Chapter 1

General Introduction

1.1 Introduction

Lectins are a group of multivalent carbohydrate-binding proteins of non immmune origin and they exert biological effects through their ability to specifically bind different carbohydrate structures (Chandra *et al.*, 2006; Bies *et al.*, 2004; Jepson *et al.*, 2004; Sharon & Lis, 1989a, 1989b, 2004; Nishi *et al.*, 2003; Vijayan & Chandra, 1999; Loris *et al.*, 1998; Liener *et al.*, 1986). Lectin was discovered in 1888 by Stillmark. He observed the ability of castor bean extracts to agglutinate animal erythrocytes. The term 'lectin' (derived from the Latin, 'legere' - to select) was first coined by Boyd and Shapleigh (1954) based on the observation that some plant seed extracts could distinguish between different human blood groups. As they were first identified in plants and their best-characterized property was the ability to agglutinate red blood cells, lectins were used to be referred to as phytohaemagglutinins. Subsequently, lectins have been identified and characterized in other forms of life such as animals, bacteria, fungi and viruses (Lis & Sharon, 1998; Loris *et al.*, 1998). Research on lectins received an added impetus on account of the appreciation of the importance of protein-carbohydrate interactions in biology. The importance of carbohydrate as the energy source in food and its crucial role in maintaining the structural integrity of plants has been recognized for long. However, only during the last few decades the important role played by carbohydrates in biological recognition has come to be recognized (Sharon

& Lis, 2004). It turns out that a substantial part of the recognition events, particularly on cell surfaces, are mediated through specific interactions between proteins and diverse carbohydrate structures. Lectins participate in many of these events. They also serve as good models for elucidating such interactions. For this reason and indeed for the insights they provide into the structural diversity of proteins and the strategies for generating ligand specificity, lectins have received considerable attention. Currently, the three dimensional structures of about 200 different lectins derived from widely different sources and several of their carbohydrate complexes accounting for more than 500 independent crystal structures are available (http://www.cermav.cnrs.fr/lectines) . These structures assume widely different folds and have a broad range of sizes and shapes. The only common thread that runs through all of them is the ability of lectins to specifically bind carbohydrates. Plant lectins account for nearly 40% of the lectins of known structures; animal lectins for nearly 25%. The remainder of the structures is of bacterial, viral and fungal lectins. Although those from plants constitute structurally the most thoroughly studied group of lectins, their biological role is not yet fully elucidated. The endogenous functions of lectins from other sources are much better understood (Barondes *et al*., 1988). It also turns out that the two groups of lectins are substantially distinct, allowing them to be treated independently. A long-term program on the structure and interactions of plant lectins and other proteins are being pursued by the authors group. The work reported in this thesis forms an integral part of this program. Therefore, a brief review of different lecins and the available structural information on plant lectins is presented in this introductory chapter.

1.2 Definition and Subdivision of Lectin

With the progress in the purification and characterization of lectins, evidences have accumulated that lectins are a very heterogeneous group of proteins, artificially classified together, solely on the basis of their capability to agglutinate cells (van Damne e*t al*., 1998). However, as soon as it was demonstrated that the agglutinin properties of lectins were based on their specific recognition of and binding to carbohydrates (Watkins & Morgan, 1952), the question arose as to what criteria are to be met by a protein in order for it to be regarded as a lectin. The first definition of lectins was based primarily on the sugar specificity and inhibition of the agglutination reactions. According to this definition, lectins are carbohydrate-binding proteins of non-immune origin which agglutinate cells or precipitate glycoconjugates (Goldstein *et al*., 1980). This definition was too restrictive as it excluded some poorly agglutinating toxins (such as ricin, abrin and modecin), which were known to contain lectin subunits, and so the definition was extended to include these toxins as well (Kocourek & Horejsi, 1983). Also, some lectins contain a second type of binding site that interacts with non-carbohydrate ligands and lectins were re-defined as carbohydrate binding proteins other than antibodies or enzymes (Barondes, 1988).

Recent advances, in the molecular cloning of lectins and lectin related proteins, argue for an update for the definition of lectins on the basis of structural/ functional criteria. Firstly, it has been demonstrated that some plant enzymes are fusion proteins of a carbohydrate domain tandemly arrayed with a catalytic domain. Class I chitinases, for instance, are built up of an N-terminal chitin-binding domain (i.e. evolutionarily related to the

four domains of, for example, wheat germ agglutinin) linked through a hinge region to a catalytic domain (Collinge *et al*., 1993). Similarly, the so-called Type 2 ribosome-inactivating proteins (RIPs, e.g. ricin and abrin) are fusion proteins of an N-terminal toxic A chain, which has the N-glycosidase activity, characteristic of all RIP, and a C-terminal carbohydrate binding B-chain (Hedge & Podder, 1992). Secondly, several carbohydrate-binding proteins possess only one binding site and are therefore unable to precipitate glycoconjugates or agglutinate cells. For instance, the non-agglutinating mannose binding proteins from orchids are very similar to the dimeric mannose-specific lectins from the same species, except that they occur as monomers and are hence incapable of agglutinating cells (van Damme *et al*., 1994). Thirdly, the molecular cloning of lectin genes led to the discovery of so-called lectin-related proteins. Some legumes contain proteins that are evolutionarily and structurally related to the lectins but are devoid of carbohydrate-binding activity, because of alterations in the structure of the binding sites. Well known examples of this group of proteins are the *Phaseolus vulgaris* arcelins, α-amylase inhibitor, etc (Mirkov *et al*., 1994).

These insights into the structural and molecular evolution of lectins and lectin genes confirm the need for a definition that is based on the presence of functionally active carbohydrate domains. As a consequence, the presence of at least one non-catalytic domain, which binds reversibly to a specific carbohydrate, should be the only criterion for a protein to be considered as a lectin. Accordingly, plant lectins can be defined as 'plant proteins possessing at least one non-catalytic domain, which binds reversibly to a specific mono- or oligosaccharide' (Peumans & van Damme, 1995). This new definition is much restrictive, than the previously proposed

definitions which includes a broad range of proteins, with different agglutination and/or glycoconjugate precipitation properties. On the basis of the overall structure of the lectin subunits (or more precisely the primary translation products of the lectin genes) four major types of lectins are distinguished, namely merolectins, hololectins, chimerolectins and superlectins (van Damme *et al.,* 1995). Merolectins are small proteins consisting exclusively of a single carbohydrate-binding domain. Due to their monovalent nature, merolectins cannot precipitate glycoconjugates or agglutinate cells. At present, only a few merolectins have been described. Well known examples are the chitin-binding hevein from the latex of the rubber tree, *Hevea brasiliensis* (van Perjis *et al*., 1991) and the monomeric mannose binding proteins from orchids (van Damme *et al.*, 1994). Hololectins are exclusively composed of carbohydrate-binding domains. However, in contrast to merolectins, they contain two or more carbohydrate domains which are identical or very homologous and bind either the same or structurally similar sugars. Since hololectins have multiple binding sites, they are fully capable of agglutinating cells and/or precipitating glycoconjugates.

Chimerolectins are basically fusion proteins composed of a carbohydrate binding domains tandemly arrayed with an unrelated domain. The latter domain may have a well defined catalytic activity but acts independently of the carbohydrate binding domain. Chimerolectins with multiple carbohydrate binding sites behave as, e.g. Type II Ribosome Inactivating Proteins (RIPs), which possess two carbohydrate-binding sites on each B chain and agglutinate cells. On the contrary, Chimerolectins with a single carbohydrate-binding site behave as merolectins. Class I

plant chitinases, for instance, which has only one chitin-binding domain per molecule cannot precipitate glycoconjugates or agglutinate cells. Superlectins are a special type of chimerolectins. They are fusion proteins built up of two tandemly arrayed carbohydrate-binding domains that are structurally different and recognise structurally unrelated sugars. At present, only a few superlectins has been described; one from tulip bulbs, TXLC-I (Cammue *et al.*, 1986) and UDA from *Urtica dioca* are examples. (Fredrick *et al.*, 2000). The primary translation product of the gene encoding this lectin is composed of an N-terminal mannose-binding domain tandemly arrayed with an unrelated GalNAc-binding domain (Van Damme *et al.*, 1996).

1.3 Occurrence and Distribution

Lectins are usually considered as a very large and heterogeneous group of proteins (Goldstein & Poretz, 1986). Although, there is no doubt indeed that numerous plant species of different taxonomic groupings contain lectins. The total number of well-documented cases is about 400. Assuming that all the close relatives of these plants also contain agglutinins and that some new lectins will be discovered in the future, the expected occurrence of lectins is still limited to a small fraction of the plant kingdom. It can be concluded, therefore, that the occurrence of at least the classical agglutinating lectins in plants is the exception rather than the rule. However, in contrast to the relative scarcity of the agglutinating lectins, chimerolectins belonging to the Class I chitinases seem to be present in almost all plant species (Collinge *et al.*, 1993).

Lectins are widely distributed throughout the plant kingdom where they have been found in a variety of tissues of a large number of different plants (Etzler, 1986). In plants, lectins are particularly localized in seeds. Howard *et al.,* 1972, reported that seed lectins are particularly seen in cotyledons where they appear during the later stages of maturation of the seeds. In addition to cotyledons, in some cases appreciable amounts of lectins have been reported in the embryos and small amounts in the seed coats (Pueppke *et al.,* 1978). Immunolocalization studies have revealed that lectins are primarily found in the protein bodies of the cotyledon cells (Herman & Shannon, 1984). During the early seedling growth, Weber and Neumann (1980) noticed the decrease in lectin concentration as the cotyledons are resorbed. A short survey of the occurrence and concentration of lectins in seeds as well as in different types of vegetative tissues reveals striking differences in the location and relative abundance of the individual lectins (Etzler, 1986). Usually, seed lectins are confined to cotyledons (e.g. legumes) or endosperm (e.g. castor bean). Normally lectins account for up to 5% of the total seed proteins. Sometimes, they become predominant protein in the seed representing 50% of the total seed protein (e.g. *Phaseolus* species). The non-seed lectins are found in all kinds of vegetative tissues such as leaves, stem, bark, bulb, tubers, corns, rhizomes, roots, fruits, flowers, ovaries, phloem sap and even in nectar (Peumans & van Damme, 1995) and are only minor, quantitatively unimportant proteins. Non-seed lectins may occur in different tissues of the same plant. The snowdrop and daffodil lectins, for instance, have been found in all vegetative tissues, although the lectin is most abundant in the bulbs (van Damme & Peumans, 1990). Similarly, the potatoe lectin occurs in tubers, stems, leaves and fruits (Kilpatrick, 1980). There are exceptions also. The ground elder berry lectin is

confined to the rhizome only (Peumans *et al.*, 1985). In the case of tulip bulbs, lectins are present in large quantities in the bulb but are almost undetectable in stem and leaves (van Damme & Peumans, 1989). Some legume lectins are found in seeds as well as in bark tissues. A thorough examination of the genes coding for these lectins revealed that the seed and bark lectins are encoded by different, though highly hommologus, genes (van Damme *et al.*, 1995).

1.4 Physicochemical Properties

1.4.1 Composition

There are no structural features common to all lectins. Many of these proteins are relatively rich in aspartic acid, serine and threonine, which comprise as much as 30% of their amino acid content and are low in sulfur-containing amino acids. Such a pattern of amino acids is characteristic of plant proteins. In contrast, lectins such as those from wheat germ, potato and pokeweed are rich in cysteine with 20, 11.5 and 18% of the total amino acid residues respectively, most or all of which are in the form of cysteine. The high content of disulfide bonds in wheat germ agglutinin endows the protein with stability to heat (Aub *et al.*, 1963; Burger & Goldberg, 1967), to proteolytic enzymes and to denaturing agents such as detergents, urea, alkali and acids (Nagata & Burger, 1972; Rice & Etzler, 1974). The potato and *Datura stramonium* lectins are rich in hydroxyproline (Lamport, 1969).

A few lectins, such as concanavalin A, wheat germ and peanut agglutinins are devoid of covalently bound sugars. Most lectins, however, are glycoproteins with carbohydrate contents that can be as high as 50%,

e.g., potato lectin. The table shown below (Table No 1.1) is on the sugar contents of certain important glycoprotein lectins. The sugar constituents in animal glycoproteins are the same as those found in other plant glycoproteins, with the exception of L-arabinose.

Table 1.1 Well-characterised glycoprotein lectins

Lectin	Man	Gal	L-Fuc	L-Ara	Glc NAc	Xyl	C-P linkage	Reference
Bandeiraea simplicifolia	5.8		1		2.6	1		Lescar *et al*., 2007
Datura stramonium				28	4.5			Kilpatrick, 1978
Phaseolus lunatus	3.2	3.7	0.5		1.3		Glc Nac- Asn	Mach *et al*., 1991
Solanum tuberosum		3		47			Ara-Hyp, Gal-ser	Matsumoto *et al*., 1983
Phaseolus vulgaris	7.3				2.8		GlcNAc- Asn	Ohtani *et al*., 1980
Glycine max	4.5				1.2		GlcNAc-Asn	Lis *et al*., 1973
Wistaria floribunda	0.77	1.63			0.65		Glc NAc-Asn	Kueokawa *et al*., 1976

The MW of lectins in plants ranges from 36,000 Da for wheat germ agglutinin (Nagata & Burger, 1972; Rice & Etzler, 1974) to 265,000 for lima bean lectin (Galbraith & Goldstein, 1970). The lower limit of MW of animal lectins is found to be 14 kDa (Lis & Sharon, 1998). Some lectins exhibit a pronounced tendency to aggregate. Thus, the MW of Con A at pH below 6 is 51,000 Da and at physiological pH it is 1, 02,000 Da (Kalb & Lustig, 1968; Mc Cubbin & Kay, 1971; Wang *et al.,* 1971). Upon storage at room temperature, soybean agglutinin and peanut agglutinin also possibly

undergo irreversible self-association to high molecular weight aggregates (Lotan *et al.,* 1975). The subunits are identical in most lectins. But lectins comprising of non-identical subunits are known as seen in soyabean agglutinin (Lotan *et al*., 1975) and the lectin from *D. biflorus* (Carter & Etzler, 1975) which are tetramers, consisting of two types of subunits (Wright et al., 1996). A different type of subunit heterogeneity was first demonstrated in Con A (Abe *et al*, 1971; Wang *et al.,* 1971). The anti-B lectin from *Bandeiraea simplicifolia* consists of a family of five closely related proteins, each of which is a tetramer of one or two types of subunits. One of the subunits is specific for a-N-acetyl galactosamine, whereas the specificity of the other is confined to a-galactose (Goldstein *et al.,* 1978). The structure of *Bandeiraea simplicifolia* isolectins is analogous to that of PHA isolectins. They have five tetrameric proteins comprising of varying proportions of two classes of subunits (Miller *et al*, 1973; Rasanen *et al*., 1973; Leavitt *et al*., 1977). These subunits show difference in properties. It is assumed that it is due to their difference in the primary structure of subunits (Miller *et al*., 1973).

1.4.2 Metal Ion Requirements

With a few exceptions, all lectins examined contain metal ions and in some cases evidence has been presented for the requirement of Mn^{2+} or Ca^{2+} (Emmerich *et al*., 1994) for activity (Table 1.2). Treatment with ethylene-diamine tetra acetic acid (EDTA) at neutral pH did not remove the metal ions from Con A (Doyle *et al*., 1984), soybean agglutinin (Jaffe *et al*., 1974) or lima bean lectin (Galbraith & Goldstein, 1970). Reversible removal of metal ions can be achieved under acidic conditions. The Mn^{2+} in lectins can be replaced by a variety of transition-metal ions without loss of

biological activity as demonstrated for Con A (Agrawal & Goldstein, 1968; Shoham *et al.,* 1973). Ca^{2+} in Con A could be replaced by $Cd^{2+,}$ but not by Ba^{2+} (Shoham *et al*., 1973). The metal ions confer a high degree of structural stability to Con A, protecting the lectin against heat inactivation (Doyle *et al.,* 1975) and hydrolysis by proteolytic enzymes (Thomasson & Doyle, 1975). Ni^{2+} alone protects Con A against proteolysis at pH 7.0 but not at pH 8.2. Some lectins require metal ions for the saccharide-binding activity (Sumner & Howell, 1936). Extensive studies by NMR have revealed a complicated set of interlocking equilibrium involving the apoprotein and various complexes with metal ions and the saccharides (Brewer *et a*l., 1983).

Table 1.2 Metal content and metal requirements for activity of lectins

Lectin	**Metal content (atoms per mole)**				**Reference**
	Mn^{2+}	**Ca^{2+}**	**Zn^{2+}**	**Metal**	
Canavalia ensiformis	4	4		Mn^{2+}	Magnuson *et a*l., 1983
Datura stramonium	<0.2		<0.2		Kilpatrick, 1978
Dolichos biflorus	1.6	5.4	2.0		Etzler *et al.,* 1970
Euonymus europeus	-	8.0	0.7		Petryniak *et al*., 1977
Lens culinaris	0.64	3.8		Mn^{2+}	Westbrook *et al*., 1984
Phaseolus lunatus	1.0	4.0		Mn^{2+}	Mach *et al*.,1991
Marasrous oreades	-		0.7		Winter *et al*., 2002
Ononis hircina	1.0		1.0		Horejsr *et al*., 1978
Pisum sativum	1	2.5		Ca^{2+}	Reeke *et al.,* 1986
Phaseolus coccineus	0.15	4.8	1.0		Perez Campos *et al.*,1997
Ricinus communis	<0.1	<0.1	<0.1		Mandal *et al.*,1989
Sarothamnus scoparius	1.5		0.8		Poretz *et al.*,1992
Glycine max	1-1.7	3.5-4.1	0.28	Mn^{2+}	Lis and Sharon, 1973
Vicia cracca	0.9		2.4		Sitohy *et al.*,2007
Phaseolus vulgaris	0.24	6.2		Mn^{2+}	Andrews , 1974
Ulex europeus I	0.42	2.0	0.82		Yanamoto *et al.,* 1992
Bandeiraea simplicifolia I	1.2	2		Ca^{2+}	Lescar ***et al.,*** 2002

1.5 Biological Role

Lectins are present abundantly in many plants. Despite this abundance, their precise biological roles in the plants to which they belong, are not well understood. The available evidences suggest two main roles for them.

- **Mediation of symbiotic relationship between nitrogen fixing microorganisms, primarily, rhizobia and leguminous plants.**

Lectins localized at the root hairs are the entry sites for rhizobia. The lectins then aggregate the rhizobia in the root nodules and make them immobile (Hamblin & Kent, 1973; Bohlool & Schmidt, 1974; Diaz *et al.,* 1989; Brewin & Kardailsky, 1997; Hirsch, 1995). Type specificity of host-parasite interactions between leguminous plants and particular strains of rhizobia infecting them is determined by lectins. The expression of the pea lectin gene in white clover roots enabled them to be nodulated by a rhizobium strain specific for the pea plant (van Eijsden *et al.,* 1995).

- **Protection of plants from predatory animals and phytopathogens.**

Abrin, a type-II ribosome-inactivating protein (RIP), was the first lectin to be recognized as a defence protein (Peumans & van Damme, 1995). Soon afterwards ricin also came to be recognized as a defence protein (Olsnes, 2004). Type-II RIPs which belong to the plant lectin family with β-trefoil fold (see later) are known to be toxic to animals and insects (Hartley & Lord, 2004; Stirpe, 2004). Lectins from *Phaselous vulgaris* (PHA), *Robinia pseudocacia* and *Sambuscus nigra* have been reported to be toxic to higher animals (Peumans & van Damme, 1995). Lectins from many plants, when ingested by animals, have resulted in toxic effects (Lis & Sharon, 1998), fungal growth in *Trichoderma viride* is inhibited by wheat germ aggiutinin

(WGA) (Mirelman *el al.,* 1975). Brambl and Gade (1985) have shown that eleven purified lectins, representing a wide spectrum of sugar specificity, inhibited the growth of fungal species *Neurospora crassa*, *Aspergilius amsteldomi* and *Botryodiplodia theobromae*. Known antifungal lectins include those which bind chitin (Peumans & van Damme, 1995) Hirsch *et al.,* 1995; Eijsden *et al.,* 1995; Kijne, 1997; Selitrennikoff, 2001). The anti-insect activity of many plants has been attributed to the presence of lectins in them. For example PHA (Chrispeels & Raikhel, 1991; Etzler, 1992) pea nut agglutinin (PNA), WGA, *Maclura pomifera* agglutinin (MPA) and lectins from potato, thorn apple and osage orange show anti-insect activity against cowpea weevil. WGA and *Bauhinea purpurea* agglutinm are toxic to *Ostrinia nubilalis* larvae. Snow drop and garlic lectin show toxic effects on cowpea weevil and tobacco hornworm (Hilder *et a*l., 1995; Peumans & van Damme, 1995).

1.6 Applications

Although, as mentioned earlier, the biological role of plant lectins in the plants to which they belong are not yet fully elucidated, their ability to specifically bind carbohydrate structures have been made use of extensively in research and bio medical applications. These applications primarily employ precipitations and aggregations.

Blood typing: Blood typing is among the oldest applications of plant lectins. Lectins from *Lotus tetragonolobus* and *Ulex europaeus,* both specific for fucose, are employed for identifying blood type O cells. *Dolichos biflorus* agglutinin is used to distinguish between Al and A2 subgroups whereas *Vicia graminea* lectin is able to select out blood type N from a mixture of M and N type cells. Peanut lectin and *Vicia villosa* aggregate T and Tn cell types,

respectively. An N-Acetylgalactosamine specific lectin from lima bean agglutinates type A red cells (Slifkin & Doyle, 1990; Sharon & Lis, 2004; Sharon, 2005).

Analysis and purification of glycoconjugates: The ability to bind specific carbohydrates makes lectins an obvious tool for differentiating glycoconjugate on the basis of sugar present in them. Glycoproteins containing rnannosc, galactosc and N-acetylneuraminic acid are purified in laboratories using concanavalin A (Con A), PNA and wheat germ agglutinin, respectively (Liener *et al.,* 1986). Jacalin, in addition to its ability to selectively bind IgAl subclass, binds a number of human plasma glycoproteins. They are IgD, C l-inhibitor, C4-binding gpI20, hemopexin, plasminogen, α l-antitrypsin, α-2-macroglobulin, 8S α3-glycoprotein and α2-HSG (Hagiwara *et al.,* 1988; Kondoh *et al.,* 1986; Aucouturier *et al.,* 1987; Hiemstra *et al.,* 1987, Hortin & Trimpe, 1990; Tonevitsky *et al.,* 1990).

Histochemical and cytochemical probes: During the process of development, cell surface carbohydrates undergo considerable changes. Lectins are used as markers to characterize these changes (Gabius, 1991). The ability to agglutinate specific cell types is used for cell separation. Peanut agglutinin (PNA), soybean agglutinin (SBA) and jacalin are particularly used in this regard (Sharon & Lis, 1989a). The ability of PNA to distinguish between mature and immature thymocytes has led to it being used in bone marrow transplantations (Reisner, 1987). The exclusive specificity of PNA for the Thomson-Friedenreich antigen (T-antigen) is exploited widely for monitoring its differential expression for both the prognosis and diagnosis of malignancies (O'Keefe & Ashman, 1982; Zabel *et al.,* 1983; Ching & Rhodes, 1989; Zebda *et al.,* 1994). Jacalin has been used as a histochemical

reagent to study tissue-binding properties in benign and malignant lesions of the breast and thyroid (Remani *et al.,* 1989; Vijayakumar *et al.,* 1992). The number of cells with intense jacalin binding increased with various stages of tumour progression. Jacalin has been used to distinguish malignant cells from benign reactive cells in serous effusions (Sujathan *et al*., 1996). The intensity of jacalin binding to cells of different types of oral carcinomas is high, particularly in the membrane of differentiated cells (Pillai *et al*., 1996). Lectins derivatised with gold particles, fluorescent dyes or enzymes are employed as histochemical and cytochemical reagents for detection of glycoconjugates in tissue sections, on cells and subcellular organelles, and in investigations of intracellular pathways of protein glycosylation (Lis & Sharon, 1998; Bies *et al*., 2004).

Mitogenic stimulation of lymphocytes: Certain lectins are potent mitogens, activating lymphocytes and inducing them to divide; phytohemagglutinin (PHA) and Con A, for example, stimulate T lymphocytes, while pokeweed mitogen (PWM) stimulates both T and B cells (Di Sabato *et al.,* 1987; Ashraf & Khan, 2003). PNA could agglutinate lymphocytes from rat, mouse and human only after their treatment with neuramindase. However, it stimulates neuraminidase-treated cells from rat and human (Novogrodsky *et al.,* 1975). Mitogenic simulation by lectins provide easy and simple means to assess the immunocompetence of patients suffering from a diversity of diseases including AIDS and to monitor the effects of various immunosuppressive and immunotherapeutic manipulations. In particular, jacalin has been applied in AIDS research since it is the only lectin known to be selectively mitogenic for human CD4+ T cells (Pineau *et al.,* 1989, 1990; Corbeau *et al.,* 1994, 1995; Favero *et al.,* 1993; Lafont *et al.,* 1993, 1994, 1996). Since CD4+ T lymphocytes act as a receptor of HIV 1, Jacalin has been used to investigate

the proliferation of PBMC in HIV-1 infected patients. *Canavalia brasiliensis* (ConBr), *Pisum arvens* (PAA+) and *Artocarpus integrifolia* (KM+) lectins are used as immunostimulatory molecules in vaccination against *Leishmania amazonensis* infection. They induce IFγ, enhance the expression of MHC II, CD80, and CD86 (CD80 and CD86 are members of the Ig gene superfamily) and reduce the level of parasite (Teixeira *et al.,* 2006).

Mapping of neuronal pathways: Another area of lectin application is neuroanatomy. Lectins are increasingly being used as tracers for mapping neuronal connections, WGA being the first to be used in this regard (Gerfen & Sawchenko, 1985). Lectin-HRP (Horse Radish Peroxidase) conjugate is readily taken up by neurons and transported along the axon thus helping to trace the central neuronal pathways. The same strategy when employed with a cytotoxic lectin (e.g. ricin) functions as suicide transporter offering a new means to tackle neurobiological disorders. PNA and WGA are used to label the interphotoreceptor matrix in cryosections of retinal tissue. PNA readily labeled the cone-associated matrix with faint binding to the rod-associated matrix, WGA labeled both the rod- and cone-associated matrices (Kristina *et al.,* 1991).

Lectin mediated drug targeting and delivery: Based on the fact that oligosaccharides encode biological information, the biorecognition between lectinsed drug delivery systems and glycosylated structures in the intestine could be exploited for improved peroral therapy (Gabor *et al.,* 2004). Basic research revealed that lectins such as WGA, Con A, PNA and jacalin can mediate mucocoadhesion, cytoadhesion and cytoinvasion of drugs (Yi et *al.,* 2001). Entering the vesicular pathway by receptor mediated endocytosis; part of the conjugated drug is accumulated within the lysosomes. Additionally,

part of the drug is supposed to be transported across the epithelium. As exemplified by lectin-grafted prodrug and carrier systems, this strategy is expected to improve absorption and probably bioavailability of poorly absorbable drugs, peptides and proteins as well as therapeutic DNA (Gabor *et al.*, 2004). Conjugation of lectins with suitable drugs or conjugation of drugs to lectins enhances the drug delivery to the epithelial cells. Lectins from mistletoe (ML I-III), stinging nettle (UDA), tomato (TL, LEA), and WGA are used to target human intestinal eptithelial Caco-2 cells. In this context, WGA, TL and Con A are used to deliver the drug formulation to the intestinal cells. UEA-I, DBA, WGA and GS-I-B4 are targeted to M cells in the nasal cavity. With its higher affinity towards cerebral endothelium cells, WGA is used to cross the blood-brain barrier, without disturbing the brain function (Bies *et al.*, 2004; Gao *et al.*, 2006).

1.7 Structural Aspects

1.7.1 The background

Starting with Con A in the 70s (Hardman & Ainsworth, 1972; Reeke *et al.* 1975), the structure determination of lectins using X-ray crystallography gathered momentum over the years. Structural work in the area has been particularly vigorous during the last couple of decades. The crystal structures of many lectins from plants, animals, bacteria, viruses and fungi are currently available (http://www.cermav.cnrs.fr/lectines). The information on the structures and interactions contained in them has been extensively reviewed (Drickamer, 1997; Lis & Sharon, 1998; Loris *et al.*, 1998; Weis *et al.*, 1996, Vijayan & Chandra, 1999; Chandra *et al.*, 2001). The structural information currently available on lectins is so extensive that it is rather difficult to review it in a

comparatively short introductory chapter like, as attempted here. In any case, this dissertation is exclusively concerned with a plant lectin. Furthermore, as indicated earlier, plant lectins to a great extent stand apart from other lectins as a distinct structural group. Therefore, the discussion here is confined to plant lectins. As can be seen from Table 1.1, plant lectins themselves constitute a large group. These lectins can be structurally classified into the following sub-groups: (i) legume (ii) cereal (iii) β-trefoil (iv) β-prism II fold and (v) β-prism I fold. As mentioned earlier, the first plant lectin fold to be identified was jelly roll as in the legume lectin, Con A. The hevein domain in cereal lectins was next identified through the structural analysis of WGA (Wright *et al.*, 1989). The β trefoil fold as in ricin was the next plant lectin fold to be recognized (Rutenber *et al.*, 1991). The β-prism II fold was first seen in snowdrop lectin (Hester *et al.*, 1995). The last known plant lectin fold, the β-prism I fold, was identified in 1996 (Sankaranarayanan *et al.*, 1996).

Table 1.3 Classification **of plant lectins structures based on source and fold. Data are from the '3Dlectindatabase' internet site by Bettler, E., Loris, R. and Imberty, A. (http://www.cermav.cnrs.fr/lectines)**

Family	Source	Common/acronym	Reference
Jelly roll lectins (Legume lectins)	*Canavalia brasiliens*	C-brasiliens	Sanz Aparicio *et al.*, 1997
	Canavalia ensiformis	Con A	Bouckaert *et al.*, 1995
	Canavalia maritina	ConM	Gadelha *et al.*, 2005
	Cratylia mollis	Cramoll	De Souza *etal.*, 2003
	Dioclea grandiflora	DGL	Rozwarski *et al.*, 1998
	Dioclea guianensis	Dguia	Wah *et al.*, 2001
	Dolichos biflorus	DBL & DB58	Hamelryck *et al.*, 1999
	Dolichos lablab	FRIL	Hamelryck *et al.*, 2000
	Erythrina orallodendron	ECorL	Elgavish *et al.*, 1998
	Erythrina crystalgalli	ECL	Svensson, 2002
	Vicia faba	Favin	Reeke *et al.*, 1986
	Griffonia simplicifolia	GS1,GS4 & GS-I-B4	Lescar *et al.*, 2002
	Lnthyrus ochms	LOL-1 & LOL-2	Bourne *et al.*, 1990
	Lens culinaris	Lentil lectin (LCL)	Loris *et al.*, 1994
	Phaseolus lunatus	Lima bean lectin(LBL)	Imberty *et al.*, 1998
	Maackia amurensis	MAL	Imberty *et al.*, 1998
	Pisum sativum	Pea lectin (PSL)	Ruzeinikov *et al.*, 1998
	Arachis hypogaea	Peanut lectin (PNA)	Ravishankar *et al.*, 1998
	Phaseolus vulgaris	PHA	Hamelryck *et al.*, 1996
	Pterocarpus angolensis	PAL	Loris *et al.*, 2003
	Robinia pseudoacacia	Bark lectin I	Robijns *et al.*, 2001
	Glycine max	Soybean (SBA)	Oslen *et al.*, 1997
	Ulex europaeus	UEA-I & UEA-II	Gohier *et al.*, 1996
	Vicia villosa	VVL-B4	Babino *et al.*, 2003
	Psophocarpus tetragonolobus	Winged bean (WBA) I & II	Manoj *et al.*, 1999

Hevein domain lectins (Cereal lectins)	*Amaranthus caudatus* *Triticum vulgaris* *Phytolacca americana* *Hevea brasiliensis* *Urtica dioica*	Ac-AMP2 Wheatgermagglutinin (WGA) Isolectin I, II & III Pokeweed lectin (PL-D1 & P1-D2) Hevein UDA	Martins *et al.,* 1996 Wright , 1990 Fujii t *et al.,* 2004 Gidrol et al., 1994 Harata *et al.,* 2000
β–prism II fold lectins	*Gslanthus nivglis* *Scilla campanulata* *Narcissus pseudonarcissu* *Alliurn sativum*	Snowdrop lectin Bluebell bulb lectin (SCA-FET & SCA-MAN) Daffodil lectin Garlic lectin	Hester *et al.,* 1995 Wright *et al.,* 1996 Sauerborn *et al.,* 1999 Chandra *et al.,* 1999
β-prism I fold lectins	*Artocarpus integrifolia* *Artocarpus hirsuta* *Musa acuminata* *Musa paradisiaca* *Calystegia sepium* *Helianthustuberosus* *Artocarpus integrifolia* *Maclura pomifera* *Morus nigra* *Parkia platycephala*	Artocarpin *Artocarpus hirsuta* lectin(AHL) Banana lectin Banana lectin Calsepa Heltuba Jacalin MPA MornigaM Parkia lectin	Pratap *et al.,* 2002 Rao *et a.,l* 2004 Meagher *et al.,* 2005 Singh *et al.,* 2005 Bourne *et al.,* 2004 Bourne *et al.,* 1999 Sankaranarayanan *et al.,* 1996 Lee *et al.,* 1998 Robijns *et al.,* 2005 Delsol *et al.,* 2005

1.7.2 Jelly roll lectins (Legume lectins)

Until a decade or so ago, legume lectins were almost considered synonymous with lectins. Although the situation has changed with the structure analysis of many lectins from other sources, those from legumes still remain the most thoroughly studied group of lectins. The first lectin to

be X-ray analysed, Con A, belongs to this family (Figure 1.1) (Hardman & Anisworth, 1972; Reeke *et al.,* 1975). Con A is a Glc/man specific lectin. Subsequently, the structures of a number of legume lectins with and without bound-carbohydrate have been determined. These include Glc/Man specific lectins such as favin (Reeke & Becker, 1986) pea lectin (PSL; Einspahr *et al*., 1988) *Lathyrus ochrus* isolectin-I (LOLI; Bourne *et al.,* 1990) lentil lectin (LenL; Loris *et al.,* 1994) *Lathyrus ochrus* isolectin II (LOLII; Bourne *et al.,* 1994) *Canavalia brasiliensis* lectin (AZD; Sanz-Aparicio *et al.,* 1997*) Dioclea grandiflora* lectin (DGL; Rozwaski *et al.,* 1998) *Cratylia mollis* (Cramoll; de Souza *et al.,* 2003*) Canavalia maritime* (ConM; Gadelha *et al,* 2005). Gal/GalNAc specific lectins analysed include *Erythrina corrallodendron* lectin (ECorL; Shaanan *et al.,* 1991; Elgavish & Shaanan, 1998) PNA (Banerjee *et al*., 1994, 1996) SBA (Dessen *et al.,* 1995) *Vicia villosa* isolectin B4 (Babino *et al.,* 2003) winged bean basic agglutinin (WBAI; Prabu *et al.,* 1999) winged bean acidic agglutinin (WBAII; Manoj *et al.,* 2000) *Dolichos biflorus* seed lectin (DBL; Hamelryck *et al.,* 1999) *Dolichos biflorus* stem and leaf lectins (DB58; Hamelryck *et al.,*1999). Also reported as the fucose or promiscuous specific lectins I and II from *Ulex europeaus* (Audette *et al.,* 2000; Loris et al., 2000), and complex carbohydrate-binding lectins such as lectin IV from *Griffonia Simplicifolia* (GS4; Delbaere *et al.,* 1993), PHA-L (Hamelryck *et al*., 1996 a). All the legume lectins mentioned above are either dimeric or tetrameric, with identical or almost identical subunits of 25-30 kDa. Each subunit has a single carbohydrate-binding site. They exhibit exquisite specificity for di, tri and tetrasaccharides, with association constants up to 1000-fold higher as compared to the monosaccharides. A few of them are specific only to oligosaccharides. Some of them are glycosylated, with the N-linked sugars accounting for 10% of their

molecular weight. They bind to sugars only in the pyranose form in the D configuration, except in the case of fucose. The substituents at C3 and C4 positions determine sugar specificity for monosaccharides (Liener *et al.*, 1986). Some lectins exhibit differential specificities for the α/β-anomers. Each subunit contains a Ca^{2+} ion and a transition metal ion, usually Mn^{2+}. The metal ions are approximately 4.5 Å apart and bridged by two aspartate residues. Both metal ions have four protein ligands and two water ligands that are conserved in all legume lectins. Although these ions do not themselves interact with the saccharide, they are essential for sugar-binding. They facilitate the orientation of the carbohydrate-binding residues for optimum interactions. Large changes have been observed in the structure of Con A upon demetallization, resulting in a loss of carbohydrate-binding ability (Bouckaert *et al.*, 1995). However, GS-I-B4 when demetallized still binds carbohydrate perhaps with less affinity as evidenced by the loss of two of the seven lectin-sugar hydrogen bonds present in the original structure. Some legume lectins possess a hydrophobic binding site that binds adenine and adenine-derived plant hormones, cytokinins (Roberts & Goldstein, 1983), which does not interfere with carbohydrate-binding (Gegg *et al.*, 1992). The legume lectin structural family also contains two proteins with no carbohydrate recognition activity, the α-amylase inhibitor (α-AI) and arcelin, a seed defence protein, both from *Phaseolus vulgaris* (Hamelryck *et al.*, 1996). Expectedly, the metals are absent in them. Most of the residues involved in metal-binding and carbohydrate recognition in other legume lectins are not conserved in α-AI and arcelin (Bompard-Giles *et al*, 1996). Peptide mimicry of carbohydrate-binding in lectins began to be recognized as an interesting line of investigation in the early 90s (Scott *et al.*, 1992; Oldenburg *et al.*, 1992). The structural biological content for this approach was provided by a series of

studies on Con A-peptide interactions by Salunke and others (Kaur *et al.*, 1997; Jain *et al,* 2000; Zhang *et al.*, 2001). The above group has investigated a complex between PNA and phorphyrin also (Goel *et al.*, 2005).

1.7.2.1 Primary structure

The subunits of legume lectins are commonly made up of a single polypeptide chain of about 250 amino acids, although a few of them (e.g. PSL and LenL) are fragmented into a heavy and a light chain. The primary structure exhibits homology by circular permutation. The N-termini of the heavy chain and single polypeptide chain of PNA coincide, both aligning with residue number 123 of Con A. This alignment continues in the same direction as in Con A pass its C-terminus and continuing along the N-terminus. Insertions and deletions do occur during alignment. This unusual sequence homology arises due to post-translational processing of Con A (Carrington *et al.*, 1985). Recently, this type of sequence homology has been observed in other classes of proteins as well (Lindqvist & Schneider, 1997). In the case of legume lectins the sequence homology observed by considering them, on account of post-translational processing. The N-and C-termini of Con A appear to be located differently from those in other legume lectins.

1. 7.2.2 Secondary and tertiary structures

The legume lectin subunit exhibits the lectin fold, which is a subset of the jelly roll fold. This fold has also been found in proteins other than legume lectins with hardly any detectable similarity with the lectins in their amino acid sequences (Srinivasan *et al.*, 1996). The secondary and tertiary folds are largely conserved and consist of a flat six stranded anti-parallel β-sheet at the back of the subunit, and a seven-stranded curved anti-parallel β-sheet at the front and a small five-stranded β-sheet which plays an important role in

holding the other two sheets together (Hardman & Ainsworth, 1972; Banerjee *et al.,* 1996; Chandra *et al.,* 2001) (Figure 1.1).

Figure 1.1: Stereo view of a subunit of PNA with bound-lactose (PDB: 2PEL).

Loops and β-bends that connect the strands of the sheets constitute about 50% of the secondary structures. α-helices are not seen except in DBL and DB58. The main hydrophobic core is located between the front and the back sheets. The carbohydrate-binding region and the metal-binding region are located mostly in the loops contiguous to the front sheet.

1.7.2.3 Quaternary structure

Although, all of them have similar tertiary structure, legume lectins exhibit a variety of quaternary structures. The oligomerization in all lectins of this family

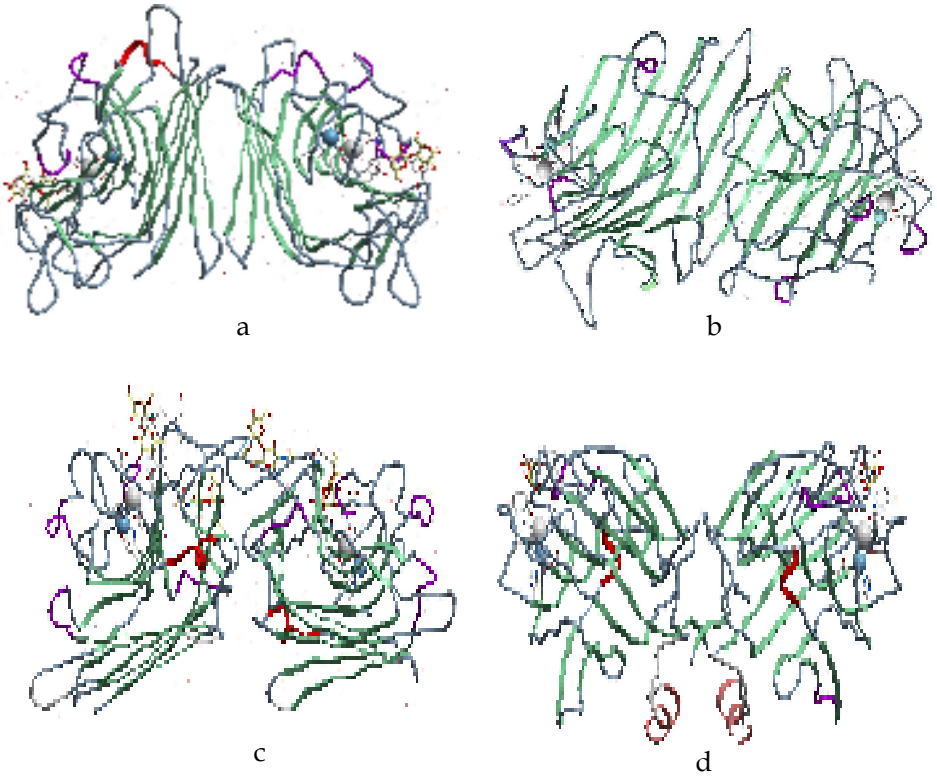

Figure 1.2: The commonly observed modes of dimerization in legume lectins: a) Con A dimer, b) ECorL dimer, c) DBL dimer, d) GS4 dimer.

involves the 6-stranded back side β-sheet. In Con A (Hardman & Ainsworth, 1972; Reeke *et al.*, 1975), the formation of the two-fold symmetric dimer involves a side-by-side arrangement of the two monomers, such that the two back side β-sheets form a contiguous 12-stranded β-sheet (Figure 1.2 (a). This is referred to as the 'canonical' mode of legume lectin dimerization. Further contacts, arising through side-chain-side-chain, side-chain-main-chain and water bridges are also observed. This canonical mode of dimerization has been observed in AZD, DGL, PSL, favin, LOL I, LOL II, LenL, SBA, UEA-I and UEA-II. The first non-canonical mode of dimerization was observed in dimeric GS4 (Delbaere *et al.*, 1990), in which there is a back-to-back arrangement of the two β sheets (Figure: 1.2 d). This difference in dimerization was suggested to be caused to avoid the burial of a glutarmic acid residue (Delbaere *et al.*, 1993). The structure analysis of G S4 was closely followed by that of ECorL, which again exhibited a non-canonical mode of dimerization (Shaanan *et al.*, 1991) (Figure: 1.2 (b). It was suggested that canonical dimerization in ECorL is prevented through steric interactions of the covalently linked carbohydrates. Peanut lectin, a tetramer, is not glycosylated (Banerjee *et al.*, 1994, 1996). Yet, the main (dimer) interface in it is non-canonical, suggesting that the variability in quaternary association in legume lectins is not necessarily caused by interactions involving covalently linked sugar. The basic lectin from winged beans (WBAI) has 63% sequence identity with ECorL. However, unlike ECorL, the glycosylation sites in WBAI are far away from the region that would constitute the interface in the canonical dimer and hence do not prevent the formation of such a dimer. Yet, WBAI forms ECorL type of dimers, demonstrating that the mode of dimerization is

primarily dictated by factors intrinsic to the protein itself (Prabu *et al*., 1999). The same conclusion was drawn from the structure analysis of the WBAII also (Manoj *et al*., 2000). The above observations show that legume lectins form an interesting family of proteins, in which small alterations in essentially the same tertiary structure, consequent to sequence changes, lead to large variations in their quaternary structures. These variations have been examined by (Prabu *et al*., 1999) and rationalized based on hydrophobic surface area buried upon oligomerization, interaction energy and shape complementarity. All three indices favour the observed mode of quaternary association in their proteins. They also showed that this variability can be described in terms of simple geometrical parameters. The dimers of DBL and DB58, display yet another mode of dimerization featuring a α-helix sandwitched between two subunits (Figure 1. 2. (c)). The tetrameric structures in this class of proteins also show interesting variations. The quaternary structure of the tetrameric lectins Con A, AZD, DGL, SBA, PHA-L and UEA-II can be described in terms of back-to-back arrangements of side-by-side dimers. Among them, lectins exemplified by Con A have an arrangement schematically illustrated in Figure1.*3.(a)*. SBAprovides an example of a different back-to-back arrangement of side-by-side dimers PNA, however, has an open quaternary structure (Banerjee et *al*., 1994, 1996) involving the association of two back-to-back dimers (Figure 1.3. (b). The recently determined structure of GS-I-B_4 has a dimeric arrangement like in PNA, but the two dimers associate symmetrically to form a closed tetramer. The tetrameric association in DBL is similar to that of PHA-L and SBA. Thus, at least four different kinds of dimeric associations and four kinds of tetrameric associations are seen in legume lectins.

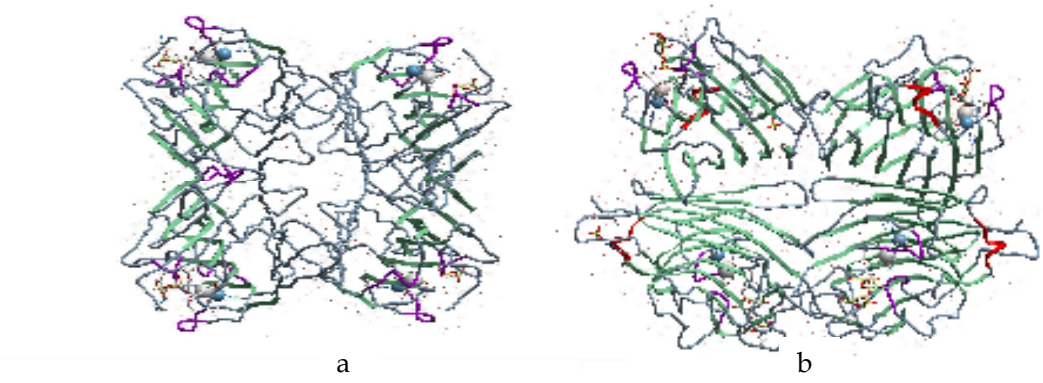

Figure 1.3: Tetrameric association of (a) Con A (b) PNA

Many plant lectins, covering a wide range of carbohydrate specificities, have been shown to specifically precipitate glycoprotins and branched multivalent oligosaccharides (Bhattacharyya & Brewer, 1986; Oben *et al.*, 1991; Sacchettini et al., 2001; Dam *et al.*, 2000), often leading to the formation of homogenous cross-linked lattices. The formation of homogeneous lattices can be in part explained by the multimeric nature of lectins. The different spacings between the carbohydrate-sites of different lectins combined with the specific distances between the different epitopes on multivalent ligands are probably the reason for this type of specificity.

1.7. 2. 4 Lectin-sugar interactions

Detailed crystallographic studies on carbohydrate complexes and thermodynamic measurements have provided a wealth of information on the

structural basis of legume leetin sugar interactions. The affinity of lectins for monosaccharides is usually weak (with Ka in the mM range); yet they are highly selective (Sharon & Lis, 1989a; Kini, 1995; Weis & Drickamer, 1996). In general, lectins that bind to Galactose (Gal) do not recognize Glucose/Mannose (Glc/Man) and vice versa. Within a subgroup, they exhibit different affinities for different derivatives. Lectins that bind Gal usually bind GalNAc, generally with a stronger affinity, but there are exceptions. PNA is one such exception. GS4 and PHA-L are two legume lectins that do not bind to monosaccharides at all. The atomic features of carbohydrate-binding have been studied in considerable detail and attempts made to identify common recognition principles (Lis & Sharon, 1998; Loris *et al*., 1998; Bouckaert *et al*., 1999). The binding sites are in the form of shallow depressions on the surface. One or two edge(s) face(s) of the carbohydrate ligand is/are bound to the protein. From a comparison of the saccharide-bound and the saccharide-free forms, it appears that the binding site is substantially preformed, with water molecules replacing the key carbohydrate hydroxyls in the free form (Elgavish & Shaanan, 1998; Delbaere *et al*., 1990, 1993; Ravishankar *et al*., 1998). Carbohydrate-binding is facilitated by a network of hydrogen bonds between sugar hydroxyls on the one hand and carboxylate, NH and side chain groups in the protein on the other. van derWaals interactions, though week, are found in large numbers and might contribute significantly to binding. Stacking by aromatic residues (for Gal specific lectins) is also observed.

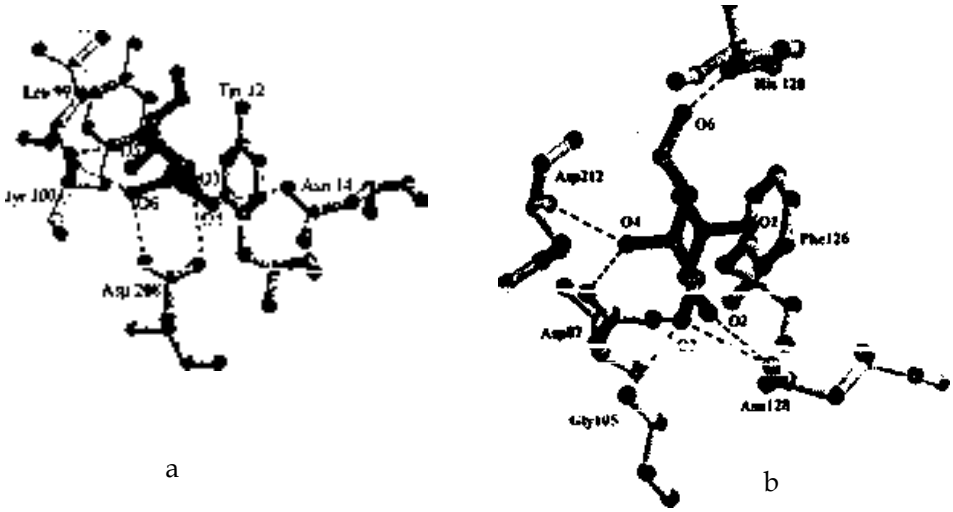

Figure l.4: (a) The mannose-binding site in Con A (PDB: 5CNA) and (b) galactose-binding site in WBAI (PDB: 2DUO).

Water mediated interactions also seem to play an important role in sugar-binding to lectins. Water has the ideal properties, i.e., ability to act both as donor and acceptor of hydrogen bonds and its compact size, to perform this function maximizes bond/volume ratio. Comparison of a series of carbohydrates bound to a lectin, or a series of lectins bound to the same ligand, reveal common waters, suggesting that they are important for protein-ligand interactions (Elgavish & Shannan 1998; Ravishankar *et al.,* 1998; Bouckaert *et al.,* 1999). Water bridges have been used in PNA as a strategy for generating carbohydrate specificity (Ravishankar *et al.*, 1998). Typically, the carbohydrate-binding site in legume lectins is made up of residues from four loops (Young & Omen, 1992), named as A, B, C and D loops (Sharma & Surolia, 1997). The first three loops contain a constellation of four invariant

combining residues: an aspartic acid in loop A, a glycine (conserved in all legume lectins except Con A) in loop B, an asparagine and an aromatic residue or leucine in loop C. They interact with the monosaccharide in all legume lectins irrespective of their specificity. The fourth loop (D), the monosaccharide specificity loop is hypervariable. The invariant aspartate and asparagine residues are also involved in calcium ion coordination, which explains the requirement of metal ions for carbohydrate binding. Another characteristic of the site is the presence of a cis-peptide bond in loop A, between the aspartic acid and its predecessor that keeps the aspartate side-chain in the proper orientation. In spite of these highly conserved residues involved in carbohydrate-binding and their very similar constellation, fine distinction between Gal and Glc/Man is achieved by the different spatial orientations that the two sugar moieties adopt in the carbohydrate-binding site (Lis & Sharon, 1998). Figure 1.4 (a) and (b) illustrate the Man-binding site in Con A (Naismith *et al.,* 1994) and the Gal-binding site in WBAI (Prabu *et al.*, 1999). The direct protein-sugar interactions common to all Man/Glc and Gal/GalNAc lectins involve four hydrogen bonds. They are (in WBAI sequence numbering) Asp87 OD1-Q3, Gh/105 N"O3, Asnl28 ND2 O3 and Asp87 OD2 O4. In addition, stacking interactions also exist in the case of all Gal/GalNAc lectins. There are interactions specific to each legume lectin, which also endow it with its characteristic carbohydrate specificity. For example, in the case of WBAI, there are additional interactions involving Asp212 and His 84. When the lectin binds a disaccharide, interactions with the second sugar residue also occur. Much of the structural work so far has been on Man/Glc and Gal/GalNAc specific legume lectins. However, the crystal structures of lectins I (Gohier *et al.,* 1996; Audette *et al*, 2000) and II (Loris *et al.,* 2000) from *Ulex europaeus*, which are fucose specific or exhibit

promiscuous binding, have also been reported. Protein-sugar interactions in them exhibit similarities as well as differences with those in Man/Glc and Gal/GalNAc lectins.

1.7.3 Cereal lectins

The best example of Cereal lectins is wheat germ aggiutinin. It also happens to be the second lectin to be X-ray analysed after Con A. The lectin is highly rich in cysteine residues. Each subunit in the dimeric lectin has a unique architecture with no regular secondary structural elements. The tertiary structure can be divided into four domains with 43, 43, 43 and 42 residues, respectively. The domains have a high structural homology and partial sequence homology. Each domain has four disulphide bridges and numerous hydrogen bonds that stabilize it. The structures of three isolectins of WGA I, II (Wright, 1989) and III (Harata *et al.,* 1995) were determined. The lectins from *Hevea brasiliensis* and *Urtica dioica* are the other lectins belonging to this class and their structures, in complex with chitobiose (Asensio *et al.,* 1995) and tri-N-acetylchitotriose (Harata & Muraki, 2000), have been determined. Recently, a couple of lectins from poke weed have also been determined. They are PL-C (Hayashida *et al.,* 2003) and two isolectins PL-D1 and PL-D2 (Fujii *et al.,* 2004). They contain different numbers of chitin-binding domains.

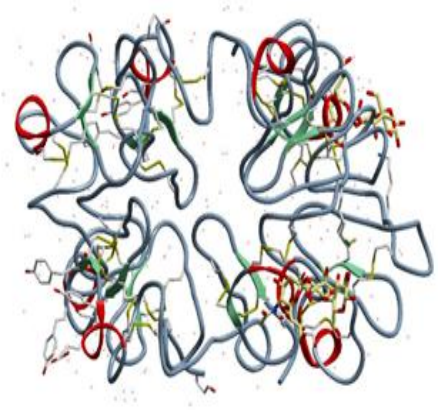

Figure 1.5: Crystal structure of a sub unit of wheat germ aggiutinin. Sugars in A, B, C and D sugar-binding sites are shared by the two subunits in the dimer (PDB1: WGC)

1.7.4 β-Trefoil lectins

Ricin was the first crystal structure to be solved from this class of lectins (Rutenber *et al.,* 1991). Subsequently, structures of abrin-A (Tahirov *et al.,* 1995), amaranthin (Transue *et al.,* 1997) and Mistletoe lectin I (Sweeney *et al.,* 1998) have also been determined. The tertiary type-II RIP, Ricin (PDB: 2AAI) structures exhibit the β-trefoil fold (Figure: 1.6). Crystal structure of the carbohydrate-binding domain of ricin and abrin-A are Type-II ribosome-inactivating proteins (RIPs) and they contain two chains each with a molecular weight of ~ 60 kDa. Two disulphide bridges link the two chains, A and B. The B chain is the lectin subunit and contains the carbohydrate-binding site, while the A subunit is responsible for the cytotoxic activity. The A chain is a globular subunit with extensive secondary structures, both α-helices and

β-sheets, with a prominent cleft that is assumed to be the site responsible for the toxic action. The B chain is made of two globular domains each possessing the β-trefoil fold. Each domain is made up of three homologous 40 residue subdomains and a link domain (Figure 1.6). Amaranthin has no appreciable sequence similarity with the B chain of the RIPs. Its carbohydrate-binding site is at a different position as compared to those in RIPs. Although amaranthin subunit has two β-trefoil domains, only one of them binds carbohydrate while the other is involved in oligomerization. Crystal structures of ebulin-1, a non-toxic type-II ribosome-inactivating protein, complexed with Gal, lactose and pteroic acid have also been reported (Pascal *et al.*, 2001)

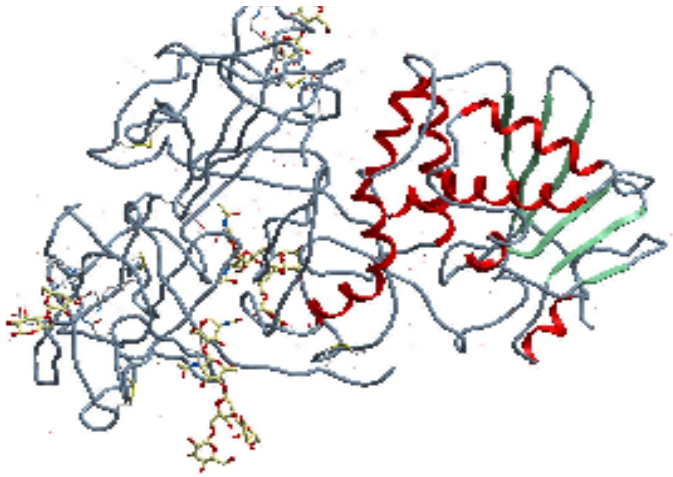

Figure 1.6: Crystal structure of the carbohydrate –binding domain of Type II RIP, Ricin (PDB: 2AAI)

1.7.5 β-prism II fold lectins

Lectins with an exclusive β-prism II fold was the fourth plant lectin fold to be identified. All of them have Specificity for mannose but differs to some degrees in their fine specificities for oligomannosides. This fold, first found in *Galanthus nivalis* agglutinin -Snowdrop lectin (Hester *et al.*, 1995), contains three four-stranded β-sheets that are nearly perpendicular to the three-fold axis, forming a 12-stranded β-barrel (Figure 1.7). One of the strands comes from a neighbouring subunit as part of strand-swapping arrangement between the two subunits. The lectin from garlic (Chandra *et al.,* 1999) is dimeric while snowdrop lectin is tetrameric in nature. Although the saccharide-binding sites and the overall structure are similar in the subunits of snowdrop and garlic lectins, their specificities to glycoproteins such as GP 120 vary considerably. These differences were attributed to the differences in oligomerization (Chandra *et al.,* 1999), implying that variation in quaternary association may be a mode of achieving oligosaccharide specificity in bulb lectins. Extensive modeling studies have also been carried out on the multivalency of garlic lectin (Ramachandraiah *et al.*, 2003). The structures of two other mannose-binding lectins, one from daffodil (Sauerborn *et al.*, 1999) and the other from bluebell (Wood *et al.,* 1999), have been determined. Both of them are tetrameric and share the same β-prism II fold. Each of the four-stranded β-sheet of this fold possesses a potential saccharide binding site. Unlike other lectins in this family, the two isoforms of gastrodianin from orchid *Gusirodia elata*, the structures of which have been determined and showed antifungal activity (Liu *et al.,* 2005).

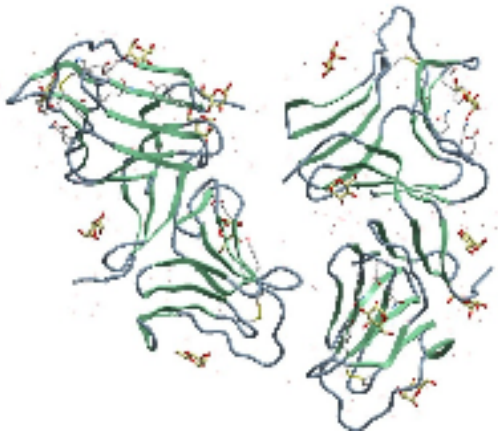

Figure 1.7: crystal structure of Garlic lectin

1.7.6 β-prism I fold lectins

The seeds of jackfruit contain two lectins, jacalin and artocarpin. It is in the structure of jacalin that β-prism I fold was first recognized as a lectin fold (Sankaranarayanan *et al.*, 1996). The lectin is galactose specific with one binding site on each subunit of the tetrameric molecule. Each monomer is made up of three Greek keys, one broken, which are arranged like the three faces of a prism. The strands in the Greek keys are parallel to the 3-fold axis of the prism. The symmetry in 3-dimensional structure is not reflected in the sequence. The break in one of the Greek keys is caused by a post-translational proteolysis, which results in the generation of a free amino group. This free amino group is important for the recognition of galactose. The extended binding site, relevant for oligosaccharide binding of jacalin has been thoroughly characterized through series of X-ray studies on jacalin-sugar complexes (Jeyaprakash *et al.*, 2003, 2004). The β-prism I fold has been shown

to be shared by two other galactose-binding lectins, those from *Maclura pomifera* (Lee *et al.,* 1998) and *Artocarpus hirsuta* (Rao *et al.,* 2004). Although all the β-prism I fold lectins have similar tertiary structure, they assume widely different quaternary structures. Jacalin, MPA and artocarpin are tetrameric. Heltuba is octameric while calsepa and banana lectins are dimeric. Very recently, another β-prism I fold lectin from *Parkia platycephala* has become available. The lectin molecule is dimeric. However, each subunit contains three β-prisms I fold domains such that the dimeric molecule contains six β-prism I fold domains (Gallego *et al.,* 2005).

1.8 Crystallization

Crystallization is obiously the first and often the rate limiting step in protein crystallography. The success of protein crystal structure determination depends much on the quality of the protein crystal grown. Over the years, a number of methods have been developed to grow quality protein crystals. Among these, vapour diffusion, batch method, free interface method and dialysis method are commonly used (McPherson, 1982; Giege & McPherson, 2001). The method of vapour diffusion is undoubtedly regarded as the most widely employed approach for crystallization. Vapour diffusion methods include hanging drop, sitting drop, sandwich, and capillary methods. The most common protocols are the hanging drop and sitting drop methods (McPherson, 1998).

1. 9 Plan of the work

Since early twentieth century, lectins were known for their cell agglutinating and carbohydrate binding property. It is now evident that lectins perform a variety of functions in nature and their mechanisms have also been unravelled. At present, hundreds of lectins are well characterised and

their number is growing fast. These lectins have different properties and characteristics and have been used for various biochemical as well as biophysical studies. The primary objective of this study was to isolate and purify lectin from *Spatholobus parviflorus* and study its biochemical and biophysical properties. It was also proposed to ananlyse its X- ray structure.
In the present study, we have isolated a lectin from the tropical plant *Spatholobus parviflorus* belonging to the family leguminosae. The investigations also involved various proces as follows.

Characterization.
Antifungal property analysis.
Crystallization.
Structure analysis using X-ray crystallography.