CHAPTER 1
INTRODUCTION

1.1 Lectins: an overview

Lectins are proteins of non-immune, non-enzymatic origin which bind to carbohydrates specifically and reversibly and without modifying them (Lis and Sharon, 1998). They are di or polyvalent in nature and therefore when they interact with cells like erythrocytes, they will not only combine with the sugars on their surfaces but will also cause cross-linking of the cells and their subsequent precipitation, a phenomenon referred to as cell agglutination. The hemagglutinating, activity of lectins is a major attribute and is used routinely for their identification and characterization. Lectins also form cross-links between polysaccharide or glycoprotein molecules in solution and induce their precipitation. Both the agglutination and precipitation reactions of lectins are inhibited by the sugar ligands for which the lectins are specific (Lis and Sharon, 1998). Lectins are ubiquitously present across different life forms and are involved in various biological processes such as cell–cell communication, host–pathogen interaction, cancer metastasis, embryogenesis, tissue development and mitogenic stimulation glycoprotein synthesis (Lis et al., 1998; Wragg et al., 1999; Vijayan et al., 1999; Loris et al., 2002). They are being explored intensely due to the fact that they act as recognition determinants in diverse biological processes like clearance of glycoproteins from the circulatory system, control of intracellular traffic of glycoproteins, adhesion of infectious agents to host cells, recruitment of leukocytes to inflammatory sites, as well as cell interactions in the immune system, malignancy, and metastasis. Investigation of lectins and their role in cell recognition, as well as the application of these proteins for the study of carbohydrates in solution and on cell surfaces, are making marked contributions to the advancement of glycobiology. During the past decade, there has been remarkable progress in elucidating the features of lectins that are important for carbohydrate binding. This was made possible by the refinement of old techniques and development of new ones. In particular, high-resolution X-ray crystallography of lectins in complex with their ligands allowed the identification of the chemical groups on the protein and on the carbohydrate that interact with each other and kind of bond formation which is responsible for such diverse interactions. Lectins represent a diverse group of oligomeric proteins that vary widely in size, structure, molecular organization and the carbohydrate binding site. Many of them belong to distinct protein families, classified based on biochemical, functional or structural properties. Lectins from

different plant species often differ on their molecular structure and specificity (Peumans et al., 1995).

1.2 History

The first report of the characterization of an erythrocyte agglutinating protein was given by William Hermann Stillmark, in his doctoral thesis, by the end of 19th century. He showed the presence of a highly toxic protein in the seeds of *Ricinus communicus* which were able to agglutinate the erythrocytes from human and animals and named it as "Ricin" (Sharon and Lis, 2004).

Figure 1.1 Picture of Castor Beans(www.globalsecurity.org/.../bio_ricin-pics.htm)

In 1891 H. Helen reported another protein, abrin, purified from seeds of jequirity beans (*Abrus precatorius*) showing hemagglutinating activity (Helen, 1891). In 1890`s Paul Elrich established some of the most fundamental principles of immunology, by using the property of these two proteins. In his studies on mice, he reported that mice were rendered immune by exposure to ricin or abrin and also showed that "antiabrin" which developed in the mice serum after exposure to abrin, could not neutralize the toxic effect of ricin and vice versa. He also observed that immunity to the toxin is transferred from mother to her offsprings. The earlier observations by Stillmark regarding agglutination of human and animal blood was extended by Karl Landsteiner from University of Viena in

1900. Landsteiner in his studies reported relative hemagglutinating activity of different seed extracts were different when tested in erythrocytes from different animals (Landsteiner and Raubitschek, 1907). Initial lectin studies were carried out with crude extract and not with the purified proteins, and this was significantly changed in 1936 by Summer and Howell when they purified a protein Concanavalin A, from the seeds of Jace beans (*Canavalia Ensiformnis)* and showed its hemagglutinating activity with horse erythrocytes (Sumner and Howell, 1936).

For several decades, lectins were known as hemagglutinins or phytohemagglutinins due to their erythrocytes binding capacity. Boyed and Shapleigh coined the term Lectin-based on the Latin word, 'Legere' which means to choose or select. This term was later standardized to embrace all sugar-specific agglutinins of non-immune origin, irrespective of their origin or blood type specificity (Boyd, 1963; Sharon and Lis, 1972). Another breakthrough came in lectin research in the early 1960s when Nowell and his colleagues observed that a lectin from red kidney bean was able to stimulate lymphocytes to undergo mitosis (Nowell, 1960). Aub in 1965 demonstrated that a lectin from wheat germs known as Wheat germ agglutinin (WGA) was able to agglutinate malignant cells (Aub et al., 1965). These investigations showed that the changes in cell surface glycan structure are associated with the onset of cancer and agglutination by lectins is shared property of all malignant cells. The lectin research changed dramatically with the use of Concanavalin A immobilized on dextran for affinity chromatography (Agarwal and Goldstein, 1965) which led to increased speed and ease of lectin purification and characterization.

1.3 Classification of Plant Lectins

The amino acid sequences and three-dimensional structures of hundreds of lectins, and almost all in complex with a ligand have been elucidated, and new sequences and structures are being added at an increasing rate. This vast data makes it possible to replace the classical way of dividing the lectins based on their origin, i.e., plants, animals, and microorganisms, to a classification based on features of sequences and structures. Advancement in sequencing techniques has made a huge impact by providing the primary structure of numerous lectins, which has allowed the identification of homologies in the

sequence of lectins from taxonomically related sources. During the past decade, the number of primary and 3D structures of lectin has steadily increased, with some 1400 3-D structures elucidated and reported in protein data bank (PDB) as of 2015. Also, many structures consisting of lectin-carbohydrate complexes, which gives detail molecular information of these interactions have been solved. Detail analysis of these structures has found out the similarities between the tertiary structures of lectins from diverse sources, in spite of no similarity at the sequence level. One such common tertiary structure, first observed in the legume lectins, and referred to as the lectin fold, consist characteristically of an elaborate jelly roll fold, derived from antiparallel beta- strands, arranged as two beta-sheets (Sharma et al*.,* 1996). This fold has been found in the legume lectins, the galactins, and in several other animal lectins, such as the pentraxins (Crennel et al*.*, 1994). Plant lectins are the heterogeneous group of proteins, till date several attempts has been made to classify the plant lectins based on their sugar-binding specificity, numbers of sugar- binding domains, 3D fold and domain and phylogenetic analysis. Recent studies involving analysis of complete plant genome and transcriptome has provided substantial evidence of the number of proteins having one or more carbohydrate binding domains in a complex structural design (Peumans et.al. 2001). Because of this heterogeneity among all the carbohydrate binding proteins, lectins have been classified based on the carbohydrate binding domains in their sequences into seven categories (Peumans et.al., 2001). These classes are legume lectin**s**, monocot-mannose binding lectins, jacalin- related lectins, Cucurbitaceae phloem lectins, chitin-binding lectin consisting of hevein domain, type-2 ribosome inactivating proteins and the Amaranthine lectin family.

1.3.1 Legume lectin

The most thoroughly studied family of the lectins is from leguminous plants, of which close to 200 members have been characterized (Peumans and Van Damme, 1998). Concanavalin A from Jack Bean, the prototype member of this family, was first isolated in 1919 by James Sumner, and it showed specificity for mannose and glucose. Concanavalin A was also the first lectin which was characterized in detail by X-Ray crystallography (Edelman et al., 1972; Hardman and Ainsworth, 1972). Other well-studied legume lectins are phytohemagglutinin (PHA) from the red kidney bean (Kornfeld and Kornfeld, 1970),

soybean agglutinin (SBA) (Lotan et al., 1974), peanut agglutinin (PNA) (Meehan et al., 1982). Typically, legume lectins are dimeric or tetrameric proteins of identical, or almost identical, subunits (or protomers) of 25-30 kDa, each with a single, small carbohydrate combining site with the same specificity. They also contain a tightly bound Ca2+ and a transition metal ion, predominantly Mn2+, per subunit. These metal ions are required for the interaction of lectins with carbohydrates (Sharon and Lis, 1990). The subunits of the legume lectins are commonly made up of single polypeptide chains of about 250 amino acids that may carry one or two *N*-linked oligosaccharides. In some lectins of this class (e.g., those from pea and lentil) the polypeptides are fragmented into a light (α) and heavy (β) chain (Casset.F et.al. 1995). The subunits are in the shape of a dome, made up largely of two antiparallel β -sheets, one of six strands and the other of seven. The six stranded sheet is almost flat while the other is concave. The strands of the sheets form jellyrolls also referred to as the lectin fold. The majority of residues not included in the β structures are in loops and β bends that connect the strands of the β sheets. The combining sites of the carbohydrate and of the metal ions are located mostly in the β folds of the seven-chain, curved sheet. Four loops located at the upper part of the dome forms the carbohydrate binding site. Legume lectins are devoid of any α-helical structure and thus belong to class of all beta proteins (Banerjee et al., 1994)

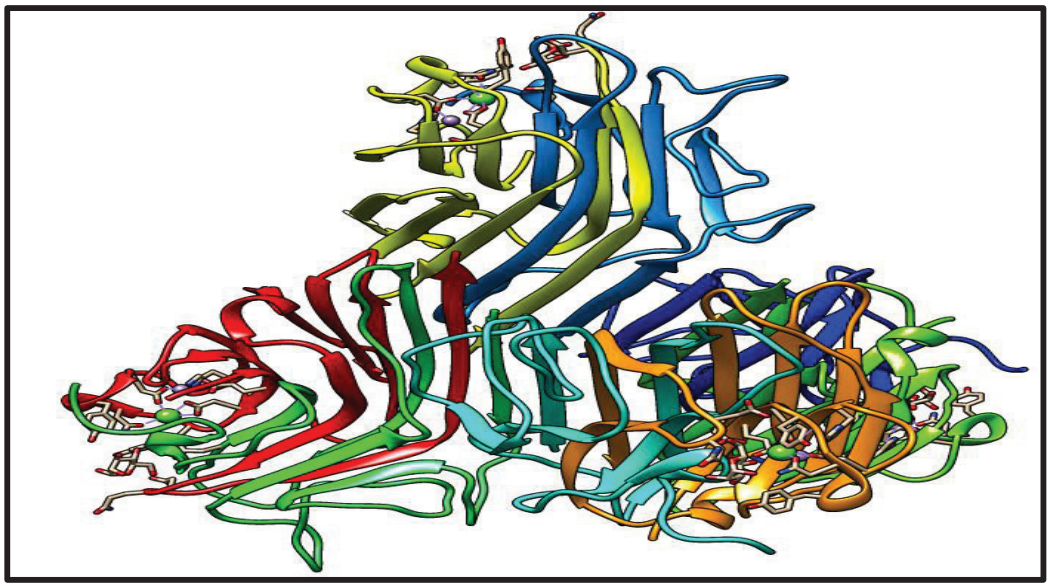

Figure 1.2 Crystal structure of Flt3receptor interacting lectin from seeds of *Dolichos lablab* (PDB Code:1QMO)

Legume lectins interact with their carbohydrate ligands through side chains of three invariant amino acids such as asparagine, aspartic acid and an aromatic amino acid (Sharon 1993, Adar and Sharon 1996) or leucine. Mutation of asparagine or aspartic acid to other amino acid leads to losing of carbohydrate binding activity. Legume lectins have shown specificity towards carbohydrate moieties which are not present in plants such as complex glycan structure with sialic acid at terminal position or Thomson-nouveau antigen (Tn). Due to their specificity towards typical animal glycans, the role of these class of lectins in plant defense against insects or predating animals has been speculated.

1.3.2 Chitin binding lectins

The chitin binding lectins are also referred as "hevein domain lectins" so named from one of the members derived from the latex of the rubber tree (*Hevea brasilienesis*) (Gidrol et al., 1994). Chitin binding lectins are exclusively found in cereals. Proteins belong to this class like legume lectins consist of two identical subunits and are exceptionally rich in cysteine. Some chitin binding proteins are built up of one or more hevein domain linked to catalytically activated chitinase domain.

Wheat germ agglutinin is one of the most well-studied lectins in this family; it is a mixture of three isolectins that differ slightly in their amino acid compositions. The isolectins are dimers of two identical 17kDa subunits and are devoid of metals. Each subunit is made up of four homologous subdomains (A to D) of 43 amino acids each; these domains are similarly folded, with four identically positioned disulfide bridges. In the primary amino acid composition of wheat germ agglutinin, 20% of residues are half cysteine and 21% are glycine (Allen et al., 1973). Unlike most of the lectins, this protein is not a glycoprotein. These proteins show more powerful binding to oligosaccharide than the monosaccharide and the binding site consists of three or four

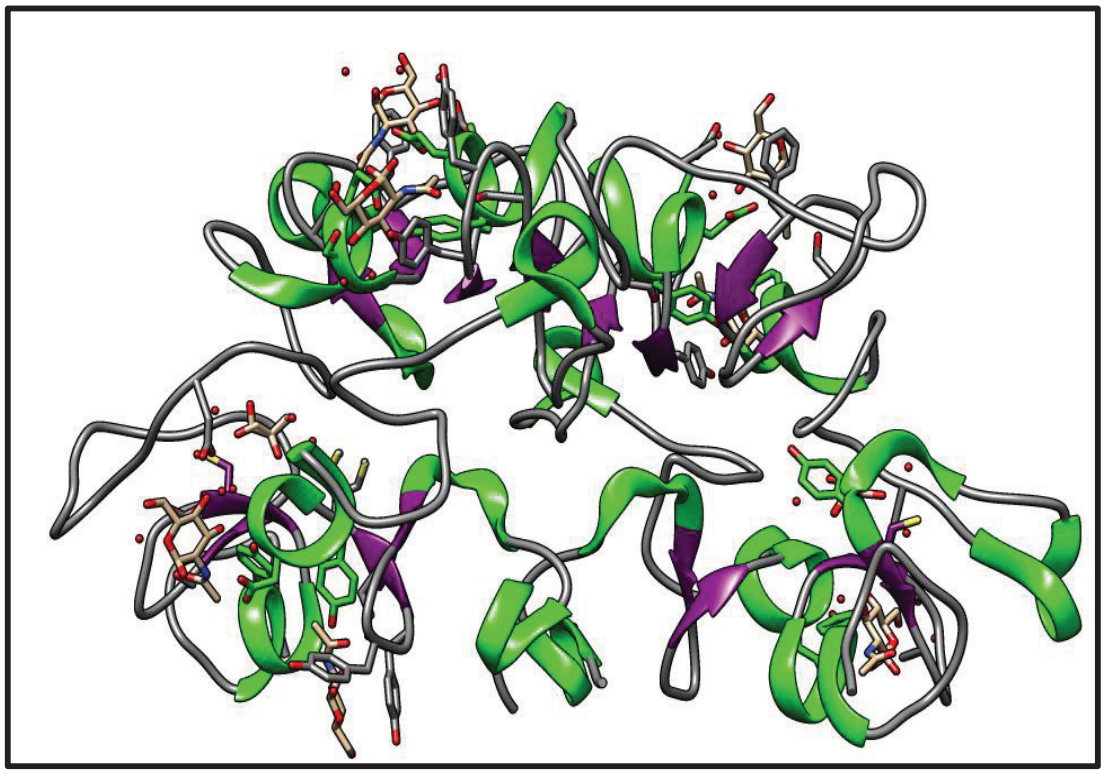

Figure 1.3 Crystal structure of Wheat Germ Agglutinin (WGA) in complex with N-Acetyl-D-glucosamine (PDB Code: 2UVO)

subsites with differing specificity (Lotan and Sharon, 1973). The binding interactions of N-acetyl-D-neuraminic acid (Greenaway and LeVine, 1973) and N,N' diacetyl-chitobiose (GlcNAc-beta-1,4-GlcNAc), were observed in crystal complexes of wheat germ agglutinin (WGA) at four independent sites/monomer (Muraki et al., 2002). There is a hydrogen bond between Ser62 and the non-reducing end of GlcNAc. In addition to the hydrogen bonding and van der Waals interactions, CH--pi interactions involving Tyr64, His66 and Tyr73 were suggested to play an essential role in the ligand binding conformation. The presence of multiple binding sites due to the internal 4-fold structure repeat and their location at the interface between the subunits forms the molecular dimer of the lectin (Muraki et al.,2002). Lectins consisting of hevein domain have also been isolated from cereals.

1.3.3 Monocot Mannose Binding Lectins

The first member of this family of lectin was isolated from the bulb of *Galanthus nivalis* (*Amaryllidaceae*). As the name suggest proteins from these family have a strong affinity towards mannose of mannose-containing N-glycans. Like *Galanthus nivalis* lectin, mannose-binding lectins are widely distributed in all monocot plant families such as Amaryllidaceae, Allicaceae, Araceae, Orchidaceae, Lillicaceae and Bromeliaceae (Barre A et al. 1996). Almost, all the monocot mannose-binding lectins are consisting of one or more subunit, and the distinguishing feature of this class of proteins is small monomer size (12 kDa), lack of metal ion requirement for carbohydrate binding and internal three-fold repeats of 36 amino acids. The sequence similarity (80 to 90%) of these class of lectins clearly suggests an evolutionary relationship. The three-dimensional structure of one of these lectins from blub of *Scillia campanulata* is described in detail. It is a flat tetrameric molecule with a central opening 16 Å wide. Each monomer contains three subdomains (1, 2, and 3) with antiparallel, four stranded, β -sheet structures (Figure1.4). There is one inter-subdomain disulfide bond, between the second and third subdomains, and the interior of the monomer is stabilized by conserved hydrophobic residues. The carboxy-terminal subdomain 1 includes one strand from the adjacent subunit, and the four subunits hence form two pairs of dimers (A-D and B-C) within the tetramer (Wright and hester, 1996).

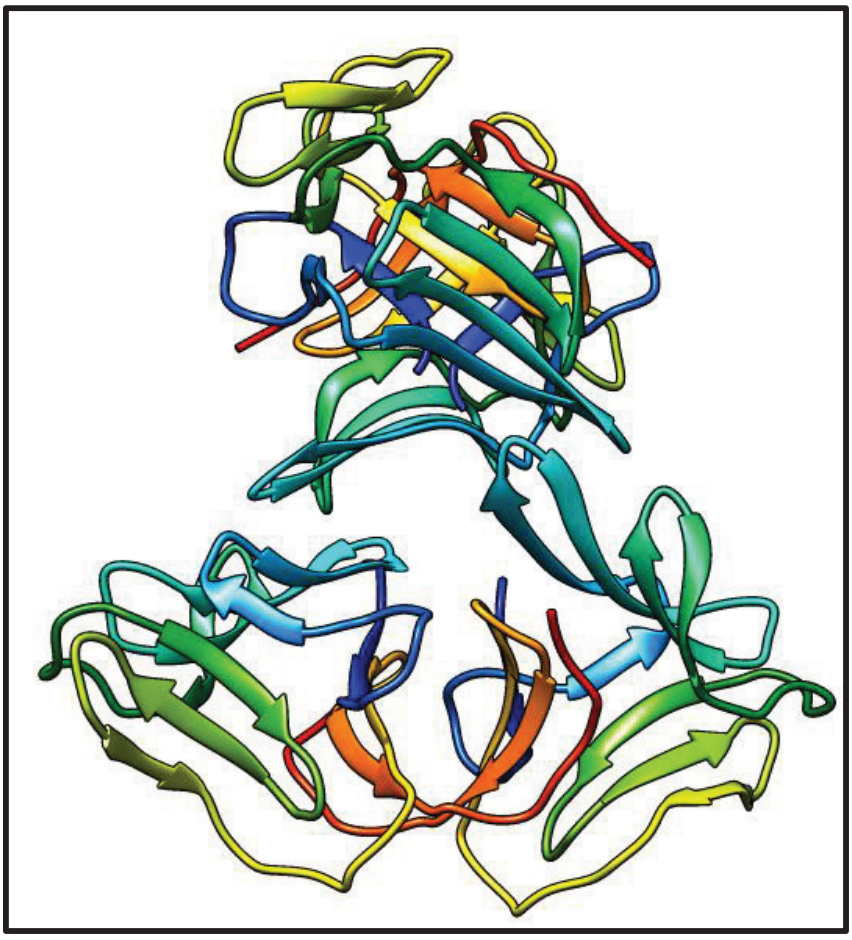

Figure 1.4 Crystal structure of mannose specific bulb lectins from *Scillia campanulata* (Blue bell) (PDB Code: 1B2P)

1.3.4 Jacalin-related lectins

Jacalin, the galactose-specific lectin from the seeds of jackfruit, *Artocarpus integrifolia* (*Moraceae*), is a tetrameric glycoprotein with a molecular weight of about 66 kDa. Each of its subunits consists of a heavy chain (β) of 133 amino acids and a light chain (α) of 20 residues The subunits of jacalin show a β-prism-I fold, made up of three four-stranded antiparallel β-sheets, arranged like the faces of a triangular prism, with loops connecting strands in the sheets (Sankaranarayanan et al., 1996). The structure of tetrameric jacalin in complex with Gal, Me-alpha-GalNAc, Me-alpha-T-antigen, GalNAcbeta 1-3Gal-alpha-O-Me has been solved through X-Ray crystallography (Jeyaprakash et al., 2003). The interactions of the disaccharide at the binding site are

predominantly through the GalNAc moiety. The Gal interaction is through water molecules by making a hydrogen bond between the anomeric oxygen of GalNAc and the pi electrons of an aromatic side-chain. (Jeyaprakash et al., 2003). Also, there are several intermolecular interactions involving the binding of carbohydrate that contributes towards the stability of the crystal structure.

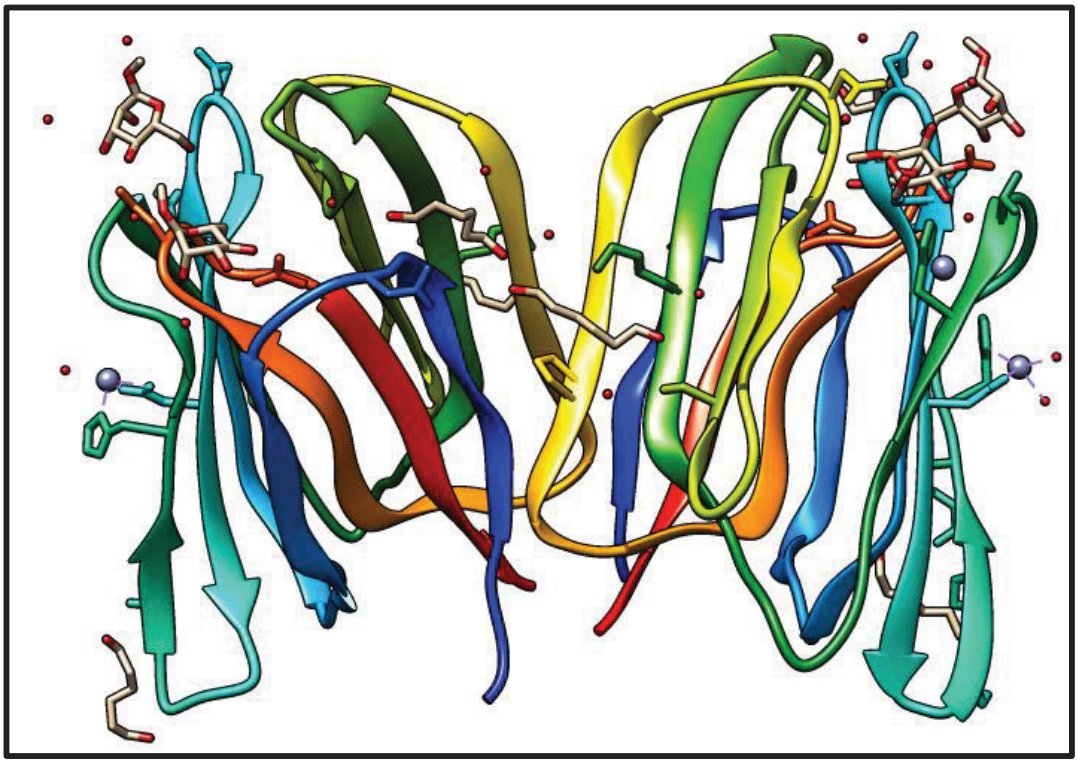

Figure 1.5 Crystal structure of Jacalin related lectin complex with Methyl-α-mannose from *Musa acuminata* (Banana) (PDB Code: 1X1V)

The family of Jacalin related lectins is subdivided into galactose-specific hemagglutinin and mannose-specific hemagglutinin (Peumans and Van Damme, 1998; Peumans et al., 2001). Galactose- specific agglutinins are made up of four identical protomers each consisting of a heavy and light chains. These chains are derived from a large proprotein through co and post-translational modifications and are accumulated in storage protein vacuoles. The mannose-specific agglutinin consists of two, four or eight intact subunits. These proteins do not undergo posttranslational modification, and the exact relationship between the two subfamilies of jacalin-related lectins is unclear.

1.3.5 Cucurbitaceae phloem lectins

The lectins from the phloem of the plants are also called phloem proteins or PP2. This family is unrelated to other plant lectins families and most of these lectins form dimers in solution with high specificity towards chitin. These class of lectins was first identified as the component of P-proteins in cucurbits (Bostwick et al., 1994). Cucurbits have been used as a model plant for many phloem studies and the *Cucurbita* spp. contain two predominant P-proteins, phloem filament protein or phloem protein 1 (PP1) and the phloem lectin or phloem protein 2 (PP2) (Smith et al., 1987). PP2 is a dimeric poly-GlcNAc-binding lectin covalently linked by disulfide bridges (Read and Northcote, 1983).

Studies showed that PP2-A1 also bind to high-mannose N-glycans and 9-acyl N-acetylneuraminic sialic acid (Beneteau et al., 2010). The *Cucumis* spp., cucumber (*Cucumis sativus*) and melon (*Cucumis melo*) contain two lectins with apparent molecular masses of approximately 26kDa, and 17kDa and the difference between these two lectins are the absence of 62 amino acids from the N-terminus region of 17kDa proteins. This suggests that the 62 amino acids present in the N-terminus region are not essential for the lectin activity. The sequence analysis of the phloem lectins present in the *Cucumis* spp. was even greater when compared with other genera within the Cucurbitaceae family. The 26kDa *Cucumis* spp. Lectins shared only 51.6% (melon) and 47.5% (cucumber) sequence identity with the phloem lectin present in the winter squash (*Cucurbita maxima*; CbmPP2). The sequence identity of 17kDa *Cucumis* spp. Lectins show 44.3% identity with melon and 39.7% sequence identity with cucumber (Dinant et al., 2003). The other lectin isolated from the pumpkin phloem contains two abundant phloem lectins, PP1, a 96 kDa protein that forms polymeric filaments *in vivo*, and PP2, a 48kDa dimeric lectin (Bostwick et al., 1992).

1.3.6 Type II ribosome inactivating proteins

Many proteins in plants are capable of inactivating the ribosome and thus called as ribosome inactivating protein or RIPs (Peumans and Van Damme, 2001). These class of proteins is widely distributed in the plant kingdom (Stripe and Barbieri, 1986). Ribosome-inactivating proteins are broadly classified into two groups: type I and type II. Type I RIPs are a diverse family of proteins comprising a single polypeptide chain, whereas type II RIPs are heterodimeric glycoproteins (Puri et al., 2012). Type II ribosome inactivating

proteins are composed of a Polynucleotide: adenosine glycosidase domain (A-chain) tandemly arranged with a carbohydrate binding domain. The lectin in this family was the first lectin isolated from plants, in 1888 from the seed of castor which contains two closely related lectins, *Ricinus communis* agglutinin (RCA) and ricin. Ricin is a heterodimeric protein with a MW of 60kDa, made up of two S- S linked chains, A and B and is one of the most poisonous substances known. RCA is a dimer of two subunits, each of which is similar to ricin but it is not toxic (Olsnes, 2004; Balint, 1974; Funatsu and Funatsu, 1968). The other type II RIP's from this family are also highly poisonous and include *Sambucus nigra* lectin, cinnamomin lectin, and ebulin (Girbes et al., 2003; Xu and Liu, 2004; Spivak and Hendrickson, 2005). The sequence of all the type II ribosome inactivating proteins have a high sequence similarity, a similar three-dimensional structure and are widely spread in the plant kingdom and as a superfamily of evolutionarily related proteins.

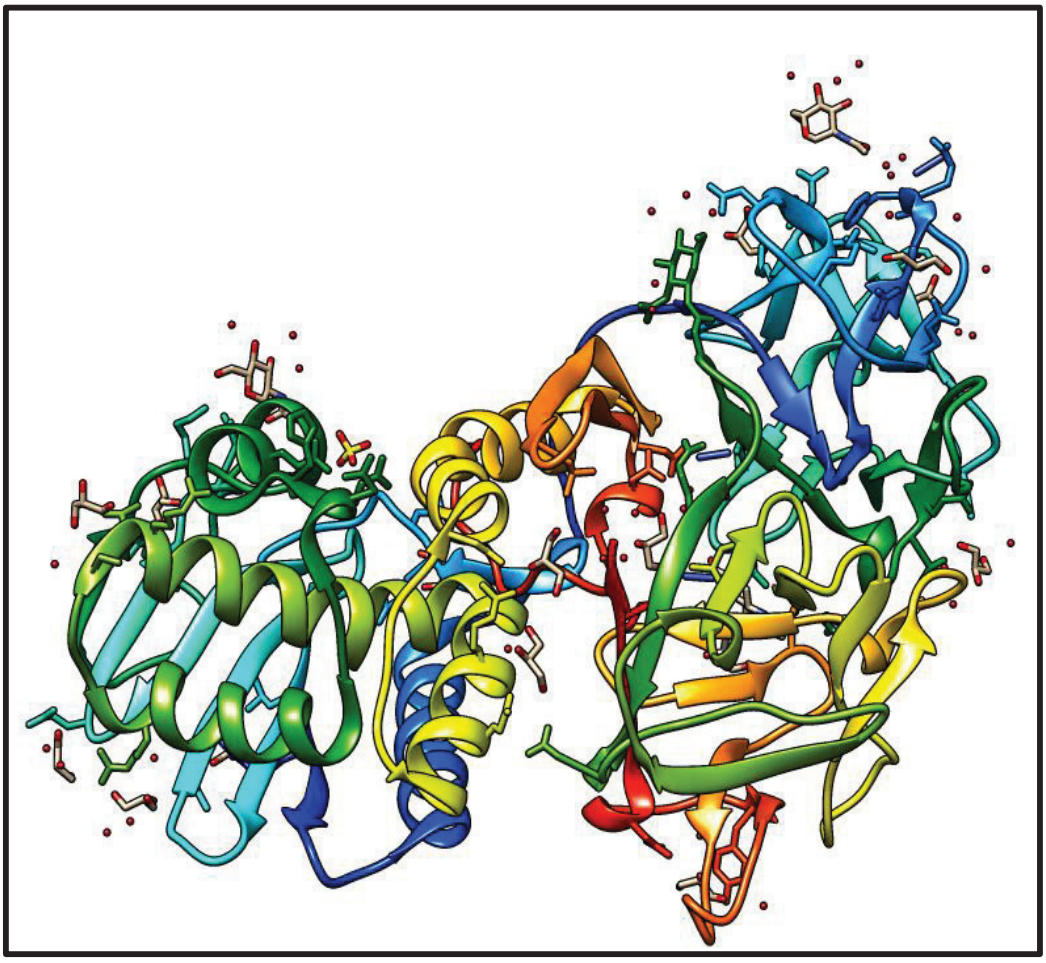

Figure 1.6 Crystal structure of *Viscum album* mistletoe lectin (PDB Code: 2RG9)

One of the most studied members of this family is ricin, and the three-dimensional structure of ricin has been determined by X-ray crystallography. The A-chain is 32,000-dalton glycoprotein with extensive secondary structures, both β-pleated sheet, and α-helix (Montfort, 1987). The B-chain is a lectin with a binding preference for galactosides from cell surface carbohydrates and is made of three copies of a primitive 40 residue folding unit, which packs around a pseudo threefold axis (Olsnes et al., 1975). In each domain, galactose binds in a shallow cleft formed by a three residue peptide kink on the bottom and an aromatic ring on the top. At the back of the cleft, an aspartate forms hydrogen bonds to the C3 and C4 hydroxyls of galactose, whereas a glutamine bonds to the C4 alcohol, helping to define specific epimer binding (Rutenber and Robertus, 1991). The A-chains are covalently associated, with a disulfide bridge between Cys157 of each of the chains. The disulfide link between the chains does not appear to be critical for toxicity except in maintaining protein-protein interaction at very low toxin concentration. The subdomains contribute conserved Trp, Leu, and Ile residues to a compact central hydrophobic core and this tight threefold binding probably drive the peptide folding and stabilizes the protein structure (Rutenber and Robertus, 1991).

1.3.7 The amaranthin lectin family

Amaranthin represents new lectin family, and all the lectins are isolated from amaranthus species till now (Wright, 1997). The seed lectin "amaranthin" isolated from the *Amaranthus caudatus* is a heterodimer (62kDa) composed of a lectin polypeptide of 33-36 kDa (Rinderle et al., 1989). The X-ray diffraction data analysis and the three-dimensional structure revealed that amaranthin doesn't resemble any other plant lectin. The polypeptide folds present in amaranthin share remarkably similar features and consists exclusively of the β structure. In the three-dimensional structure of amaranthin, the beta-sheet subdomains are inter-related by a pseudo three-fold symmetry which assembles to form beta-prism or beta-barrel structures, and the beta-barrel structures are stabilized by a hydrophobic core (Wright, 1997). Amaranthus lectins are specific for GalNAc and recognize human T cells (Porras et al., 2005; Hernandez et al., 2004). The lectin isolated from other *Amaranthus* species like *Amaranthus hypochondriacus, Amaranthus*

leucocarpus, Amaranthus viridis and Amaranthus spinosus are very similar to amaranthin (Ozeki et al., 1996; Kaur et al., 2006 a; Zenteno et al., 1992).

1.4 Physiological role of plant lectins

The variety in the sugar binding specificity and the structures between the different lectin families indicates lectins have varied functions. To understand the physiological role of plant lectins, it is very important to identify the potential lectin binding receptors. A break though was achieved in 1970's when it proved that the plant lectins not only play a major role in the plant itself but also recognize the foreign bodies or glycoconjugates of other organisms. Based on these observations scientist at that time proposed two main theories. The first theory was that lectins protect plants against pathogenic microorganisms, insects, and predatory animals, and the second theory assumed that they are lectins involved in the association between the symbiosis of leguminous plants and their nitrogen-fixing bacteria.

1.4.1 Role of lectins in plant symbiosis & Nitrogen fixation

Nearly three decades ago the role of lectins as a mediator between the nitrogen-fixing bacteria and leguminous plants was proposed by (Bohlool and Schmidt, 1974; Hamblin and Kent, 1973; Dazzo and Hubbell, 1975). It was proposed that lectins presents in seeds of leguminous plants recognize the carbohydrate moieties present on the cell surface of rhizobia and assist in the initial phase attachment of Rhizobium to the root epidermal cells. When rhizobia encounter the root hairs, they bind to it and releases nod factors in the root cortex of plants. The nodule formation and rhizobial entry are host-strain specific, and this specificity is determined by the type of nod factors produced by a particular rhizobial strain. This ability to recognize the specific type of nod factor is different in different leguminous species. For example, the rhizobia that infect and nodulate the soybeans cannot nodulate garden peas, and *vice versa* (Bohlool and Schmidt, 1974) .

1.4.2 Lectin in plant defense

The plant lectins play an important role in plant defense. The toxicity of various plant lectins for animals and their growth inhibitory effect on fungus, bacteria, virus and pests led to assumptions that lectin has a strong role to play in plant defense.

1.4.2.1 Anti-bacterial activity of lectins

Lectin plays a major role in the plant defense against bacteria because of a large number of glycoproteins or glycoconjugates present on the bacterial cell surface, and many studies indicate plant lectins can recognize carbohydrate moieties present on the peptidoglycans, building block of cell wall component of many bacteria (Ayouba et al., 1994). These lectin-carbohydrate interactions cannot alter the structure or the permeability of the membrane or disturb the normal intracellular processes of invading microbes. How these plant lectins prevent the growth of micro-organisms is yet to be answered, but one of the hypothesis suggests, lectin agglutinates micro-organisms preventing subsequent growth and colonization. The role of plant lectins as antimicrobial agent began in 1936 when Sumner and Howell reported Concanavalin A can agglutinate various Mycobacterium species (Sumner and Howell, 1936). Plant lectins also prevent the formation of biofilm which is necessary bacteria to survive and grow (Hasan et al., 2014). Examples of different lectins inhibiting the growth of different bacterial strains are given in table 1.1.

Table 1.1 Plant lectins exhibiting antibacterial activity

Natural source of lectin	Bacterial strain	Reference
Potato (*Solanum tuberosum*) lectin	*Pseudomonas solanacearum*	Sequeira and Graham,1977
Archidendron jiringa lectin	*Staphylococcus aureus, Bacillus subtilis*	Charungchitrak et al., 2011
Curcuma longa lectin	*Bacillus subtilis, Staphylococcus aureus, Escherichia coli, and Pseudomonas aeruginosa*	Petnual et al., 2010

1.4.2.2 Antifungal Activity

Only a few of the large numbers of lectins and hemagglutinins that have been purified have been explored for antifungal activity. Chitinase-free chitin-binding stinging nettle (*Urtica dioica*) lectin has been reported to retard the growth of fungus by interrupting the cell wall synthesis (Van Parijs et al., 1991). The other plant lectins, which inhibit the growth of fungus includes small chitin-binding merolectins with one chitin-binding domain, e.g., hevein from rubber tree latex (Van Parijs et al., 1991) and chitin-binding polypeptide from *Amaranthus caudatus* seeds (Broekaert et al., 1992). The only plant lectins that show fungicidal proteins are the chimeric lectin belonging to the class I chitinases. However, the antifungal activity of these proteins is due to their chitinase domain and not due to lectin domain (Wong et al., 2010).

Table 1.2 Plant lectins exhibiting Antifungal activity

Natural source of lectin	Fungal species inhibited	Sugar specificity	Reference
Amaranthus viridis (Green Amaranth) seeds	*Botrytis cinerea, Fusarium Oxysporum*	Asialofetuin, fetuin, T-antigen, N-Acetyl-D-lactosamine, N-Acetyl-D-galactosamine	N.Kaur, et al., 2006
Withania somnifera (Ashwagandha/ Indian ginseng/Winter cherry/Ajagandha/Kanaje Hindi) leaves	*Fusarium moniliforme, Macrophomina phaseolina*	Methyl-D-Mannopyranoside	Ghosh. M, 2009
Zea mays (maize) endosperm	*Aspergillus flavus*	D-galactose	Baker et al., 2009
Capsicum frutescens (red cluster pepper) seeds	*Aspergillus flavus, Fusarium Moniliforme*	D-mannose, glucose	Ngai et al., 2007
Astragalus mongholicus (huangqi) roots	*Borrytis cinerea, Colletrichum sp., Droschslara turia, Fusarium oxysporum*	D-galactose, lactose	Yan, Q et al., 2005
Pouteria torta (pouteria trees/eggfruits) seeds	*Saccharomyces carevisiae, C. musae, Fusarium oxysporum*	Fetuin, asialofetuin, heparin, orosomucoid, ovoalbumin	Boleti A et al., 2007

1.4.2.3 Anti-insect activity of lectins

Plant lectins can have severe effects on fecundity, growth, and development of an insect. A wide range of the plant lectins screened against insects demonstrated that the *Galanthus nivalis* agglutinin (GNA) has the highest anti-insect potential (Rahbe et al., 1995). The snowdrop lectin isolated from *Galanthus nivalis* is shown to be toxic towards rice brown plant hopper (*Nilaparvata lugens*) (Vandenborre, 2011). The *Galanthus nivalis* lectin specifically binds to high-mannose glycans, in particular to terminal mannose residues (Hilder et al., 1995). Lectins are one of the promising agents against insect pests and have been engineered as transgenic and being explored into a variety of crops including wheat, rice, tobacco, and potatoes (Lam and Ng, 2011). Sauvion and others indicate that neonates of the green peach aphid (*Myzus persicae*) fed with an artificial diet containing GNA, *Narcissus pseudonarcissus* lectin (NPA), *Allium sativum* agglutinin (ASA, garlic), or Concanavalin A (ConA, jack-bean *Canavalia ensiformis*), showed increased mortality and a negative effect on weight gain (Sauvion et al., 1996). The purified lectins from snowdrop (*Galanthus nivalis*) and garlic *(Allium sativum*) used for feeding trials indicated that they are moderately active against chewing insects, such as the cowpea weevil and the tobacco horn worm (*Spodoptera litoralis)*. Lectins increase the mortality or delay the development of insect when it is incorporated into an artificial diet, e.g., *Arisaema jacquemontii* lectin adversely affected the development of *Bactrocera cucurbitae* larvae (Kaur et al., 2006b). The glycosylation patterns in the insect may change depending on the developmental stage or other factors, and therefore, lectin-binding patterns may differ accordingly.

Table 1.3 Plant lectins showing anti-insect activity

Natural source of lectin	Insect affected	Anti-insect effect	Sugar specificity	Reference
Allium sativum (garlic) bulbs	*Acyrthosiphon Pisum*	Increased mortality	Mannose	Fitches et al., 2008
Arisaema intermedium & *Arisaema wallichianum* (Araceae)	*Bactrocera Cucurbitae*	(1) Prolonged period of development (2) Inhibited pupation and emergence	Not found	Kaur et al., 2009
Myracrodruon urundeuva (aroeira preta) bark	*Aedes aegypti*	Increased mortality	N-Acetyl-D-glucosamine	Sa et al., 2009
Galanthus nivalis	*Nilaparvata lugens*	Increase mortality	D-Mannose	Foissac et al., 2000

1.5 Applications of plant lectins

Lectins are ubuquitous across the living kingdom, and they are instrumental in protein-carbohydrate recognition, such as cell-cell communication, host defense, tumor metastasis. Because of their specificity for carbohydrates, lectins have been used as tools for analytical and preparative purposes.

1.5.1 Plant lectin in cancer treatment

Carbohydrate-binding properties of lectins make them a very useful marker to differentiate cancer cells from normal cells using various histochemical, biochemical techniques. Lectins are widely used in histochemical studies of tumors from epithelial tissues. The antitumor effect of lectins is well documented, and the mechanism of action for the antitumor effect of lectins has been elucidated to widen their application as antitumor drugs (Table 1.4). Lectins elicit apoptosis in different cancer cell lines and, activation of different caspases (mainly caspase-3) is usually observed (Bhutia et al., 2009). Lectins interact with the FAS receptor to activate the apoptotic pathway. In studies the Bcl family members (anti-apoptotic factors) were down-regulated and G0/G1 arrest was frequently observed in lectin-mediated apoptosis (Bhutia et al., 2009, b; Liu et al., 2009; Lam and Ng 2010).

Table 1.4 Effect of plant lectins on growth of cancer cell lines

Lectin	Cell line whose proliferation was inhibited	Reference
Black soybean (*Glycine max*) lectin	Breast cancer MCF7 cells and hepatoma HepG2 cells	Lin et al., 2008
Del Monte banana lectin	leukemia L1210 cells and hepatoma (HepG2) cells	Cheung et al., 2009
Autumn purple bean lectin	hepatoma HepG2	Fang et al., 2010
Sophora flavescens Abrus agglutinin	HeLa cells Dalton's lymphoma cells HeLa cells	Liu et al., 2008, Bhutia et al., 2009

A study showed that European Mistletoe (*Viscum album,* L.) lectin stabilized with alginate/chitosan microcapsules coated by a biodegradable polymer wall can be used to protect the lectin from acidic pH in the stomach, and used in the cancer therapy (Lyu et al.

2004).

1.5.2. Lectins as antiviral agents

Lectins have also been investigated for their antiviral effects e.g. banana (*Musa acuminata*) lectin has been reported to bind directly to the gp120 envelope protein of HIV-1 and block entry of the virus into the cell and decrease the product of early reverse transcription (Swanson et al., 2010). Studies report the inhibition of HIV-1 reverse transcriptase by Extralong autumn purple bean lectin (Fang et al., 2010) and mushroom *Russula delica* lectin (Zhao et al., 2009). Besides these lectins mentioned above, several other lectins such as concanavalin A, wheat germ agglutinin, *Lens culinaris* agglutinin, *Vicia faba* agglutinin, *Pisum sativum* agglutinin and phytohaemagglutinin are reported to inhibit fusion of HIV-infected cells with CD4 cells by binding to HIV-1 gp120 envelope molecule (Hansen et al., 1989).

1.5.3 Blood Type Specificity

In 1945 William et al. discovered that lectins could be blood group specific; some lectins being able to agglutinate the red cells of one type but not those of another. He discovered that lima bean lectin would agglutinate red cells of human blood type A but not those of O or B and based on these observations Boyd coined the term 'lectin' which is Latin for Legere, or 'to choose' (Sharon and Lis, 1972). The seeds of *Lotus tetragonobolus* can agglutinate group O specifically, and *Bandairaea simplicofolia* is specific to group B. The specificity of lectins is so sharply defined that they can differentiate among blood subgroups e.g. *Dolichos biflorens* lectin reacts more vigorously with blood group Al than A2. Other blood groups can be distinguished by lectins, such as M and N blood types (Landsteiner and Raubitschek, 1907).

1.5.4 Lectin microarray

Lectin microarray is used for analyzing the glycan structure present on glycoproteins or glycolipid. Lectin microarray has become a very useful technique in the field of glycomics and glycoproteomics. Lectin microarray has been successfully utilized in many areas of biology such as detection of cancer cells or bacteria typing. In lectin

microarray system, fluorescent-tagged lectins are immobilized on a single chip in microarray formation. Lectin microarray has been successfully utilized in the quantitative analysis of lectin-glycoproteins interactions by analyzing the dissociation constants (Kd) between the immobilized lectins and glycoproteins. Lectin microarray involving *wisteria floribunda* agglutinin was successful in detecting the difference between cancerous and normal epithelial cells with high accuracy (Matsuda et al., 2008). Lectin microarray was successfully utilized in bacteria typing in different *E. coli* strains based on the change in the surface glycan structure (Hsu et al., 2006).

1.6 Non-carbohydrate ligand binding

According to the scientific definition of lectin, Lectins are carbohydrate-binding proteins of non-immune origin. But in last few years, many studies have shown the presence of some of the lectins which possess one or more binding sites for non-carbohydrate ligands, which are independent of the carbohydrate ligand binding sites.

Grass family of the plant contains many important crop plants such as wheat, rice, maize, and sorghum. These plants contain a gene which codes for a chimeric lectin which binds with β-glucosidase enzyme and aggregates it. The studies showed that carbohydrate binding site and β-glucosidase binding site are present on lectin domain but are separate from each other (Blanchard et al., 2001). Some plant lectins were found to strongly bind, *in vitro*, some hydrophobic biomolecules like adenine and adenine-based plant hormones (Roberts and Goldstein, 1982, 1983a, b; Maliarik and Goldstein, 1988; Gegg et al., 1992). Equilibrium studies in solution showed that legume lectins bind adenine and related molecules in a stoichiometry of one (or two) per lectin molecule. Interestingly, the crystal structure of *Dolichos biflorus* seed lectin complex with adenine shows two identical hydrophobic cavities within the lectin matrix, with each cavity accommodating two adenine molecules (Hamelryck et al., 1999). Besides adenine and adenine-related phytohormones, binding of other biomolecules such as kinetin, zeatin, isopentyl adenine as well as abscisic and gibberellic acids to Wheat Germ Agglutinin (WGA) has been reported (Bogoeva et al., 2004).

1.7 Motivation and objectives of the work

A highly efficient coding system is necessary to allow cells to communicate swiftly through complex surface recognition and interactions. Complex carbohydrate surpasses amino acids and nucleotides in their information storing capacity and serves as important ligand throughout the process of recognition in a living system. Ability to decipher the biological code written in carbohydrate moieties is one of the biggest strength of lectins. The multivalency of lectins makes sure accurate signal decoding and transmission of information. The emerging medical importance of protein-carbohydrate recognition in combating infections and the spread of tumors or in targeting drugs also explains the need to isolate continuously and characterization of novel lectins from a different source. Among plant lectins, legume lectins have been studied in great details providing the information about their primary and tertiary structures. This in-depth analysis has led to the utilization of this lectins in many other areas of research as well as commercial applications.

There are many class of plant lectins which have not been characterized in great details as legume lectins and also, there is evidence of proteins which are not classified as lectins but show carbohydrate binding capacity. Plant lectins will continue to have an important role and the understanding of their carbohydrate-binding specificity, and the three-dimensional structures might allow the lectins to be engineered to recognize specific glycoconjugates. This will help the biologists to identify the phenomenal part of carbohydrates which play in the cellular interactions and regulation and also contribute to a better understanding of the natural functions of these interesting plant proteins.

In continuation of previous work from our laboratory, on plant and microbial lectins, we propose to isolate and characterize lectins from selected important crop plants. These three plants are *Solanum tuberosum* (potato), *Sorghum bicolor* (jowar), and *Gossypium aerboreum* (cotton). The purified novel lectins will be characterized with respect to hemagglutination activity and carbohydrate specificity. Also purified protein will be characterized with respect to its secondary structure, effect of carbohydrate binding on secondary structure, pH, and thermal stability. Fluorescence quenching studies will be done to find out the thermodynamic parameters of carbohydrate binding and to probe the

exposure of tryptophan residue and their microenvironment in purified proteins. Bioinformatics analysis will give an in-depth idea of purified proteins and their sequence similarity with other known lectins, their evolutionary pattern. Molecular Docking and Molecular Dynamics Studies will give insight into the protein-ligand interactions. Protein crystallization will be carried out to obtain the crystals to solve the 3-D structure of purified proteins and understand the mechanistic basis for their sugar interactions.