoglucose: protein transglucosylase I. *European Journal of Biochemistry* 157(3):539-545 (1986).

109. Morris, G. M., Huey, R., Lindstrom, W., Sanner, M. F., Belew, R. K., Goodsell, D. S. and Olson, A. J. AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *Journal of Computational Chemistry* 30(16):2785-2791 (2009).

110. Mudgil, D., Barak, S. and Khatkar, B. S. Effect of enzymatic depolymerization on physicochemical and rheological properties of guar gum. *Carbohydrate Polymers* 90(1):224-228 (2012).

111. Mudgil, D., Barak, S. and Khatkar, B. S. Guar gum: processing, properties and food applications-A Review. *Journal of Food Science and Technology*:1-10 (2014).

112. Naoumkina, M. and Dixon, R. A. Characterization of the mannan synthase promoter from guar (*Cyamopsis tetragonoloba*). *Plant Cell Reports* 30(6):997-1006 (2011).

113. Naoumkina, M., Torres-Jerez, I., Allen, S., He, J., Zhao, P. X., Dixon, R. A. and May, G. D. Analysis of cDNA libraries from developing seeds of guar (*Cyamopsis tetragonoloba* (L.) Taub). *BMC plant Biology* 7(1):62 (2007).

114. Naoumkina, M., Torres-Jerez, I., Allen, S., He, J., Zhao, P. X., Dixon, R. A. and May, G. D. Analysis of cDNA libraries from developing seeds of guar (*Cyamopsis tetragonoloba* (L.) Taub). *BMC Plant Biology* 7:62 (2007).

115. Overbeeke, N., Fellinger, A. J., Toonen, M. Y., Wassenaar, D. and Verrips, C. T. Cloning and nucleotide sequence of the α-galactosidase cDNA from *Cyamopsis tetragonoloba* (guar). *Plant Molecular Biology* 13:541-550 (1989).

116. P. Rakesh, J. U. N., Singh .M. and Henry A. Guar and its polysaccharide: a review. *Journal of Arid Legumes* 7(1):1-9 (2010).

117. Pathak, R., Singh, M. and Henry, A. Genetic diversity and interrelationship among clusterbean (*Cyamopsis tetragonoloba*) genotypes for qualitative traits. *Indian Journal of Agricultural Sciences* 81(5):402 (2011).

118. Pathak, R., Singh, S. K., Singh, M. and Henry, A. Molecular assessment of genetic diversity in cluster bean (*Cyamopsis tetragonoloba*) genotypes. *Journal of Genetics* 89(2):243 (2010).

119. Pauly, M., Gille, S., Liu, L., Mansoori, N., de Souza, A., Schultink, A. and Xiong, G. Hemicellulose biosynthesis. *Planta* 238(4):627-642 (2013).

120. Pauly, M. and Keegstra, K. Cell-wall carbohydrates and their modification as a resource for biofuels. *The Plant Journal* 54(4):559-568 (2008).

121. Pear, J. R., Kawagoe, Y., Schreckengost, W. E., Delmer, D. P. and Stalker, D. M. Higher plants contain homologs of the bacterial *cel*A genes encoding the catalytic subunit of cellulose synthase. *Proceedings of the National Academy of Sciences* 93(22):12637-12642 (1996).

122. Perez-Iratxeta, C. and Andrade-Navarro, M. A. K2D2: estimation of protein secondary structure from circular dichroism spectra. *BMC Structural Biology* 8(1):25 (2008).

123. Pitkänen, L., Tuomainen, P., Mikkonen, K. S. and Tenkanen, M. The effect of galactose side units and mannan chain length on the macromolecular characteristics of galactomannans. *Carbohydrate Polymers* 86:1230– 1235 (2011).

124. Punia, A., Arora, P., Yadav, R. and Chaudhury, A. Optimization and inference of PCR conditions for genetic variability studies of commercially important cluster bean varieties by RAPD analysis. *Asian Pacific Journal of Molecular Biology and Biotechnology* 17:33-38 (2009).

125. Punia, A., Yadav, R., Arora, P. and Chaudhury, A. Molecular and morphophysiological characterization of superior cluster bean (*Cymopsis tetragonoloba*) varieties. *Journal of Crop Science and Biotechnology* 12(3):143-148 (2009).

126. Rau, U., Vandamme, E. J., De Baets, S. and Steinbuchel, A. Biopolymers. Vol. 6: Polysaccharides II: Polysaccharides From Eukaryotes. In: Weinheim Wiley-VCH, Germany, (2002).

127. Rautengarten, C., Ebert, B., Herter, T., Petzold, C. J., Ishii, T., Mukhopadhyay, A., Usadel, B. R. and Scheller, H. V. The interconversion of UDP-arabinopyranose and UDP-arabinofuranose is indispensable for plant development in Arabidopsis. *The Plant Cell* 23(4):1373-1390 (2011).

128. Ray, P. M. Principles of plant cell expansion. *In physiology of plant growth during plant growth*, edited by D J Consgrove and D P Kenivel (Ann. Soc. Plant Physiol. Rockville, Md.) 1-17 (1987).

129. Ray, P. M., Shininger, T. L. and Ray, M. M. Isolation of Beta-glucan synthetase particles from plant cells and identification with Golgi membranes. *Proceedings of the National Academy of Sciences* 64(2):605-612 (1969).

130. Reid, J. S., Edwards, M. and Dea, I. Biosynthesis of galactomannan in the endosperms of developing fenugreek (*Trigonella foenum-graecum* L.) and guar (*Cyamopsis tetragonoloba*[L.] Taub.) seeds. *Food Hydrocolloids* 1(5):381-385 (1987).

131. Reid, J. S. G., Edwards, M., Gidley, M. J. and Clark, A. H. Enzyme specificity in galactomannan biosynthesis. *Planta* 195(4):489-495 (1995).

132. Reid, J. S. G. and Edwards, M. E. Galactomannans and other cell wall storage polysaccharides in seeds. *Food Science and Technology NewYork -MARCEL DEKKER*:155-155 (1995).

133. Reid, J. S. G., Edwards, M. E., Dickson, C. A., Scott, C. and Gidley, M. J. Tobacco transgenic lines that express fenugreek galactomannan galactosyltransferase constitutively have structurally altered galactomannans in their seed endosperm cell walls. *Plant Physiology* 131(3):1487-1495 (2003).

134. Reiter, W.-D. Biosynthesis and properties of the plant cell wall. *Current Opinion in Plant Biology* 5(6):536-542 (2002).

135. Reiter, W. D. Biochemical genetics of nucleotide sugar interconversion reactions. *Current Opinion in Plant Biology* 11(3):236-243 (2008).

136. Richmond, T. A. and Somerville, C. R. The cellulose synthase superfamily. *Plant Physiology* 124(2):495-498 (2000).

137. Richter, S., Voss, U. and Jurgens, G. Post-Golgi traffic in plants. *Traffic* 10(7):819-828 (2009).

138. Rothschild, A., Wald, F. A., Bocca, S. N. and Tandecarz, J. S. Inhibition of UDP-glucose: protein transglucosylase by a maize endosperm protein factor. *Cellular and Molecular Biology* 42(5):645-651 (1996).

139. Sagi, G., Katz, A., Guenoune-Gelbart, D. and Epel, B. L. Class 1 reversibly glycosylated polypeptides are plasmodesmal-associated proteins delivered to plasmodesmata via the Golgi apparatus. *The Plant Cell* 17(6):1788-1800 (2005).

140. Sambrook, J. W. and Russel, D. Molecular cloning: a laboratory manual. *Quarterly Review of Biology* 76(3):348-349 (2001).

141. Sandhu, A. P. S., Randhawa, G. S. and Dhugga, K. S. Plant cell wall matrix polysaccharide biosynthesis. *Molecular Plant* 2(5):840-850 (2009).

142. Scheible, W. R. d. and Pauly, M. Glycosyltransferases and cell wall biosynthesis: novel players and insights. *Current Opinion in Plant Biology* 7(3):285-295 (2004).

143. Seifert, G. J. Nucleotide sugar interconversions and cell wall biosynthesis: how to bring the inside to the outside. *Current Opinion in Plant Biology* 7(3):277-284 (2004).

144. Selth, L. A., Dogra, S. C., Rasheed, M. S., Randles, J. W. and Rezaian, M. A. Identification and characterization of a host reversibly glycosylated peptide that interacts with the Tomato leaf curl virus V1 protein. *Plant Molecular Biology* 61(1-2):297-310 (2006).

145. Selth, L. A., Randles, J. W. and Rezaian, M. A. Host responses to transient expression of individual genes encoded by Tomato leaf curl virus. *Molecular Plant-Microbe Interactions* 17(1):27-33 (2004).

146. Sharma, P. and Gummagolmath, K. C. Reforming guar industry in India: issues and strategies. *Agricultural Economics Research Review* 25(1) (2012).

147. Singh, A. K., Singh, N., Tiwari, A., Sinha, M., Kushwaha, G. S., Kaur, P., Srinivasan, A., Sharma, S. and Singh, T. P. First structural evidence for the mode of diffusion of aromatic ligands and ligand-induced closure of the hydrophobic channel in heme peroxidases. *Journal of Biological Inorganic Chemistry* 15(7):1099-1107 (2010).

148. Singh, D. G., Lomako, J., Lomako, W. M., Whelan, W. J., Meyer, H. E., Serwe, M. and Metzger, J. r. W. β-Glucosylarginine: a new glucose-protein bond in a self-glucosylating protein from sweet corn. *FEBS letters* 376(1):61-64 (1995).

149. Somerville, C., Bauer, S., Brininstool, G., Facette, M., Hamann, T., Milne, J., Osborne, E., Paredez, A., Persson, S. and Raab, T. Toward a systems approach to understanding plant cell walls. *Science* 306(5705):2206-2211 (2004).

150. Soumya, R. S., Ghosh, S. and Abraham, E. T. Preparation and characterization of guar gum nanoparticles. *International Journal of Biological Macromolecules* 46(2):267-269 (2010).

151. Srivastava, M. and Kapoor, V. P. Seed galactomannans: an overview. *Chemistry & Biodiversity* 2(3):295-317 (2005).

152. Street, T. O., Lavery, L. A. and Agard, D. A. Substrate binding drives large-scale conformational changes in the Hsp90 molecular chaperone. *Molecular Cell* 42(1):96-105 (2011).

153. Sullad, A. G., Manjeshwar, L. S. and Aminabhavi, T. M. Novel pH-sensitive hydrogels prepared from the blends of poly (vinyl alcohol) with acrylic acid-graft-guar gum matrixes for isoniazid delivery. *Industrial & Engineering Chemistry Research* 49(16):7323-7329 (2010).

154. Tamura, K., Stecher, G., Peterson, D., Filipski, A. and Kumar, S. MEGA6: molecular evolutionary genetics analysis version 6.0. *Molecular Biology and Evolution* 30(12):2725-2729 (2013).

155. Thompson, J. D., Higgins, D. G. and Gibson, T. J. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research* 22(22):4673-4680 (1994).

156. Towbin, H., Staehelin, T. and Gordon, J. Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. *Proceedings of the National Academy of Sciences* 76(9):4350-4354 (1979).

157. Vaten, A., Dettmer, J., Wu, S., Stierhof, Y.-D., Miyashima, S., Yadav, S. R., Roberts, C. J., Campilho, A., Bulone, V. and Lichtenberger, R. Callose biosynthesis regulates symplastic trafficking during root development. *Developmental Cell* 21(6):1144-1155 (2011).

158. Viotti, C., Bubeck, J., Stierhof, Y. D., Krebs, M., Langhans, M., van den Berg, W., van Dongen, W., Richter, S., Geldner, N. and Takano, J. Endocytic and secretory traffic in Arabidopsis merge in the trans-Golgi network/early endosome, an independent and highly dynamic organelle. *The Plant Cell* 22(4):1344-1357 (2010).

159. Warren, C. The Supreme Court in United States History: Cosimo, Inc. (1963).

160. Wiederstein, M. and Sippl, M. J. ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. *Nucleic Acids Research* 35(2):W407-W410 (2007).

161. Wiseman, T., Williston, S., Brandts, J. F. and Lin, L. N. Rapid measurement of binding constants and heats of binding using a new titration calorimeter. *Analytical Biochemistry* 179(1):131-137 (1989).

162. Wu, A. M., Ling, C. and Liu, J. Y. Isolation of a cotton reversibly glycosylated polypeptide (GhRGP1) promoter and its expression activity in transgenic tobacco. *Journal of Plant Physiology* 163(4):426-435 (2006).

163. Yadav, H., Prasad, A. K., Goswami, P., Pednekar, S., Haque, E. and Shah, M. Guar Industry Outlook 2015. (2013).

164. Yoon, S. J., Chu, D. C. and Juneja, L. R. Chemical and physical properties, safety and application of partially hydrolized guar gum as dietary fiber. *Journal of Clinical Biochemistry and Nutrition* 42(1):1 (2008).

165. Yoon, S. J., Chu, D. C. and Juneja, L. R. Chemical and physical properties, safety and application of partially hydrolized guar gum as dietary fiber. *Journal of Clinical Biochemistry and Nutrition* 42(1):1 (2008).

166. Zavaliev, R., Sagi, G., Gera, A. and Epel, B. L. The constitutive expression of Arabidopsis plasmodesmal-associated class 1 reversibly glycosylated polypeptide impairs plant development and virus spread. *Journal of Experimental Botany* 61(1):131-142 (2010).

167. Zhang, G. F. and Staehelin, L. A. Functional compartmentation of the Golgi apparatus of plant cells immunocytochemical analysis of high-pressure frozen-and freeze-substituted sycamore maple suspension culture cells. *Plant Physiology* 99(3):1070-1083 (1992).

168. Zhao, G. R. and Liu, J. Y. Isolation of a cotton RGP gene: a homolog of reversibly glycosylated polypeptide highly expressed during fiber development. *Biochimica et Biophysica Acta (BBA)-Gene Structure and Expression* 1574(3):370-374 (2002).

169. Zhao, G.-R., Liu, J.-Y. and Du, X.-M. Molecular cloning and characterization of cotton cDNAs expressed in developing fiber cells. *Bioscience, Biotechnology, and Biochemistry* 65(12):2789-2793 (2001).

170. Zhong, R. and Ye, Z.-H. Regulation of cell wall biosynthesis. *Current Opinion in Plant Biology* 10(6):564-572 (2007).

Appendix i: Manual crystallization of Cryo crystal screen with different buffer and salt composition.

	1	2	3	4	5	6	7	8	9	10	11	12
A	**0.5M Li_2SO_4 20µl Hepes sodium 0.075M (7.5µl) pH 6.8 Glycerol 25µl MQ 47.5µl**	**1.0 M Li_2SO_4 40µl Hepes sodium 0.075M (7.5µl) pH 6.8 Glycerol 25µl MQ 27.5µl**	**1.125M Li_2SO_4 45µl Hepes sodium 0.075M (7.5µl) pH 6.8 Glycerol 25µl MQ 22.5µl**	**1.5M Li_2SO_4 60µl Hepes sodium 0.075M (7.5µl) pH 6.8 Glycerol 25µl MQ 7.5µl**	**2.0M Li_2SO_4 80µl Hepes sodium 0.075M (7.5µl) pH 6.8 Glycerol 12.5µl MQ 0.0µl**	**1.125M Li_2SO_4 45µl Hepes sodium 0.075M (7.5µl) pH 6.8 Glycerol 12.5µl MQ 35.0µl**	**0.5M Li_2SO_4 20µl Hepes sodium 0.075M (7.5µl) pH 6.8 Glycerol 15.0µl MQ 57.5µl**	**1.0M Li_2SO_4 40µl Hepes sodium 0.075M (7.5µl) pH 6.8 Glycerol 15.0µl MQ 37.5µl**	**1.125M Li_2SO_4 45µl Hepes sodium 0.075M (7.5µl) pH 6.8 Glycerol 15.0µl MQ 32.5µl**	**1.5M Li_2SO_4 60µl Hepes sodium 0.075M (7.5µl) pH 6.8 Glycerol 15µl MQ 17.5µl**	**2.0M Li_2SO_4 79µl Hepes sodium 0.075M (7.5µl) pH 6.8 Glycerol 13.5µl MQ 0.0µl**	**1.125M Li_2SO_4 45µl Hepes sodium 0.075M (7.5µl) pH 6.8 Glycerol 13.0µl MQ 35.5µl**
B	**0.5M Li_2SO_4 20µl Hepes sodium 0.075M (7.5µl) pH 7.0 Glycerol 25µl MQ 47.5µl**	**1.0 M Li_2SO_4 40µl Hepes sodium 0.075M (7.5µl) pH 7.0 Glycerol 25µl MQ 27.5µl**	**1.125M Li_2SO_4 45µl Hepes sodium 0.075M (7.5µl) pH 7.0 Glycerol 25µl MQ 22.5µl**	**1.5M Li_2SO_4 60µl Hepes sodium 0.075M (7.5µl) pH 7.0 Glycerol 25µl MQ 7.5µl**	**2.0MLi_2SO_4 80µl Hepes sodium 0.075M (7.5µl) pH 7.0 Glycerol 12.5µl MQ 0.0µl**	**1.125M Li_2SO_4 45µl Hepes sodium 0.075M (7.5µl) pH 7.0 Glycerol 12.5µl MQ 35.0µl**	**0.5M Li_2SO_4 20µl Hepes sodium 0.075M (7.5µl) pH 7.0 Glycerol 15.0µl MQ 57.5µl**	**0.5M Li_2SO_4 40µl Hepes sodium 0.075M (7.5µl) pH 7.0 Glycerol 15.0µl MQ 37.5µl**	**1.125M Li_2SO_4 45µl Hepes sodium 0.075M (7.5µl) pH 7.0 Glycerol 15.0µl MQ 32.5µl**	**1.5M Li_2SO_4 60µl Hepes sodium 0.075M (7.5µl) pH 7.0 Glycerol 15µl MQ 17.5µl**	**2.0M Li_2SO_4 79µl Hepes sodium 0.075M (7.5µl) pH 7.0 Glycerol 13.5µl MQ 0.0µl**	**1.125M Li_2SO_4 45µl Hepes sodium 0.075M (7.5µl) pH 7.0 Glycerol 13.0µl MQ 35.5µl**
C	**0.5M Li_2SO_4 20µl Hepes sodium 0.075M (7.5µl) pH 7.5 Glycerol 25µl MQ 47.5µl**	**1.0 M Li_2SO_4 40µl Hepes sodium 0.075M (7.5µl) pH 7.5 Glycerol 25µl MQ 27.5µl**	**1.125M Li_2SO_4 45µl Hepes sodium 0.075M (7.5µl) pH 7.5 Glycerol 25µl MQ 22.5µl**	**1.5M Li_2SO_4 60µl Hepes sodium 0.075M (7.5µl) pH 7.5 Glycerol 25µl MQ 7.5µl**	**2.0M Li_2SO_4 80µl Hepes sodium 0.075M (7.5µl) pH 7.5 Glycerol 12.5µl MQ 0.0µl**	**1.125M Li_2SO_4 45µl Hepes sodium 0.075M (7.5µl) pH 7.5 Glycerol 12.5µl MQ 35.0µl**	**0.5M Li_2SO_4 20µl Hepes sodium 0.075M (7.5µl) pH 7.5 Glycerol 15.0µl MQ 57.5µl**	**0.5M Li_2SO_4 40µl Hepes sodium 0.075M (7.5µl) pH 7.5 Glycerol 15.0µl MQ 37.5µl**	**1.125M Li_2SO_4 45µl Hepes sodium 0.075M (7.5µl) pH 7.5 Glycerol 15.0µl MQ 32.5µl**	**1.5M Li_2SO_4 60µl Hepes sodium 0.075M (7.5µl) pH 7.5 Glycerol 15µl MQ 17.5µl**	**2.0M Li_2SO_4 79µl Hepes sodium 0.075M (7.5µl) pH 7.5 Glycerol 13.5µl MQ 0.0µl**	**1.125M Li_2SO_4 45µl Hepes sodium 0.075M (7.5µl) pH 7.5 Glycerol 13.0µl MQ 35.5µl**
D	**0.5M Li_2SO_4 20µl Hepes sodium 0.075M (7.5µl) pH 8.0 Glycerol 25µl MQ 47.5µl**	**1.0 M Li_2SO_4 40µl Hepes sodium 0.075M (7.5µl) pH 6.8 Glycerol 25µl MQ 27.5µl**	**1.125M Li_2SO_4 45µl Hepes sodium 0.075M (7.5µl) pH 6.8 Glycerol 25µl MQ 22.5µl**	**1.5M Li_2SO_4 60µl Hepes sodium 0.075M (7.5µl) pH 8.0 Glycerol 25µl MQ 7.5µl**	**2.0M Li_2SO_4 80µl Hepes sodium 0.075M (7.5µl) pH 8.0 Glycerol 12.5µl MQ 0.0µl**	**1.125M Li_2SO_4 45µl Hepes sodium 0.075M (7.5µl) pH 8.0 Glycerol 12.5µl MQ 35.0µl**	**0.5M Li_2SO_4 20µl Hepes sodium 0.075M (7.5µl) pH 8.0 Glycerol 15.0µl MQ 57.5µl**	**0.5M Li_2SO_4 40µl Hepes sodium 0.075M (7.5µl) pH 6.8 Glycerol 15.0µl MQ 37.5µl**	**1.125M Li_2SO_4 45µl Hepes sodium 0.075M (7.5µl) pH 8.0 Glycerol 15.0µl MQ 32.5µl**	**1.5M Li_2SO_4 60µl Hepes sodium 0.075M (7.5µl) pH 8.0 Glycerol 15µl MQ 17.5µl**	**2.0M Li_2SO_4 79µl Hepes sodium 0.075M (7.5µl) pH 8.0 Glycerol 13.5µl MQ 0.0µl**	**1.125M Li_2SO_4 45µl Hepes sodium 0.075M (7.5µl) pH 8.0 Glycerol 13.0µl MQ 35.5µl**

(Continued……..)

	1	2	3	4	5	6	7	8	9	10	11	12
E	**0.5M Li_2SO_4 20µl Hepes sodium 0.075M (7.5µl) pH 6.8 Glycerol 10µl MQ 62.5µl**	**1.0 M Li_2SO_4 40µl Hepes sodium 0.075M (7.5µl) pH 6.8 Glycerol 10µl MQ 42.5µl**	**1.125M Li_2SO_4 45µl Hepes sodium 0.075M (7.5µl) pH 6.8 Glycerol 10µl MQ 37.5µl**	**1.5M Li_2SO_4 60µl Hepes sodium 0.075M (7.5µl) pH 6.8 Glycerol 10µl MQ 32.5µl**	**2.0M Li_2SO_4 80µl Hepes sodium 0.075M (7.5µl) pH 6.8 Glycerol 10.0µl MQ 2.5µl**	**1.125M Li_2SO_4 45µl Hepes sodium 0.075M (7.5µl) pH 6.8 Glycerol 10.0µl MQ 33.5µl**	**0.5M Li_2SO_4 20µl Hepes sodium 0.075M (7.5µl) pH 6.8 Glycerol 5.0µl MQ 67.5µl**	**1.0M Li_2SO_4 40µl Hepes sodium 0.075M (7.5µl) pH 6.8 Glycerol 5.0µl MQ 47.5µl**	**1.125M Li_2SO_4 45µl Hepes sodium 0.075M (7.5µl) pH 6.8 Glycerol 5.0µl MQ 42.5µl**	**1.5M Li_2SO_4 60µl Hepes sodium 0.075M (7.5µl) pH 6.8 Glycerol 5µl MQ 27.5µl**	**2.0M Li_2SO_4 80µl Hepes sodium 0.075M (7.5µl) pH 6.8 Glycerol 5.0µl MQ 7.5µl**	**1.125M Li_2SO_4 45µl Hepes sodium 0.075M (7.5µl) pH 6.8 Glycerol 8.0µl MQ 39.5µl**
F	**0.5M Li_2SO_4 20µl Hepes sodium 0.075M (7.5µl) pH 7.0 Glycerol 10µl MQ 62.5µl**	**1.0 M Li_2SO_4 40µl Hepes sodium 0.075M (7.5µl) pH 7.0 Glycerol 10µl MQ 42.5µl**	**1.125M Li_2SO_4 45µl Hepes sodium 0.075M (7.5µl) pH 7.0 Glycerol 10µl MQ 37.5µl**	**1.5M Li_2SO_4 60µl Hepes sodium 0.075M (7.5µl) pH 7.0 Glycerol 10µl MQ 32.5µl**	**2.0M Li_2SO_4 80µl Hepes sodium 0.075M (7.5µl) pH 7.0 Glycerol 10.0µl MQ 2.5µl**	**1.125M Li_2SO_4 45µl Hepes sodium 0.075M (7.5µl) pH 7.0 Glycerol 10.0µl MQ 33.5µl**	**0.5M Li_2SO_4 20µl Hepes sodium 0.075M (7.5µl) pH 7.0 Glycerol 5.0µl MQ 67.5µl**	**1.0M Li_2SO_4 40µl Hepes sodium 0.075M (7.5µl) pH 7.0 Glycerol 5.0µl MQ 47.5µl**	**1.125M Li_2SO_4 45µl Hepes sodium 0.075M (7.5µl) pH 7.0 Glycerol 5.0µl MQ 42.5µl**	**1.5M Li_2SO_4 60µl Hepes sodium 0.075M (7.5µl) pH 7.0 Glycerol 5µl MQ 27.5µl**	**2.0M Li_2SO_4 80µl Hepes sodium 0.075M (7.5µl) pH 7.0 Glycerol 5.0µl MQ 7.5µl**	**1.125M Li_2SO_4 45µl Hepes sodium 0.075M (7.5µl) pH 7.0 Glycerol 8.0µl MQ 39.5µl**
G	**0.5M Li_2SO_4 20µl Hepes sodium 0.075M (7.5µl) pH 7.5 Glycerol 10µl MQ 62.5µl**	**1.0 M Li_2SO_4 40µl Hepes sodium 0.075M (7.5µl) pH 7.5 Glycerol 10µl MQ 42.5µl**	**1.125M Li_2SO_4 45µl Hepes sodium 0.075M (7.5µl) pH 7.5 Glycerol 10µl MQ 37.5µl**	**1.5M Li_2SO_4 60µl Hepes sodium 0.075M (7.5µl) pH 7.5 10µl MQ 32.5µl**	**2.0M Li_2SO_4 80µl Hepes sodium 0.075M (7.5µl) pH 7.5 Glycerol 10.0µl MQ 2.5µl**	**1.125M Li_2SO_4 45µl Hepes sodium 0.075M (7.5µl) pH 7.5 Glycerol 10.0µl MQ 33.5µl**	**0.5M Li_2SO_4 20µl Hepes sodium 0.075M (7.5µl) pH 7.5 Glycerol 5.0µl MQ 67.5µl**	**1.0M Li_2SO_4 40µl Hepes sodium 0.075M (7.5µl) pH 7.5 Glycerol 5.0µl MQ 47.5µl**	**1.125M Li_2SO_4 45µl Hepes sodium 0.075M (7.5µl) pH 7.5 Glycerol 5.0µl MQ 42.5µl**	**1.5M Li_2SO_4 60µl Hepes sodium 0.075M (7.5µl) pH 7.5 Glycerol 5µl MQ 27.5µl**	**2.0M Li_2SO_4 80µl Hepes sodium 0.075M (7.5µl) pH 7.5 Glycerol 5.0µl MQ 7.5µl**	**1.125M Li_2SO_4 45µl Hepes sodium 0.075M (7.5µl) pH 7.5 Glycerol 8.0µl MQ 39.5µl**
H	**0.5M Li_2SO_4 20µl Hepes sodium 0.075M (7.5µl) pH 8.0 Glycerol 10µl MQ 62.5µl**	**1.0 M Li_2SO_4 40µl Hepes sodium 0.075M (7.5µl) pH 8.0 Glycerol 10µl MQ 42.5µl**	**1.125M Li_2SO_4 45µl Hepes sodium 0.075M (7.5µl) pH 8.0 Glycerol 10µl MQ 37.5µl**	**1.5M Li_2SO_4 60µl Hepes sodium 0.075M (7.5µl) pH 8.0 Glycerol 10µl MQ 32.5µl**	**2.0M Li_2SO_4 80µl Hepes sodium 0.075M (7.5µl) pH 8.0 Glycerol 10.0µl MQ 2.5µl**	**1.125M Li_2SO_4 45µl Hepes sodium 0.075M (7.5µl) pH 8.0 Glycerol 10.0µl MQ 33.5µl**	**0.5M Li_2SO_4 20µl Hepes sodium 0.075M (7.5µl) pH 8.0 Glycerol 5.0µl MQ 67.5µl**	**1.0M Li_2SO_4 40µl Hepes sodium 0.075M (7.5µl) pH 8.0 Glycerol 5.0µl MQ 47.5µl**	**1.125M Li_2SO_4 45µl Hepes sodium 0.075M (7.5µl) pH 8.0 Glycerol 5.0µl MQ 42.5µl**	**1.5M Li_2SO_4 60µl Hepes sodium 0.075M (7.5µl) pH 8.0 Glycerol 5µl MQ 27.5µl**	**2.0M Li_2SO_4 80µl Hepes sodium 0.075M (7.5µl) pH 8.0 Glycerol 5.0µl MQ 7.5µl**	**1.125M Li_2SO_4 45µl Hepes sodium 0.075M (7.5µl) pH 8.0 Glycerol 8.0µl MQ 39.5µl**

Appendix ii: Manual crystallization of PEG ion crystal screen with different buffer and salt composition.

	1	2	3	4	5	6	7	8	9	10	11	12
A	**PEG 15%(37.5µl) $C_2H_2MgO_4.2H_2O$ 200mM(20µl) pH 6.75, dH_2O 42.5µl**	**PEG 16%(40µl) $C_2H_2MgO_4.2H_2O$ 200mM(20µl) pH 6.75, dH_2O 40µl**	**PEG 17%(42.5µl) $C_2H_2MgO_4.2H_2O$ 200mM(20µl) pH 6.75, dH_2O 37.5µl**	**PEG 18%(45µl) $C_2H_2MgO_4.2H_2O$ 200mM(20µl) pH 6.75, dH_2O 35µl**	**PEG 19%(47.5µl) $C_2H_2MgO_4.2H_2O$ 200mM(20µl) pH 6.75, dH_2O 32.5µl**	**PEG 20%(50µl) $C_2H_2MgO_4.2H_2O$ 200mM(20µl) pH 6.75, dH_2O 30µl**	**PEG 21%(52.5µl) $C_2H_2MgO_4.2H_2O$ 200mM(20µl) pH 6.75, dH_2O 27.5µl**	**PEG 22%(55µl) $C_2H_2MgO_4.2H_2O$ 200mM(20µl) pH 6.75, dH_2O 25µl**	**PEG 23%(57.5µl) $C_2H_2MgO_4.2H_2O$ 200mM(20µl) pH 6.75, dH_2O 22.5µl**	**PEG 24%(60µl) $C_2H_2MgO_4.2H_2O$ 200mM(20µl) pH 6.75, dH_2O 20µl**	**PEG 25%(62.5µl) $C_2H_2MgO_4.2H_2O$ 200mM(20µl) pH 6.75, dH_2O 17.5µl**	**PEG 26%(65µl) $C_2H_2MgO_4.2H_2O$ 200mM(20µl) pH 6.75, dH_2O 15µl**
B	**PEG 15%(37.5µl) $C_2H_2MgO_4.2H_2O$ 200mM(20µl) pH 7.0, dH_2O 42.5µl**	**PEG 16%(40µl) $C_2H_2MgO_4.2H_2O$ 200mM(20µl) pH 7.0, dH_2O 40µl**	**PEG 17%(42.5µl) $C_2H_2MgO_4.2H_2O$ 200mM(20µl) pH 7.0, dH_2O 37.5µl**	**PEG 18%(45µl) $C_2H_2MgO_4.2H_2O$ 200mM(20µl) pH 7.0, dH_2O 35µl**	**PEG 19%(47.5µl) $C_2H_2MgO_4.2H_2O$ 200mM(20µl) pH 7.0, dH_2O 32.5µl**	**PEG 20%(50µl) $C_2H_2MgO_4.2H_2O$ 200mM(20µl) pH 7.0, dH_2O 30µl**	**PEG 21%(52.5µl) $C_2H_2MgO_4.2H_2O$ 200mM(20µl) pH 7.0, dH_2O 27.5µl**	**PEG 22%(55µl) $C_2H_2MgO_4.2H_2O$ 200mM(20µl) pH 7.0, dH_2O 25µl**	**PEG 23%(57.5µl) $C_2H_2MgO_4.2H_2O$ 200mM(20µl) pH 7.0, dH_2O 22.5µl**	**PEG 24%(60µl) $C_2H_2MgO_4.2H_2O$ 200mM(20µl) pH 7.0, dH_2O 20µl**	**PEG 25%(62.5µl) $C_2H_2MgO_4.2H_2O$ 200mM(20µl) pH 7.0, dH_2O 17.5µl**	**PEG 26%(65µl) $C_2H_2MgO_4.2H_2O$ 200mM(20µl) pH 7.0, dH_2O 15µl**
C	**PEG 15%(37.5µl) $C_2H_2MgO_4.2H_2O$ 200mM(20µl) pH 7.25, dH_2O 42.5µl**	**PEG 16%(40µl) $C_2H_2MgO_4.2H_2O$ 200mM(20µl) pH 7.25, dH_2O 40µl**	**PEG 17%(42.5µl) $C_2H_2MgO_4.2H_2O$ 200mM(20µl) pH 7.25, dH_2O 37.5µl**	**PEG 18%(45µl) $C_2H_2MgO_4.2H_2O$ 200mM(20µl) pH 7.25, dH_2O 35µl**	**PEG 19%(47.5µl) $C_2H_2MgO_4.2H_2O$ 200mM(20µl) pH 7.25, dH_2O 32.5µl**	**PEG 20%(50µl) $C_2H_2MgO_4.2H_2O$ 200mM(20µl) pH 7.25, dH_2O 30µl**	**PEG 21%(52.5µl) $C_2H_2MgO_4.2H_2O$ 200mM(20µl) pH 7.25, dH_2O 27.5µl**	**PEG 22%(55µl) $C_2H_2MgO_4.2H_2O$ 200mM(20µl) pH 7.25, dH_2O 25µl**	**PEG 23%(57.5µl) $C_2H_2MgO_4.2H_2O$ 200mM(20µl) pH 7.25, dH_2O 22.5µl**	**PEG 24%(60µl) $C_2H_2MgO_4.2H_2O$ 200mM(20µl) pH 7.25, dH_2O 20µl**	**PEG 25%(62.5µl) $C_2H_2MgO_4.2H_2O$ 200mM(20µl) pH 7.25, dH_2O 17.5µl**	**PEG 26%(65µl) $C_2H_2MgO_4.2H_2O$ 200mM(20µl) pH 7.25, dH_2O 15µl**
D	**PEG 15%(37.5µl) $C_2H_2MgO_4.2H_2O$ 200mM(20µl) pH 7.5 dH_2O 42.5µl**	**PEG 16%(40µl) $C_2H_2MgO_4.2H_2O$ 200mM(20µl) pH 7.5, dH_2O 40µl**	**PEG 17%(42.5µl) $C_2H_2MgO_4.2H_2O$ 200mM(20µl) pH 7.5, dH_2O 37.5µl**	**PEG 18%(45µl) $C_2H_2MgO_4.2H_2O$ 200mM(20µl) pH 7.5, dH_2O 35µl**	**PEG 19%(47.5µl) $C_2H_2MgO_4.2H_2O$ 200mM(20µl) pH 7.5, dH_2O 32.5µl**	**PEG 20%(50µl) $C_2H_2MgO_4.2H_2O$ 200mM(20µl) pH 7.5, dH_2O 30µl**	**PEG 21%(52.5µl) $C_2H_2MgO_4.2H_2O$ 200mM(20µl) pH 7.5, dH_2O 27.5µl**	**PEG 22%(55µl) $C_2H_2MgO_4.2H_2O$ 200mM(20µl) pH 7.5, dH_2O 25µl**	**PEG 23%(57.5µl) $C_2H_2MgO_4.2H_2O$ 200mM(20µl) pH 7.5, dH_2O 22.5µl**	**PEG 24%(60µl) $C_2H_2MgO_4.2H_2O$ 200mM(20µl) pH 7.5, dH_2O 20µl**	**PEG 25%(62.5µl) $C_2H_2MgO_4.2H_2O$ 200mM(20µl) pH 7.5, dH_2O 17.5µl**	**PEG 26%(65µl) $C_2H_2MgO_4.2H_2O$ 200mM(20µl) pH 7.5, dH_2O 15µl**

(Continued……)

	1	2	3	4	5	6	7	8	9	10	11	12
E	**PEG 20%(50µl) $C_2H_2MgO_4.2H_2O$ 150mM(15µl) pH 6.75, dH_2O 35µl**	**PEG 20%(50µl) $C_2H_2MgO_4.2H_2O$ 160mM(16µl) pH 6.75, dH_2O 34µl**	**PEG 20%(50µl) $C_2H_2MgO_4.2H_2O$ 170mM(17µl) pH 6.75, dH_2O 33µl**	**PEG 20%(50µl) $C_2H_2MgO_4.2H_2O$ 180mM(18µl) pH 6.75, dH_2O 32µl**	**PEG 20%(50µl) $C_2H_2MgO_4.2H_2O$ 190mM(19µl) pH 6.75, dH_2O 31µl**	**PEG 20%(50µl) $C_2H_2MgO_4.2H_2O$ 200mM(20µl) pH 6.75, dH_2O 30µl**	**PEG 20%(50µl) $C_2H_2MgO_4.2H_2O$ 210mM(21µl) pH 6.75, dH_2O 29µl**	**PEG 20%(50µl) $C_2H_2MgO_4.2H_2O$ 220mM(22µl) pH 6.75, dH_2O 28µl**	**PEG 20%(50µl) $C_2H_2MgO_4.2H_2O$ 230mM(23µl) pH 6.75, dH_2O 27µl**	**PEG 20%(50µl) $C_2H_2MgO_4.2H_2O$ 240mM(24µl) pH 6.75, dH_2O 26µl**	**PEG 20%(50µl) $C_2H_2MgO_4.2H_2O$ 250mM(25µl) pH 6.75, dH_2O 25µl**	**PEG 20%(50µl) $C_2H_2MgO_4.2H_2O$ 260mM(26µl) pH 6.75, dH_2O 24µl**
F	**PEG 20%(50µl) $C_2H_2MgO_4.2H_2O$ 150mM(15µl) pH 7.0, dH_2O 35µl**	**PEG 20%(50µl) $C_2H_2MgO_4.2H_2O$ 150mM(15µl) pH 7.0, dH_2O 34µl**	**PEG 20%(50µl) $C_2H_2MgO_4.2H_2O$ 150mM(15µl) pH 7.0, dH_2O 33µl**	**PEG 20%(50µl) $C_2H_2MgO_4.2H_2O$ 150mM(15µl) pH 7.0, dH_2O 32µl**	**PEG 20%(50µl) $C_2H_2MgO_4.2H_2O$ 150mM(15µl) pH 7.0, dH_2O 31µl**	**PEG 20%(50µl) $C_2H_2MgO_4.2H_2O$ 150mM(15µl) pH 7.0, dH_2O 30µl**	**PEG 20%(50µl) $C_2H_2MgO_4.2H_2O$ 150mM(15µl) pH 7.0, dH_2O 29µl**	**PEG 20%(50µl) $C_2H_2MgO_4.2H_2O$ 150mM(15µl) pH 7.0, dH_2O 28µl**	**PEG 20%(50µl) $C_2H_2MgO_4.2H_2O$ 150mM(15µl) pH 7.0, dH_2O 27µl**	**PEG 20%(50µl) $C_2H_2MgO_4.2H_2O$ 150mM(15µl) pH 7.0, dH_2O 26µl**	**PEG 20%(50µl) $C_2H_2MgO_4.2H_2O$ 150mM(15µl) pH 7.0, dH_2O 25µl**	**PEG 20%(50µl) $C_2H_2MgO_4.2H_2O$ 150mM(15µl) pH 7.0, dH_2O 24µl**
G	**PEG 20%(50µl) $C_2H_2MgO_4.2H_2O$ 150mM(15µl) pH 7.25, dH_2O 35µl**	**PEG 20%(50µl) $C_2H_2MgO_4.2H_2O$ 150mM(15µl) pH 7.25, dH_2O 34µl**	**PEG 20%(50µl) $C_2H_2MgO_4.2H_2O$ 150mM(15µl) pH 7.25, dH_2O 33µl**	**PEG 20%(50µl) $C_2H_2MgO_4.2H_2O$ 150mM(15µl) pH 7.25, dH_2O 32µl**	**PEG 20%(50µl) $C_2H_2MgO_4.2H_2O$ 150mM(15µl) pH 7.25, dH_2O 31µl**	**PEG 20%(50µl) $C_2H_2MgO_4.2H_2O$ 150mM(15µl) pH 7.25, dH_2O 30µl**	**PEG 20%(50µl) $C_2H_2MgO_4.2H_2O$ 150mM(15µl) pH 7.25, dH_2O 29µl**	**PEG 20%(50µl) $C_2H_2MgO_4.2H_2O$ 150mM(15µl) pH 7.25, dH_2O 28µl**	**PEG 20%(50µl) $C_2H_2MgO_4.2H_2O$ 150mM(15µl) pH 7.25, dH_2O 27µl**	**PEG 20%(50µl) $C_2H_2MgO_4.2H_2O$ 150mM(15µl) pH 7.25, dH_2O 26µl**	**PEG 20%(50µl) $C_2H_2MgO_4.2H_2O$ 150mM(15µl) pH 7.25, dH_2O 25µl**	**PEG 20%(50µl) $C_2H_2MgO_4.2H_2O$ 150mM(15µl) pH 7.25, dH_2O 24µl**
H	**PEG 20%(50µl) $C_2H_2MgO_4.2H_2O$ 150mM(15µl) pH 7.5, dH_2O 35µl**	**PEG 20%(50µl) $C_2H_2MgO_4.2H_2O$ 150mM(15µl) pH 7.5, dH_2O 34µl**	**PEG 20%(50µl) $C_2H_2MgO_4.2H_2O$ 150mM(15µl) pH 7.5, dH_2O 33µl**	**PEG 20%(50µl) $C_2H_2MgO_4.2H_2O$ 150mM(15µl) pH 7.5, dH_2O 32µl**	**PEG 20%(50µl) $C_2H_2MgO_4.2H_2O$ 150mM(15µl) pH 7.5, dH_2O 31µl**	**PEG 20%(50µl) $C_2H_2MgO_4.2H_2O$ 150mM(15µl) pH 7.5, dH_2O 30µl**	**PEG 20%(50µl) $C_2H_2MgO_4.2H_2O$ 150mM(15µl) pH 7.5, dH_2O 29µl**	**PEG 20%(50µl) $C_2H_2MgO_4.2H_2O$ 150mM(15µl) pH 7.5, dH_2O 28µl**	**PEG 20%(50µl) $C_2H_2MgO_4.2H_2O$ 150mM(15µl) pH 7.5, dH_2O 27µl**	**PEG 20%(50µl) $C_2H_2MgO_4.2H_2O$ 150mM(15µl) pH 7.5, dH_2O 26µl**	**PEG 20%(50µl) $C_2H_2MgO_4.2H_2O$ 150mM(15µl) pH 7.5, dH_2O 25µl**	**PEG 20%(50µl) $C_2H_2MgO_4.2H_2O$ 150mM(15µl) pH 7.5, dH_2O 24µl**

Publications

- N.A. Kuravadi, S. Verma, S. Pareek, P. Gahlot, **S. Kumari**, U.K. Tanwar, P. Bhatele, M. Choudhary, K.S. Gill, V. Pruthi, S.K. Tripathi, K.S. Dhugga and G.S. Randhawa. Guar: An Industrial Crop from Marginal Farms. Agricultural Sustainability: Progress and Prospects in Crop Research, Elsevier Science, Academic Press, 2013, Pages 47–60 (Book Chapter).

Conference Abstracts

1. **Shilpi Kumari,** Pallavi Gahlot, Shivendra Singh, Pravindra Kumar, Sangharsh Kumar Tripathi, Kanwarpal Singh Dhugga and Gurshran Singh Randhawa. ***In silico* characterization of reversibly glycosylated polypeptide in guar involved in cell-wall polysaccharides metabolism**. Indraprastha International conference on Biotechnology (IICB) 2013 poster presentation at GGSIU New Delhi 22nd-25thOctober 2013. (Page 153).

2. **Shilpi Kumari,** Pallavi Gahlot, Shivendra Singh, Pravindra Kumar, Sangharsh Kumar Tripathi, Kanwarpal Singh Dhugga and Gurshran Singh Randhawa. **Structural Analysis and 3D-Modeling of reversibly glycosylated polypeptide in guar involved in cell-wall polysaccharides metabolism**. International Conference on Biotechnology and Bioinformatics (ICBB-2014). **Got Third prize** for Poster presentation in February 1st and 2nd 2014 ICSCCB, Pune, India. (Page 129).

3. **Shilpi Kumari,** Pallavi Gahlot, Shivendra Singh, Pravindra Kumar, Sangharsh Kumar Tripathi, Kanwarpal Singh Dhugga and Gurshran Singh Randhawa. **Computational studies for structural and functional characterization of *Cyamopsis tetragonoloba* reversibly glycosylated protein (RGP).** INDO-US International Conference /Workshop on "Recent Advances in Structural Biology & Drug Discovery" (RASBDD-2014) poster presentation on at Department of Biotechnology Indian Institute of Technology Roorkee (IIT Roorkee) October 9-11, 2014. (Page 50).

4. Pallavi Gahlot**, Shilpi Kumari,** Shivendra Singh, Pravindra Kumar, Sangharsh Kumar Tripathi, Kanwarpal Singh Dhugga and Gurshran Singh Randhawa. **Expression and characterization of guar β-mannan synthase gene soluble domain in heterologous system**. Asian congress on Biotechnology (ACB-2013)-Bioprocessing for sustainable development on December 15th-19th 2013, New Delhi. India. (Page 246-247).

5. Gahlot P.[1] *, **Kumari S.[1]**, Singh S.[1], Kumar P.[1], Sondhi S.M. [2], Dhugga K.S. [3]and Randhawa G.S.[1] **Heterologous expression, characterization and *in silico* approach for structural analysis of phosphomannose isomerase (PMI) from *Cyamopsis tetragonoloba*** INTERNATIONAL PROTEOMICS CONFERENCE 6th Annual Meeting Proteomics : From Discovery to Function.PSI journal of proteins and proteomics, volume 5 number 3 JPP148 2014, ISSN: 0975-8151.