REVIEW ARTICLE

Structural biology of plant lectins and macromolecular crystallography in India

Nukathoti Sivaji, Kaza Suguna, Avadhesha Surolia and Mamannamana Vijayan*

Molecular Biophysics Unit, Indian Institute of Science, Bengaluru 560 012, India

Earnest attempts at initiating macromolecular crystallography, which is central to modern biology, started in India in the 1970s. The most important component of the efforts was on plant lectins. The lectin programme was continuously and almost exclusively supported by the Department of Science and Technology, Government of India. Lectins are quintessential carbohydrate-binding proteins and have acquired considerable importance in relation to biological recognition at cell surface. The plant lectin programme at Bengaluru spanned 40 years and was instrumental in making significant contributions to glycobiology. The work has also produced important results pertaining to protein folding, quaternary association and strategies for generating ligand specificities. The role of the lectin programme in mentoring scientists has been important and many of the senior macromolecular crystallographers in the country were initially trained under this programme. Studies on plant lectins led to those on mycobacterial and archeal lectins. The lectin work has also served as a springboard for the initiation of several other long-range programmes which, in addition to yielding important results, have also helped in training scientists, many of whom are again leaders of structural biology in India. Perhaps, majority of macromolecular crystallographers in India are those who have been trained at Bengaluru and their descendants. Particularly in recent years, there has been an influx of scientists trained in other centres, who have added vibrance to the macromolecular crystallography community. The community is now reasonably coherent, with constructive interactions among its members, and with several common programmes and shared facilities.

Keywords: Ligand specificity, macromolecular crystallography, plant lectins, protein folding, quaternary association.

LECTINS are often described as multivalent carbohydrate-binding proteins of non-immune origin, which specifically recognize different sugar structures[1–3]. They were originally discovered in plants and their best-known property was the ability to agglutinate red blood cells. Therefore, they were described as phytohemagglutinines. Many plant lectins have specificity for different blood groups. Subsequently, lectins were found in animals, algae, bacteria, fungi, yeast and viruses as well[4]. Most recently, in our laboratory, we have established the presence of lectins in archea. Thus lectins exist in all the three domains of life, viz. eukaryotes, eubacteria and archea. It would therefore appear that lectins are a group of important proteins which presumably originated before the three domains diverged.

Sugar and their polymers, viz. oligosaccharides and polysaccharides, abound in nature. They often occur in combination with proteins and lipids. Cellulose, a polymer of the sugar glucose, is the most abundant biomolecule in nature as it forms a structural component of plants. There are also many other polysaccharides with important structural roles. Sugars can polymerize in different ways. For instance, some sugars polymerize in ways different from that observed in structural polysaccharides, to yield important food materials, e.g. sucrose, starch, etc. in the form of carbohydrates. In spite of their importance as structural materials and components of food, they received much less attention than proteins and nucleic acids. The situation changed with the realization that most of the recognitive processes on the cell surface are mediated by sugars. Consequently, lectins which are quintessential sugar-binding molecules with recognitive functions, began to receive considerable attention. To start with, much of this attention was on plant lectins. It is in these early stages that we initiated our work on plant lectins which enabled us to make important, globally competitive contributions in the area. Furthermore, structural biological studies on plant lectins turned out to be an important vehicle for the initiation and development of macromolecular crystallography in India. In what follows, we outline the historical context and review our major findings in that context. We also describe briefly, succeeding and parallel efforts at Bengaluru on related and other systems.

Background

India has had a distinguished tradition in X-ray crystallography. G. N. Ramachandran (G.N.R.) and his colleagues

*For correspondence. (e-mail: mv@iisc.ac.in)

gave a head start to the country in structural biology, in addition to contributing to the foundations of crystallography. However, most of their efforts in structural biology were computational in nature. G.N.R. very much wanted to initiate macromolecular crystallography, the most important component of structural biology in India, but the time was not ripe for doing so. A few Indians were involved in early macromolecular crystallography studies abroad. However, one of us (M.V.) was the first trained macromolecular crystallographer to return home.

When M.V. returned to India and joined the Indian Institute of Science (IISc), Bengaluru in 1971, the available resources were totally inadequate to start structural work on macromolecules like proteins. The interactions between crystallographers and biochemists were also inadequate. Therefore, much of his early work was on crystalline complexes involving amino acids and peptides, with emphasis on aggregation and interaction patterns. This effort turned into a major programme with its implications for supramolecular association and chemical evolution. However, he was constantly on the lookout for avenues to initiate macromolecular crystallographic studies. It is in this context, that interaction with A.S. became important. In 1978, M.V. was a comparatively young faculty member in the Molecular Biophysics Unit (MBU) at IISc. By then, A.S. had obtained his doctorate working on the biochemistry and physical chemistry of plant lectins and was at the Indian Institute of Chemical Biology, Kolkata. Their meeting at Hyderabad was arranged during late 1978 by a common friend, S. K. Poddar, who was a faculty member of the Department of Biochemistry at IISc. That was the beginning of structural biology studies on plant lectins in India. Forty years later, in 2018, the very last structural biology paper of (M.V. and A.S.) on plant lectins appeared in *Glycobiology*[5]. During these four decades M.V. and A.S. have collaborated on plant lectins. In the meantime, A.S. joined MBU as a faculty member in 1981. That naturally further facilitated the collaboration. K.S. joined MBU in 1988 with considerable experience in work on plant lectins in a US laboratory. Since then, the three of us (M.V., A.S. and K.S.) formed a core group in the structural biology programme at Bengaluru on plant lectins. Of course, as indicated later, many other distinguished scientists contributed to the work as students and postdoctoral fellows.

Early efforts and the national mandate

When we initiated crystallographic work at Bengaluru in 1978, lectin crystallography was at its infancy. Three-dimensional structure of only conconavalin A (Con A) from jack bean[6,7] and wheat germ agglutinin from *Triticum aestivum*[8] were then available. Our first effort at crystallization was on *Ricinus communis* agglutinin (RCA). Despite efforts lasting a couple of years, the lectin could not be crystallized. We had better luck with peanut agglutinin (PNA), which readily crystallized. A crystallization note outlining this work was published in early 1982 (ref. 9). In the meantime, M.V. had initiated a programme on the hydration of proteins and its consequences based on water-mediated transformation. Yet another major programme at MBU involved fibre diffraction studies on DNA by V. Sasisekharan in collaboration with S. K. Brahmachari and others. The Department of Science and Technology (DST), Government of India (GoI) had by then formulated their Thrust Area Programme. In 1982, V. Sasisekharan, M.V., A.S. and Brahamachari submitted, a comprehensive X-ray project involving work on DNA and proteins. After much deliberation, the project was sanctioned in 1983. In the course of discussions, it was agreed, particularly at the instance of then DST Secretary, S. Varadarajan and Chairman of the Programme Advisory Committee, N. Seshagiri, that the Bengaluru group should function as a national nucleus for the development of macromolecular X-ray crystallographic studies in India. The handsome grant provided by DST in 1983 was perceived as a first step in this direction. Within a few years, the work on fibre diffraction was wound up. Funds for studies based on water-mediated transformations were obtained from other sources. The lectin programme remained with DST. Overall support for the X-ray facility at MBU and that for the lectin programme has been provided by DST from 1983 to date. Indeed, during a period, the programme on lectins used to be projected as a flagship programme of DST.

In the early 1980s, the PNA molecule was large even by global standards. The available experimental and computational facilities in India were inadequate to deal with the problem. Furthermore, the unusual nature of the quaternary association of the protein, elucidated subsequently, also contributed to the difficulty in solving the structure. It was established that PNA is a tetramer like ConA. Most of the early efforts were to solve the structure using molecular replacement on the basis of this similarity and the known substantial sequence identity between PNA and ConA[10]. These efforts failed. The success during the first 10 years of lectin crystallography at Bengaluru was in crystallizing and characterizing important lectins[11–14], which formed a springboard for further detailed structural work. In addition to new crystal forms of PNA, several crystal forms of jacalin, a lectin from jackfruit (*Artocarpus integrifolia*) seeds and winged bean (*Psophocarpus tetragonolobus*) lectin were prepared during this period. Eventually, the structure of peanut lectin was determined by the multiple isomorphous replacement method in the early 1990s (ref. 15). This was closely followed by the structure analysis of jacalin, which revealed a new lectin fold[16]. These two major structure determinations, which attracted widespread global attention, truly launched the structural biology studies of lectins in the country.

They also have had great impact on the development of macromolecular crystallography studies in India.

Technological bottlenecks

Early 1990s also marked the advent of position-sensitive detectors in India. Till then, the method of choice was oscillation photography coupled with microdensitometry. In the absence of ready access to a synchrotron facility, collection of X-ray intensity data from a given protein typically took several months and often involved dozens of crystals. That left us at a considerable disadvantage in comparison with those in developed countries, who had ready access to synchrotron facilities on which data collection from a protein could be completed within a day or a few days, often using a single crystal. For us, the availability of position-sensitive detectors substantially changed the scenario. Using these devices, data from a protein crystal could now be collected in a few days. The absence of ready access to synchrotron facilities was still a major problem, but it was not as formidable as before. The first position-sensitive detector for X-ray intensity data collection was installed at Bengaluru as a national facility accessible to all macromolecular crystallographers in the country. Over the decades, different kinds of position-sensitive detectors have become routine tools for data collection and dozens of them are now distributed in different macromolecular crystallography laboratories in the country. The Bengaluru group has played an effective role in making this possible.

Yet another problem in the 1980s and the 1990s had to do with computational facilities. India was often under sanctions and the advanced countries, particularly USA, were hesitant to supply computer facilities to us. For example, even though GoI sanctioned funds for a supercomputer to IISc, on the occasion of its Platinum Jubilee in 1984, we could not procure one, because USA was unwilling to permit us to obtain a supercomputer on normal commercial terms. The same difficulty was encountered in relation to graphics workstations as well. These difficulties led to delays in the macromolecular crystallography studies. In the course of time, computer technology changed and one could then go ahead with the work even without mainframe supercomputers. Eventually, computation ceased to be a major problem. Thus, with the availability of position-sensitive detectors and the removal of the computational bottleneck, technological impediments that slowed down macromolecular crystallography studies were substantially removed, although the absence of an Indian synchrotron facility affected our competitive edge.

Structural information – an overview

Con A, the first lectin to be structure analysed, had what came to be described as the legume lectin fold[6,7]. The structure of wheat germ agglutinin was determined next, again in the 1970s. This lectin had the havein fold[8]. Type-II ribosome inactivating proteins (RIPs) have two chains connected by a disulphide bridge. One of them is a lectin chain made up of two identical domains. Through the structure analysis of ricin from *Ricinus communis*, it was established that each lectin domain has a β-trefoil fold[17]. In the first half of the 90s, β-prism II was identified as a lectin fold through the X-ray analysis of snow-drop lectin[18]. Our work on jacalin, one of the two lectins in jackfruit seeds, brought to light a fifth lectin fold described as β-prism I (ref. 16). Most of the plant lectin structures determined to date have one of the five folds mentioned above. We have dealt with four of these five in our laboratory. As shown in Table 1, these studies involved 90 crystal structures of 12 lectins and their sugar complexes. Most of the plant lectins are either specific to galactose (Gal)/*N*-acetyl galactosamine (GalNAc) or mannose (Man)/glucose (Glc) at the primary binding site. This specificity, where known, is also indicated in the table. The specificity of different lectins for oligosaccharides such as blood group substances, arises out of interactions at secondary binding sites.

There are a few other groups in India that have worked on the structural biology of plant lectins, often not as part of a long-range programme. For instance Dinkar Salunke, then at the National Institute of Immunology, New Delhi, used plant lectins to explore molecular mimicry[19–24]. T. P. Singh (All India Institute of Medical Sciences, New Delhi) and S. Ramakumar (IISc) have worked on type-II RIPs[25–28]. Others who have contributed significantly to the work on plant lectins include C.G. Suresh (National Chemical Laboratory, Pune)[29,30] and M. Haridas (Kannur University)[31–34]. However, the bulk of the structural biology studies on plant lectins in India emanated from our group. The primary focus of the present article is on this effort with its implications to macromolecular crystallography in India. Furthermore, the emphasis in what follows is on aspects of general interest rather than on specific details of individual lectins.

Structure of subunits (folds)

Lectin molecules generally exist as oligomers. The monomers, also referred to as subunits or protomers, of lectins exhibit only a limited number of structures (folds). Figure 1 shows the folds of the lectins under review here.

The legume lectin fold, also called β-sandwich or jelly roll fold, was structurally characterized in its essential features before the first legume lectin structure was determined in our laboratory. However, studies on PNA[15,35–48], basic winged bean lectin (WBAI)[8,49–53], acidic winged bean lectin (WBAII)[54], *Erythrina corallodendron* lectin (EcorL)[55–56] and *Dolichos lablab* lectin[57] (the last mentioned studied in collaboration with Nadimpalli Siva

Table 1. Basic data on plant lectins and their complexes studied at Bengaluru using crystallography

Lectin	Source	Specificity at primary binding site	No. of independent crystals analysed	Important references
Legume lectins				
Peanut lectin (PNA)	*Arachis hypogaea*	Gal	21	15, 35–48
Basic winged bean lectin (WBAI)	*Psophocarpus tetragonolobus*	Gal/GalNAc	14	8, 49–53
Acidic winged bean lectin (WBAII)	*Psophocarpus tetragonolobus*	Gal/GalNAc	2	54
Erythrina corallodendron lectin (EcorL)	*Erythrina corallodendron*	Gal/GalNAc	4	55, 56
Dolichos lablab lectin	*Dolichos lablab*	Man/Glc	4	57
β-prism I fold lectins				
Jacalin	*Artocarpus integrifolia*	Gal	19	16, 58–64
Artocarpin	*Artocarpus integrifolia*	Man/Glc	5	65, 66
Banana lectin	*Musa paradisiaca*	Man/Glc	4	67, 68
β-prism II fold lectins				
Garlic lectin	*Allium sativum*	Man	2	69–71
Remusatia vivipara lectin (RVL)	*Remusatia vivipara*	Man	1	72
β-trefoil lectins				
Snake gourd seed lectin (SGSL)	*Trichosanthes anguina*	Gal/GalNAc	3	5, 73
Bitter gourd seed lectin (BGSL)	*Momordica charantia*	Gal/GalNAc	9	74

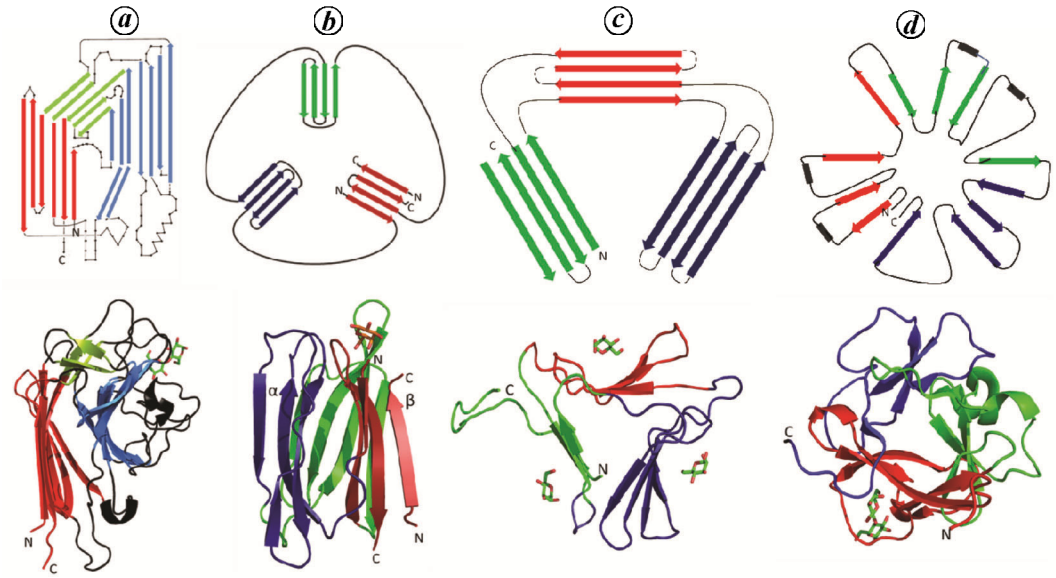

Figure 1. Subunit structures of four families of plant lectins studied in the laboratory. The top panel illustrates the folds and the bottom panel shows the corresponding three-dimensional structures. Bound sugars are shown as sticks. ***a***, Legume lectins (PNA; PDB code: 2PEL); ***b***, β-prism I (Jacalin; PDB code: 1UGW); ***c***, β-prism II (Garlic lectin; PDB code: 1BWU); ***d***, β-trefoil (one lectin domain of SGSL; PDB code: 4HR6).

Kumar, from the University of Hyderabad), brought out some of the unappreciated aspects of this fold and provided insights into the variability within the fold[77]. The features, outlined below, incorporate this additional information as well. The fold is robust and remains intact even when circular permutations in the sequence or glycosylation occur.

The legume lectin fold, as observed in PNA for example, essentially consists of a six-stranded, flat 'back β-sheet', a seven-stranded, curved 'front β-sheet' and a smaller, five-stranded 'top β-sheet' connecting the first two. The β sheets are well conserved among legume lectins. Variations occur through the elaboration of the loops. The subunit with a legume lectin fold has two hydrophobic cores, one bounded by the three sheets and the other between the curved front sheet and the long overhanging loops that connect the strands in it. The legume lectin fold is found in several other families of proteins as well. A comparison across families reveals that the presence of the second hydrophobic core is related to the curvature of the front sheet and the lengths of the loops that overhang in front of it. When the loops are long, the sheet curves and a second hydrophobic core is generated between the sheet and the loops.

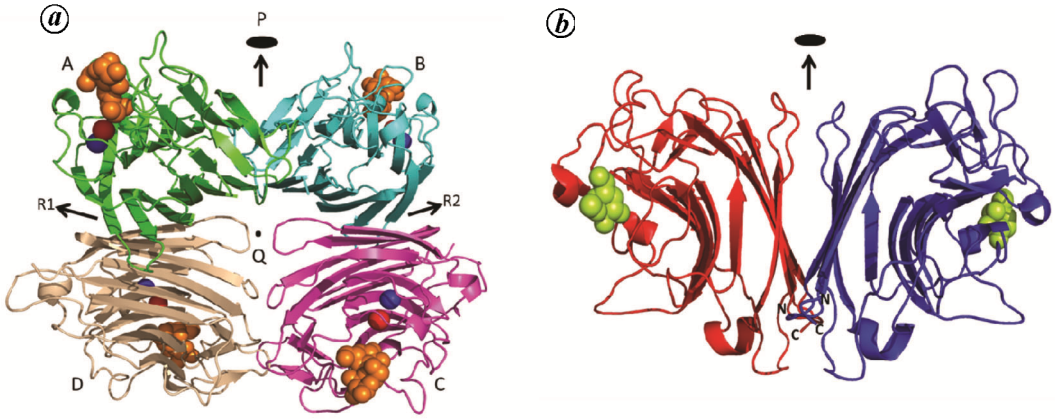

Figure 2. ***a***, Structure of PNA (PDB code: 2PEL). The four subunits are coloured differently. The bound lactose molecules are shown in orange space-filling representation. ***b***, Structure of WBAI (PDB code: 1WBL). The bound Me-α-Gal molecules are shown in limon space-filling representation.

The β-prism I fold was identified as a lectin fold through the X-ray analysis of jacalin[16,58–64]. Since then, a few plant lectins with β-prism I fold have been reported. In our laboratory, the structure of artocarpin, the second lectin from jackfruit seeds, was subsequently determined[65,66]. Artocarpin is a single-chain subunit which is 149 amino acid residues long. Two chains, an α-chain consisting of 133 amino acid residues and a β-chain comprising 20 residues, constitute the subunit of jacalin. The two chains are produced by a post-translational proteolysis of a single-chain processor which has 44.2% sequence identity with the single-chain artocarpin subunit. The two lectins have the same fold and nearly the same three-dimensional structure. The two chains of jacalin form an integral part of the same three-dimensional structure. The β-prism I fold essentially consists of three Greek keys aligned as three faces of a threefold symmetric prism, the strands in the Greek keys being parallel to the threefold axis (Figure 1 *b*). Loops which connect the strands are sequestered at the two ends of the prism. The threefold symmetry in the three-dimensional structure is not reflected in the sequence. Jackfruit tree is a dicot. A monocot lectin, viz. banana lectin, was also studied in our laboratory[67,68] and elsewhere. Banana lectin also exhibits the β-prism I fold. Unlike in the dicot lectins, the threefold symmetry of the structure is reflected in the amino acid sequence of banana lectin.

The β-prism II fold (Figure 1 *c*) was characterized as a lectin fold for the first time through the X-ray crystallographic analysis of snowdrop lectin[18]. The first lectin with a β-prism II fold to be determined in our laboratory was from garlic[69–71]. *Remusatia vivipara* lectin (RVL), the structure of which was determined subsequently by us[72] in collaboration with B. M. Swamy (Karnataka University, Dharwad), also exhibits the β-prism II fold. The subunit of β-prism II fold lectins is nearly as long as that of β-prism I fold lectins. They also exhibit threefold symmetry. However, in contrast to the β-prism I fold, strands in the β-prism II fold are perpendicular to the threefold axis and symmetry of the structure is reflected in the sequence as well.

The β-trefoil fold was well characterized in other proteins, much before it was observed in the lectin chains of ricin[17] and abrin[28]. Type-II RIPs consists of a catalytic chain and a lectin chain. The lectin chain is made up of two similar domains, each having three homologous subdomains α, β and γ (Figure 1 *d*). Each sub-domain has a four-stranded β-sheet formed by three strands of one subdomain and the fourth strand from a neighbouring subdomain. The twist in the sheet is extensive due to which a cylindrical β-barrel is formed. There are non-toxic homologues of type-II RIPs as well. Snake gourd seed lectin (SGSL)[5,73] and bitter gourd seed lectin (BGSL)[74] studied in our laboratory (the former in collaboration with M. J. Swamy of the University of Hyderabad) are such homologues. Much of the discussion presented below is concerned with the additional variations in the structure and function of lectins, based on the folds outlined above, generated through the variability in multimeric association (quaternary structure) and ligand specificity[75–78].

Quaternary association

Many proteins, perhaps a majority of them, exist as multimers. A well-established paradigm of protein architecture is that such multimers should have a 'closed' structure with point group symmetry. This paradigm is almost universally followed. Thus, a dimeric protein would have twofold symmetry. The subunits in a tetrameric protein would be related to one another either by 222 symmetry or a fourfold axis. When the tetramer is a dimer of a dimers, the expected symmetry is 222. The most striking feature of the structure of PNA is that the tetrameric molecule does not possess 222 or fourfold symmetry[15,79]. As illustrated in Figure 2, the tetramer is a dimer of dimers. Subunits A and D form a dimer, while subunits B and C form the other dimer. Both are

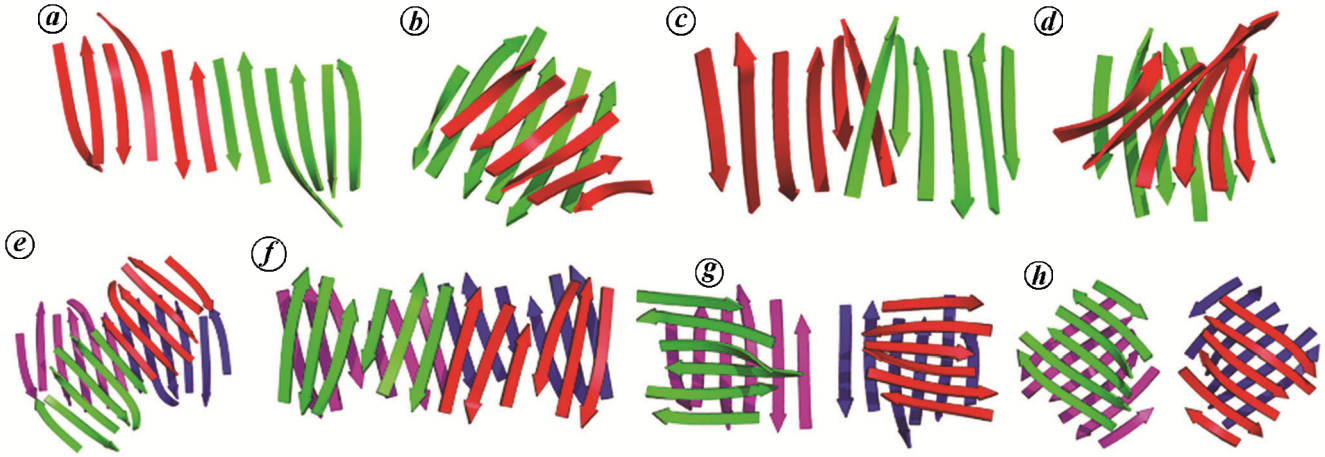

Figure 3. Schematic representation of the back β-sheets in dimeric legume lectins, viz. (***a***) pea lectin, (***b***) GS-IV lectin, (***c***) EcorL and (***d***) *Dolichos biflorus* leaf lectin (PDB codes: 2BQP, 1LEC, 1LTE and 1LUL respectively) and in tetrameric legume lectins, viz. (***e***) ConA, (***f***) soybean agglutinin, (***g***) PNA, and (***h***) *Griffonia simplicifolia* I-B4 lectin (PDB codes: 3CNA, 2SBA, 2PEL and 1LED respectively). Sheets from the four subunits are coloured differently. Image adapted from ref. 76.

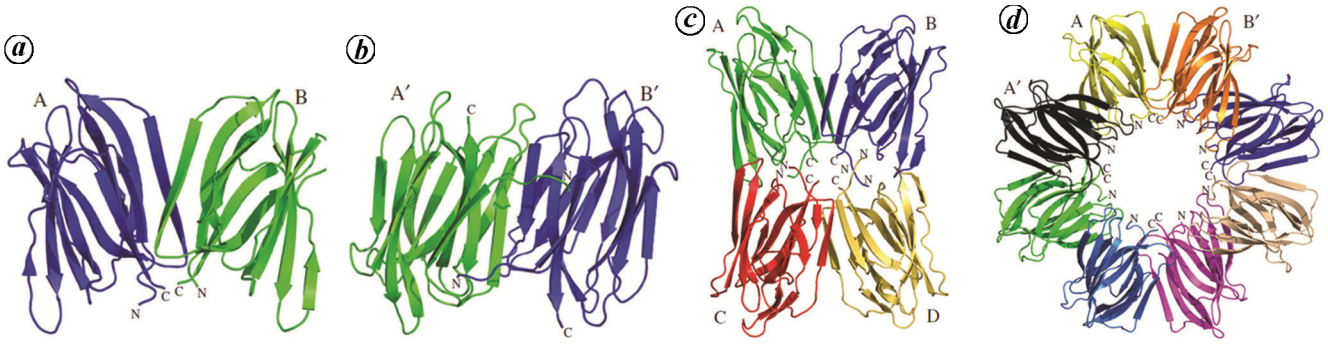

Figure 4. Quaternary structure in (***a***) banana lectin (PDB code: 1X1V), (***b***) calsepa (PDB code: 1OUW), (***c***) artocarpin (PDB code: 1J4S) and heltuba (PDB code: 1C3K). Image adapted from ref. 82.

individually twofold symmetric. The two dimers are also related to each other by a twofold symmetry. However, the three two-fold axes do not intersect to give 222 symmetry. Considering the two dimers, interactions exist only between subunits A and B. Subunits C and D are not in contact with each other. The structure of peanut lectin established that multimers with 'open' structures can also exist.

The work on PNA also led to a thorough examination of the variability in quaternary association of legume lectins. Con A, the first legume lectin to be structure analysed, is a dimer of dimers. In each dimer, the back β-sheet forms a contiguous 12-stranded β-sheet resulting from a side-to-side disposition of the two subunits. The same type of association was found in dimeric pea lectin and favin as well. Therefore, this type of demerization came to be described as 'canonical'. Subsequently, a back-to-back arrangement of back β-sheets, as in AD and BC in PNA, was found in *Griffonia simplicifolia* lectin (GS IV)[80]. It was suggested that this non-canonical mode of dimerization in GS IV was caused by interactions involving covalently bound sugar and to avoid the burial of a glutamic acid residue. Soon afterwards, a related handshake mode of dimerization was observed in EcorL[81]. Presence of a glycosylation site at the interface was suggested to be the reason why EcorL cannot form canonical dimers. PNA is not glycosylated, yet it exhibits a non-canonical mode of dimerization. This indicates that the departure from canonical mode is not necessarily caused by interactions involving covalently bound sugar.

The issue was further examined using the structures of winged bean lectins[49,54]. WBAI has a sequence identity of 54% with EcorL. However, unlike in EcorL, the glycosylation site in the former is far removed from the inter-subunit interface of the canonical dimer. Thus WBAI would associate into a canonical dimer if the formation of such a dimer is prevented by covalently linked sugar. On the contrary, it would form a handshake-type dimer as in EcorL, if the non-canonical mode is adopted for reasons intrinsic to the protein. It was found that the crystal structure of WBAI exhibited the handshake mode of dimerization (Figure 2), in consonance with differential scanning calorimetry studies carried out on WBAI and EcorL[82,83].

The same was true for WBAII as well. Subsequently, the crystal structure of non-glycosylated recombinant EcorL (rEcorL) was determined[55]. This also adopts a handshake mode of dimerization. It would thus appear that the modes of dimerization in legume lectins are caused by factors intrinsic to the protein. Once the oligomeric structure is set, glycosylation does not alter it.

A detailed examination showed that legume lectins are a family of proteins in which small alterations in essentially the same tertiary structure lead to large changes in quaternary association[79]. All legume lectins of known three-dimensional structure are dimers or tetramers (dimers of dimers). A variety of modes of quaternary association are observed in the dimeric as well as tetrameric lectins (Figure 3). The back β-sheet is involved in oligomerization in all of them. The observed variability in quaternary association could be rationalized to a substantial extent using computational approaches.

Research on legume lectins was the first among detailed studies on the variability in the quaternary association of proteins with essentially the same tertiary structure. Subsequently such variability was recognized to be widespread. Among plant lectins, those with β-prism I fold also exhibited substantial variability in quaternary association in spite of all of them having nearly the same tertiary structure[78]. This variability is most striking in mannose binding β-prism I fold lectins, the subunits of which are made up of a single polypeptide chain (Figure 4). Among them, banana lectin and calsepa are dimeric, artocarpin is a tetramer and heltuba is an octamer. The modes of dimeric association in banana lectins and calsepa are different. Tetrameric artocarpin and octameric heltuba are made up of banana lectin-type dimers. The variability in the quaternary association in β-prism I fold lectins was also studied and rationalized using energy minimization of different models and molecular dynamics (MD) simulations[84].

Substantial variability in quaternary association is exhibited by β-prism II fold lectins as well (Figure 5). For instance, *Gastrodia elata* lectin (GEL) is a monomer, a rare example of monomeric lectin[85]. Typical of dimeric β-prism II fold lectins is garlic lectin, in which dimerization involves a strand exchange[69]. Snowdrop lectin is a tetramer made up of two garlic lectin-type domains[18]. The structural basis of this variability could also be understood to a substantial extent.

The type-II RIPs, including the two non-toxic homologues studied by us, are technically monomeric molecules. However, most of them contain a lectin chain and a toxic chain held together by a disulphide bridge (Figure 6). It has been demonstrated that the structural integrity of the two-chain molecule is likely to be preserved even when the disulphide bridge is disrupted[5]. There are also a few type-II RIPs, including BGSL, in which two two-chain units are further linked by a disulphide bridge, giving rise to a four-chain molecule[74]. The differences in the mutual orientation of different chains in type-II RIPs are akin to those in the mode of quaternary association in multisubunit proteins. This aspect of the structures of type-II RIPs has been explored only to a limited extent.

Protein–carbohydrate interactions and strategies for generation of ligand specificity

The most important characteristic feature of lectins is their ability to distinguish different sugars structures (specificity) and to bind to appropriate sugars using different interactions (affinity). High affinity is often generated through multivalency (see later) and interactions

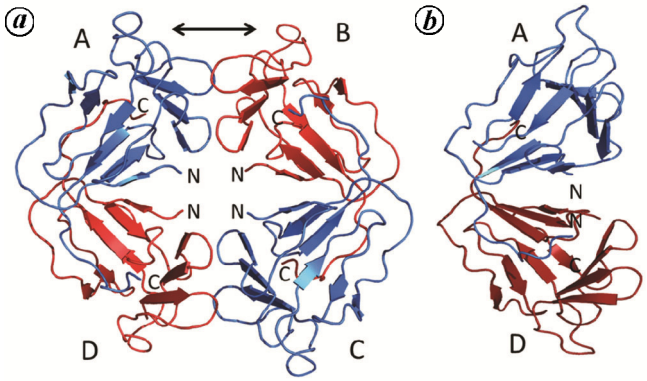

Figure 5. Structures of (***a***) tetrameric snowdrop lectin (PDB code: 1JPC) and (***b***) dimeric garlic lectin (PDB code: 1BWU). The branched mannopentose of GP120 is considered to bind in the region spanning A and B (indicated by the double arrow).

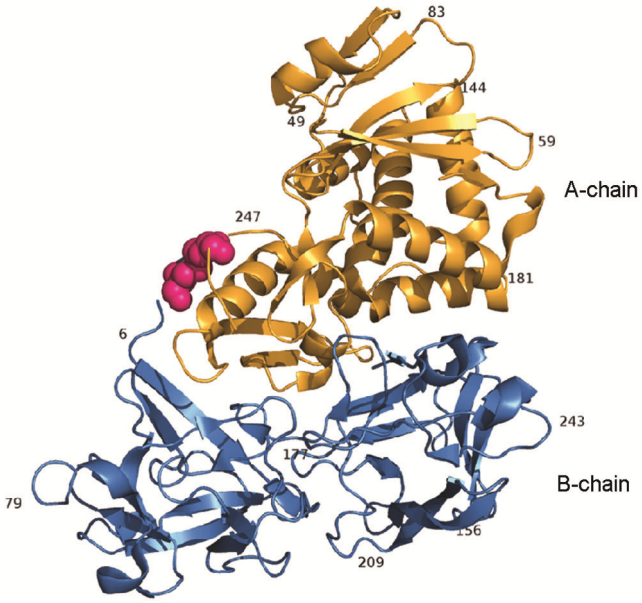

Figure 6. Structure of SGSL (PDB code: 4HR6). The catalytic chain (A-chain) is shown in orange and the lectin chain (B-chain) in marine. The disulphide bond between the two chains is shown as spheres.

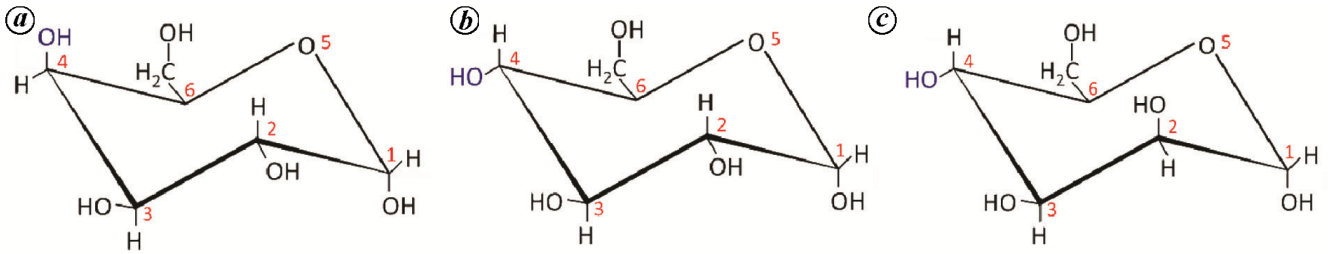

Figure 7. Chemical structure of (***a***) galactose, (***b***) glucose and (***c***) mannose. In galactose substitution of the hydroxyl group at 2 is equatorial and at 4 is axial. In glucose,substitution at 2 remains equatorial while that at 4 is also equatorial. In mannose, substitution at 2 is axial and at 4 is equatorial. The figure is that of the α-sugars in which the hydroxyl group at 1 is axial. In β-sugars, the hydroxyl group would be equatorial.

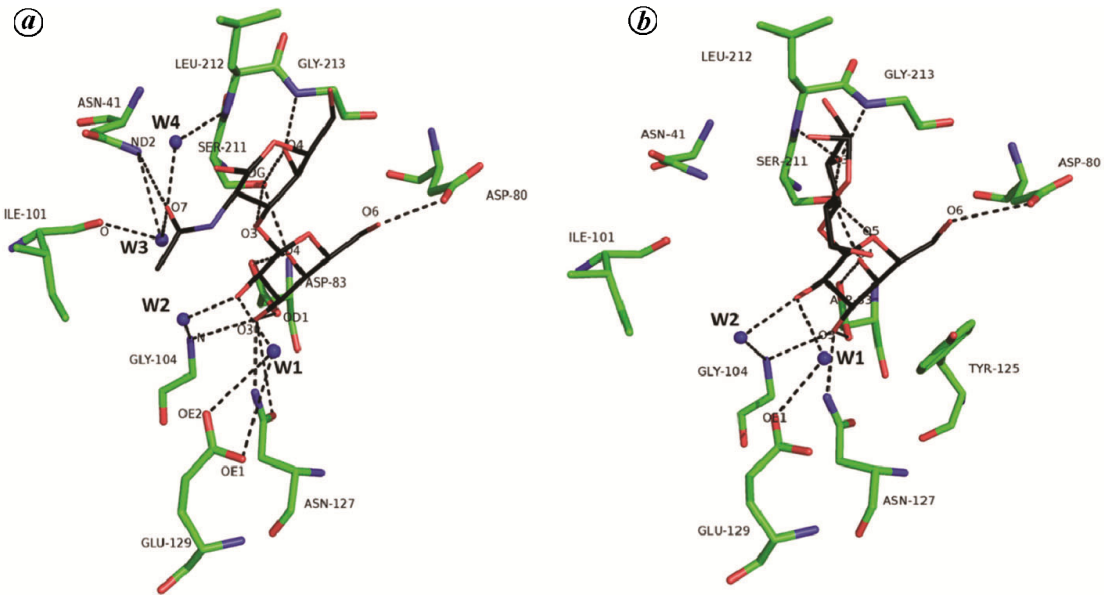

Figure 8. Protein–sugar interactions in the complexes of PNA with (***a***) T-antigen and (***b***) lactose. The main difference between the two is the water-mediated interactions involving W3 and W4, which exist only in the T-antigen complex.

with different regions of the concerned oligosaccharide (oligomers of different monomeric sugars). The first level of specificity is based on interactions with the terminal sugar. The location at which this terminal sugar binds in the lectin molecule is often referred to as the primary binding site. On this basis, most plant lectins can be classified as Gal/GalNAc-binding lectins and Man/Glc-binding lectins. Gal, Man and Glc have the same chemical structures (Figure 7). The difference is only in the configuration at ring carbon atoms. The main difference between Man and Glc on the one hand and Gal and GalNAc on the other hand, is in the orientation of O4. This difference leads to differences in the distribution of partial charges in the ring in such a way as to facilitate the stacking of an aromatic ring against the rings in Gal and GalNAc. Therefore, the combining site of Gal/GalNAc-specific lectin is always associated with aromatic residues.

PNA is among the very few lectins which binds Gal, but not GalNAc. The reason for this could be deduced from MD studies[44]. A point mutant of PNA which binds to Gal and GalNAc with equal affinity could also be produced[39]. EcorL binds Gal and GalNAc equally well. It was also possible through site-directed mutagenesis to produce a mutant of EcorL which showed preference for one over the other[56]. These mutational studies have provided further insights into the structural basis of the carbohydrate specificity of these lectins.

Legume lectins could be Man/Glc-specific or Gal/GalNAc-specific. The same is true with β-prism I fold lectins. β-prism II fold lectins are always Man/Glc-specific, while type-II RIPs are always Gal/GalNAc-specific. Within each group of lectins, further levels of specificity are generated through interactions involving sites other than the primary site as well. It is important to delineate the interactions at primary and secondary sites to elucidate the structural basis of carbohydrate specificity and its consequences. That is what we have done in relation to the lectins we have undertaken for study. Details of this effort can be obtained from primary publications. Here we only deal with some representative examples to illustrate a few general principles.

PNA is specific to Gal at the monosaccharide level. At the disaccharide level, the lectin has the highest affinity for Gal β1–3 GalNAc, which is the proper name for tumour-associated T-antigen. That confers diagnostic value to PNA. The affinity of the lectin for the T-antigenic disaccharide is 20 times more than that for, say, lactose which is a similar disaccharide. This higher affinity is achieved on account of water-mediated interactions between the lectin and the acetamido group of T-antigen[38] (Figure 8). Thus PNA demonstrates how water bridges can be used as a strategy for generating ligand specificity. Yet another strategy was brought to light when comparing WBAI and WBAII. Both have similar structures and are specific to Gal/GalNAc at the monosaccharide level. While WBAII can bind to blood-group substances A, B and O, WBAI can bind only to A and B and not O. Different blood groups are distinguished on the basis of the oligosaccharides that are attached to the surface of red blood cells. In all the three blood groups, the terminal residue of the oligosaccharide is Gal. However, the way the rest of the oligosaccharide is attached to Gal is different in A and B on the one hand, and O on the other hand. The attachment involves a β-link in O, while it is an α-link in A and B. In WBAI, the β-link results in a steric clash with a long loop, while no steric clash exists when it is an α-link[54] (Figure 9). In WBAII, the concerned loop is much shorter and hence can support

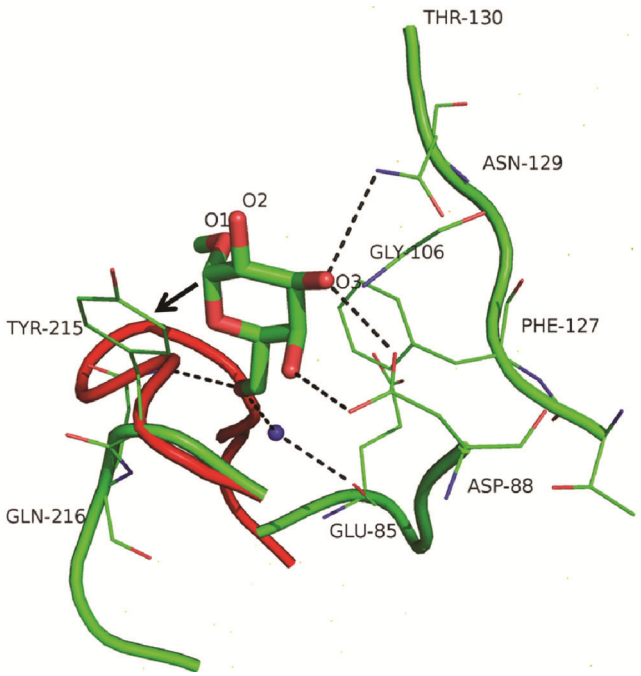

Figure 9. Protein–ligand interactions in the WBAI-Me-α-Gal complex (PDB code: 1F9K). The polypeptide stretches in the binding region are shown in thick green lines. A stretch from the polypeptide chain observed in WBAII is shown in red. The expected orientation of β-substituted oligosaccharides is schematically shown by an arrow. Such oligosaccharides can bind WBAI but not WBAII on account of steric interactions with the polypeptide strand illustrated in red.

β-linked oligosaccharides as well. This is an example where loop length has been used as a strategy for generating ligand specificity, as suggested in a comprehensive analysis[86].

The structures of a large number of sugar complexes of β-prism I fold lectins have been determined and the extended binding sites have been characterized[59–64]. However, the structure of Me–α–Gal complex of jacalin[16] and the Me–α–Man complex of artocarpin[65] define the basic framework for comparing the structural basis for sugar specificity of β-prism I fold lectins (Figure 10). It may be noted that one of the three Greek keys in jacalin is broken on account of post-translational proteolysis. In both cases, the loops at one end of the prism constitute the carbohydrate-binding site. There is one carbohydrate-binding site per subunit in both the cases. The binding sites are similar, but with a couple of important differences. The positively charged amino terminus of the chain generated by post-translational proteolysis, strongly interacts with O4 of Gal in jacalin. O4 in Gal has a different orientation from that in Man. This particular lectin–Gal interaction found in jacalin exists in other Gal-specific β-prism I fold lectins as well[78]. Thus the structure analysis brings to light post-translational modification as a strategy for generating ligand specificity. The combining site of Gal-specific β-prism I fold lectins also contains aromatic residues, unlike in the case of Man-specific ones. Another interesting observation has to do with the type of linkage. Jacalin has higher affinity for α-galactosides compared to that for β-galactosides. It was established that the difference in binding energy and therefore in affinity is caused by a distortion in the geometry of the sugar in β-galactosides[62]. Thus distortion of the ligand is a strategy for modulating affinity.

All Man-specific β-prism I fold lectins have similar primary binding sites where the mannose residue binds. However, differences exist in specificity when binding to oligosaccharides is considered. At least in a couple of instances, difference in specificity is generated by a difference in loop length[66]. Among the Man-specific β-prism I fold lectins of known three-dimensional structures, banana lectin occupies a special place[67]. In this lectin, the threefold symmetry of the structure is to an extent reflected in the sequence. Furthermore, each subunit carries two binding sites, which has interesting implications.

Another issue studied using β-prism I fold lectins has to do with the overall orientation of the oligosaccharide in relation to binding. Linear oligosaccharides have a 'reducing end' and a 'non-reducing end'. There are occasions when the reducing end occupies the primary binding site of the lectin, while in other instances the non-reducing end occupies the site. The structural basis for the choice of one or the other end was elucidated using crystallography and modelling[63,64,68].

The β-prism II fold lectins are in general mannose-specific. Most such lectins analysed so far have three

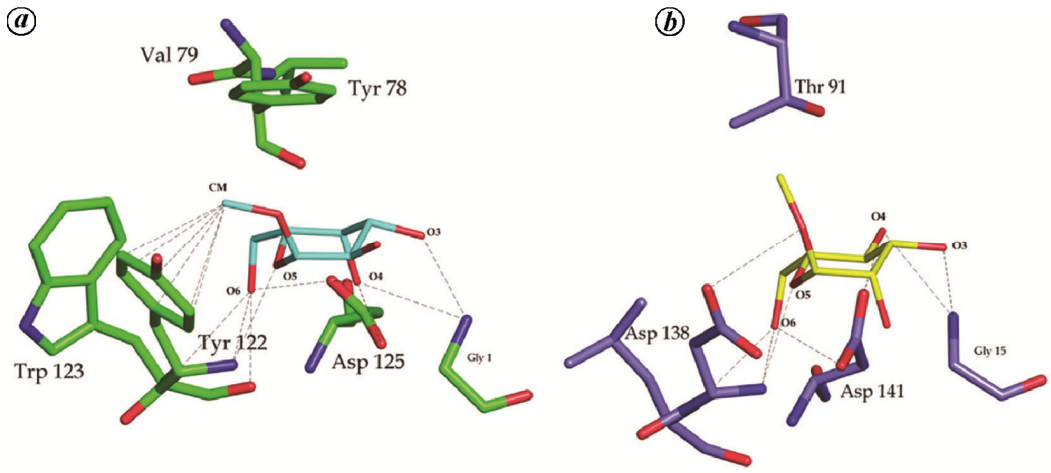

Figure 10. Sugar-binding sites of (***a***) jacalin (PDB code: 1JAC) and (***b***) artocarpin (PDB code: 1J4U). The bound sugars are Me–α–Gal and Me–α–Man respectively. Image reproduced from ref. 76.

binding sites per subunit, each Greek key carrying a site. RVL, analysed in our laboratory, has only one binding site per subunit[72]. The interactions of this family of lectins with mannose at the primary binding site are substantially conserved. However, differences occur in the interactions with oligosaccharides. An interesting insight in this respect is obtained when garlic lectin and snowdrop lectin are compared[69]. The former is a dimer, while the later is a tetramer. Snowdrop lectin binds strongly to the HIV surface glycoprotein gp120, while garlic lectin does so weakly. Structural studies indicated that the branched mannopentose of gp120 crosslinks two dimers in the tetramer of snowdrop lectin. Garlic lectin is a dimer without a partner to be cross-linked (Figure 5). This is an instance where oligomerization is used as a strategy for generating ligand specificity.

As mentioned earlier, type-II RIPs have a lectin chain made up of two β-trefoil domains, designated as 1 and 2. The trefoil in each domain is designated as α, β and γ. In most RIPs 1α and 2γ carry carbohydrate-binding sites. Defect in binding sites would contribute to non-toxicity on account of the inability of the protein to enter into the cell. There are some RIPs, mostly non-toxic, which carry only one binding site in the lectin chain[74]. In the case of SGSL, the crystal structure showed the presence of two sugar-binding sites, 1α and 2γ, while thermodynamic measurements indicated the existence of only one binding site. Detailed MD calculations led to an interesting observation. Both the sites can bind sugar, but only the ligand at one site is retained in a dynamic situation. Therefore, there is a subtle relationship between binding and retention[5].

Plasticity, water structure, conformational selection and multivalency

The number of independent crystal structures has been large in the cases of PNA[15,35–48], jacalin[16,58–64] and WBAI[8,49–53]. A detailed comparison of the different structures of each protein gave valuable information on the comparatively rigid and flexible regions of the molecule, and relatively invariant water molecules surrounding the protein molecule. For example, in the case of legume lectins, the back β-sheet was found to be more flexible than the front β-sheet. This is presumably related to the exclusive involvement of the back β-sheet in the variability in quaternary association. The functionally important water molecules are well conserved. An interesting and somewhat unexpected result was the occurrence of a substantial number of conserved water molecules in the loop regions. Nearly half of a typical legume lectin molecule is made up of loops which are not stabilized by hydrogen bonds, unlike in the case of helices and sheets. It would appear that this stabilizing role is undertaken by water bridges involving highly conserved water molecules[48].

The first MD simulation in our laboratory was carried out in the early years of this century on PNA and its complexes, in a collaborative effort involving John Helliwell and Gail Bradbrook (University of Manchester, Manchester)[44]. The calculations, among other aspects, provided fresh insights into the peculiar sugar specificity of the lectin and rationalization for one of the mutational studies. Subsequently, MD simulations have often been used to supplement information from crystallographic studies. One important result to emanate from this effort has to do with conformational selection. In the context of ligand binding, conformational selection involves using one or more possible conformations from an ensemble of native protein conformations for ligand binding. Induced fit involves the creation of a distortion in the existing native conformation to enable the ligand to bind. It turned out that in many cases both these mechanisms co-exist[81]. MD simulations also provided critical information on the variability of quaternary association.

Multivalency is considered to be an important property of lectins, but it is poorly understood in structural terms. Through crystallography, thermodynamics, dynamic light

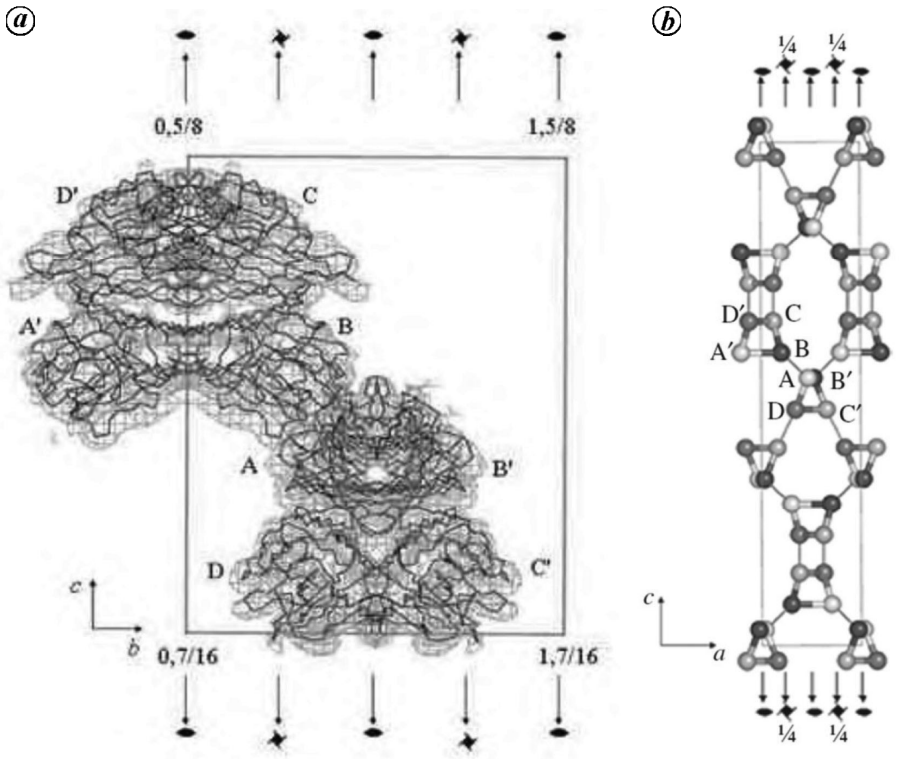

Figure 11. (***a***) Electron density for PNA molecules in a complex with a bidentate ligand and (***b***) arrangement of cross-linked molecules in the crystal. Image reproduced from ref. 46.

scattering and modelling studies, it was possible to characterize one mode of aggregation of PNA and a bivalent sugar using multivalency[46] (Figure 11). Similar studies, but not involving crystallography, on garlic lectin also led to modes of aggregation involving multivalency[71]. These studies demonstrated that interactions of multivalent lectins with multivalent sugars could lead to an ensemble of a finite number of distinct crosslinked periodic arrays.

Evolutionary implications

Structural and related studies on plant lectins brought to light several evolutionarily interesting results. Subsequent work on mycobacterial and archeal lectins introduced additional interesting dimensions to them. These additional dimensions are of course beyond the scope of this article.

At the level of folds, β-prism I fold and β-prism II fold exhibit substantial similarities. Both involve threefold symmetric arrangement of Greek keys or Greek key-like motifs. However, the Greek keys are arranged differently in the two cases. In β-prism I fold lectins, the strands in the Greek keys are parallel to the threefold axis, while they are perpendicular to the threefold axis in β-prism II fold lectins. The binding sites also exhibit differences. A simultaneous genomic search of the two types of lectins and a careful examination of structural results led to significant new insights[87].

The number of binding sites in each subunit varies between one and three in both types of lectins. The number is most often one, occasionally two and rarely three, in β-prism I fold lectins; it is always one in those from dicots. Banana lectin is from a monocot. The number of binding sites in each subunit of β-prism II fold lectins is most often three. The diversity in number of binding sites is unrelated to the taxonomical position of the organism. However, a reasonable correlation is seen between the symmetry in the sequence and number of binding sites. The β-prism I fold lectins from plants are found with comparable frequency in dicots and monocots. The β-prism II fold plant lectins mostly occur in monocots. Both the families of lectins are found in other domains of life as well.

A detailed analysis of the relevant genomes indicates that the two types of β-prism lectins perhaps had a common ancestry. Both types of lectins can be considered to have evolved through gene duplication and fusion of a primitive gene coding for a carbohydrate-binding motif involving Greek-key topology. The Greek keys assemble in different ways in the two types of lectins (Figure 12). Since their assembly, the two appear to have followed different evolutionary paths. There is perhaps further divergence in the path between dicots and monocots. In general, the number of binding sites is higher in monocots than in dicots. Dicots are characterized by a cork cambium layer to protect them, while such a layer is absent in monocots. Consequently, monocots probably have a higher dependence on other defence mechanisms, including that involving lectins. This could perhaps be the reason why monocot lectins carry a higher number of binding sites.

National impact – role in mentoring

A significant impact of lectin crystallography at Bengaluru has been in training leaders of structural biology in India. Many of these leaders cut their macromolecular crystallography teeth on plant lectins. A majority of Ph D students who worked on these proteins returned to India after postdoctoral stints abroad. Those who worked on peanut lectin at Bengaluru include Dinakar Salunke (International Centre for Genetic Engineering and Biotechnology (ICGEB), New Delhi), Shekhar Mande (Council of Scientific and Industrial Research (CSIR), New Delhi), Rahul Banerjee (Saha Institute of Nuclear Physics, Kolkata) and Ravishankar Ramachandran (Central Drug Research Institute (CDRI), Lucknow). R. Sankaranarayanan (Centre for Cellular and Molecular Biology (CCMB), Hyderabad) was involved in the solution of the structure of jacalin, while K. Sekar (IISc) was involved in the efforts as a postdoctoral fellow and faculty member. Nagasuma Chandra (IISc) was primarily responsible for solving the structure of garlic lectin when she was a postdoctoral fellow in our laboratory. Gosu Ramachandriah (Jubilant Biosys, Bengaluru) worked on the same protein for his Ph D programme. Venkatesh Pratap (CDRI) solved the structure of artocarpin, as a Ph D student. N. Manoj (Indian Institute of Technology-Madras (Chennai)), was involved in work on winged bean lectins, when he was a student. Kiran Kulkarni (National Chemical Laboratory (NCL), Pune) worked on basic winged bean lectin and EcorL during his Ph D programme. Subsequently, S. Thamotharan (Sastra University, Thanjavur), participated in the work on EcorL. Desh Deepak Singh (Panjab University, Chandigarh) solved the structure of banana lectin when he was a postdoctoral fellow in our laboratory, with the involvement of K. Saikrishnan (IISER, Pune). No other macromolecular crystallography programme has perhaps produced as many leaders of Indian science as the lectin crystallography programme.

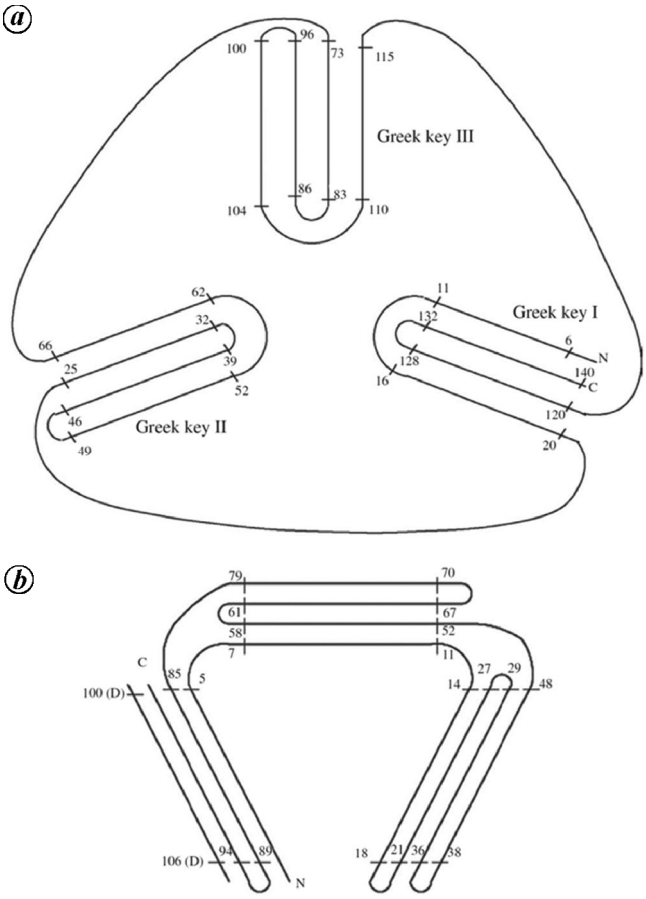

Figure 12. Schematic representation of the topology of (***a***) banana lectin fold and (***b***) garlic lectin fold. In (***b***), the strand from the other subunit involved in swapping to form the complete third Greek key has been labelled as 100(D) to 106(D). Image reproduced from ref. 84.

Towards microbial lectins

It is generally felt that most of the structural issues connected with plant lectins have already been elucidated. A substantial component of the structural information available globally has emanated from Bengaluru. It was perhaps time to move on to lectins from other sources. In view of the other concerted efforts that have been established at Bengaluru and elsewhere in India, in which three of us (K.S., A.S. and M.V.) were participants, it was appropriate to initiate work on mycobacterial lectins. A thorough genomic search led to the identification of some 94 lectins from 20/30 mycobacterial species/strains[88]. Two of them have been cloned, expressed and purified[89,90]. In addition to carrying out physico-chemical studies[91], the crystal structure of one of them has been determined[92]. Further efforts in this area are in progress. Work is also being done on lectins from the pathogenic organism, *Entamoeba histolytica*.

In relation to lectins, a domain of life that has remained unexplored is archea. Through an extensive genomic search, 46 lectins could be identified from 29 archeal species[93]. One of these lectins has been cloned, expressed and purified[94]. X-ray crystallography studies on it are in progress. With the identification of lectins in archea and the experimental work on one of them, it has been established that lectins exist in all three domains of life. It would also appear that lectins evolved to the present form well before the three domains diverged.

Future efforts on lectins at Bengaluru would revolve around those from microbes, including the ones mentioned above.

Parallel efforts at Bengaluru

For providing a correct perspective of the overall situation, it is important to briefly indicate the efforts on other

systems as well. The most important effort which progressed in parallel with the crystallography of lectins, especially in the early days, was the monumental work of M. R. N. Murthy and H. S. Savithri on the plant viruses, sesbania mosaic virus and physallis mottle virus. As indicated earlier, another parallel effort involved water-mediated transformations in proteins, pursued by one of us (M.V.). Yet another notable contribution, made by M.A. Viswamitra and S. Ramakumar, pertains to Xylanases.

At the turn of the century, by a deliberate choice, work had started on microbial proteins, particularly those from pathogens. The most important component of this effort was that on mycobacterial proteins with special emphasis on tuberculosis (TB) proteins. An effort was also made to develop a network on structural biology of TB proteins in the country. At present, more than 10 groups in India work on the crystallography of TB proteins. More than 10% of the structure determinations of TB proteins in the world has been carried out in India. The work on TB proteins at Bengaluru has encompassed many different systems. Those who contributed to the extensive programme on mycobacteiral proteins include crystallographers M. Vijayan, B. Gopal, Nagasuma Chandra, K. Sekar, M. R. N. Murthy, K.S. and S. Ramakumar, and biochemists K. Muniyappa, Umesh Varshney, Dipankar Chatterji, A.S., Sandhya Visweswariah and V. Nagaraja.

Work on the malaraial parasite has also been significant though less extensive and involved M. R. N. Murthy, K.S., A.S., Namita Surolia, P. Balaram and Hema Balaram. Other proteins studied include those from pathogens *Salmonella typhimurium* (M. R. N. Murthy and H. S. Savithri), *Entamoeba histolytica* (M. R. N. Murthy and Alok Bhattacharya), rotavirus (K.S. and Durga Rao). *Staphylococcus epidermidis* (B. Gopal) and *Bacillus subtilis* (B. Gopal). Work has also been carried out on phospholipases (K. Sekar), aspartic proteases (K.S.), heat shock proteins (K.S. and Utpal Tatu), methyl transferases (S. Ramakumar and D. N. Rao), pyridoxal phosphate dependent enzymes (M. R. N. Murthy and H. S. Savithri), and proteins from extremophiles (K. Sekar).

The above-mentioned studies also resulted in training future leaders of structural biology in India. Among them, H. S. Subramanya (Institute of Bioinformatics and Applied Biotechnology, Bengaluru), S. Srikrishna (Institute of Microbial Technology (IMTech), Chandigarh). S. Ramaswamy (InStem, Bengaluru), P. Gayathri (IISER, Pune) and Sangeetha (Amity University, Delhi) worked on plant viruses and related systems. Those who worked on mycobacterial proteins have also now begun to occupy independent positions in different organizations in India. Sunando Datta (IISER, Bhopal), K. Saikrishnan (IISER, Pune), Siddhartha Roy (Indian Institute of Chemical Biology, Kolkata), Prem Singh Kaushal (Regional Centre for Biotechnology, Faridabad), Jeyaraman Jeyakanthan (Alagappa University, Karaikudi), R. Krishna (Pondicherry University, Puducherry) and Krishan Gopal Thakur (IMTech, Chandigarh) are among them. C. G. Suresh (NCL, Pune), V. Radha Kishan (GVK Biosciences, Hyderabad), H. G. Nagendra (Sir M. Visveswariah Institute of Technology, Bengaluru), B. K. Biswal (National Institute of Immunology, New Delhi), Sudarsana Kumar (Mahatma Gandhi University, Kottayam), C. Sadasivan (Kannur Univesity), N. T. Saraswathi (SASTRA University) and R. Sankaranaryanan (Saraswathi Narayanan College, Madurai) were trained in the programme involving water-mediated transformations in protein crystals. Scientists who were trained in other macromolecular crystallography projects at IISc and are now occupying independent research positions include B. Gopal (IISc), Balaji Prakash (Centre for Food Technology Research Institute, Mysuru), R. Natesh (IISER, Thiruvananthapuram), Udupi Ramagopal (Poornaprajna Institute of Scientific Research, Bengaluru), B. V. L. S. Prasad (Amity university, Delhi) and Shankar P. Kanaujia (IIT, Guwahati).

Concluding remarks

Macromolecular crystallography in India, to a substantial extent, radiated from IISc. From humble beginnings, work in the area is now in progress in about 40 institutions across the country and involves nearly 80 research groups. A majority of the group leaders are those trained in Bengaluru, or their academic descendants. Happily, those trained in other schools, including distinguished ones abroad, also joined the community, giving it added vibrancy. The macromolecular crystallography community in India now is several hundred strong, and is reasonably coherent with multi-institutional programmes and common facilities. The role of the programme on the structural biology of plant lectins in the development of macromolecular crystallography in India has been important, especially in the early stages. Indeed, many of the senior leaders of macromolecular crystallography in India were initially trained at Bengaluru in the crystallography of plant lectins.

In addition to its historical role in the development of an area in the country, the work on plant lectins has been important in its scientific quality and impact. It provides a rare example of a well-focused multidisciplinary programme carried out, without break, for a long period of 40 years. It established our strong presence in a global niche area which assumed considerable importance during the period in which we have been engaged with it. In addition to interesting insights pertaining to individual plant lectins, the work at Bengaluru on them led to a body of results of general interest on protein folding, quaternary association and strategies for generating ligand specificities. Furthermore, the plant lectin programme provided a springboard for several other efforts, including those on mycobacterial proteins. Although the work on the structural biology of plant lectins has now been wound up,

as we have now learnt almost everything that we need to know on the general features of the structure and interactions of plant lectins, it has seamlessly led to structural and related studies on mycobacterial and archeal lectins which constitute comparatively unexplored areas.

1. Sharon, N. and Gallagher, J., Glycobiology marches on with a subtle blend of tastes and flavours. *Curr. Opin. Struct. Biol.*, 2007, **17**, 495–497.
2. Arnaud, J., Audfray, A. and Imberty, A., Binding sugars: from natural lectins to synthetic receptors and engineered neolectins. *Chem. Soc. Rev.*, 2013, **7**, 4798–4813.
3. Vijayan, M. and Chandra, N. R., Lectins. *Curr. Opin. Struct. Biol.*, 1999, **9**, 707–714.
4. Pérez, S., Sarkar, A., Rivet, A., Breton, C. and Imberty, A., Glyco3d: a portal for structural glycosciences. *Meth. Mol. Biol.*, 2015, **1273**, 241–258; http://glyco3d.cermav.cnrs.fr.
5. Chandran, T., Sivaji, N., Surolia, A. and Vijayan, M., Ligand binding and retention in snake gourd seed lectin (SGSL). A crystallographic, thermodynamic and molecular dynamics study. *Glycobiology*, 2018, **28**, 968–977.
6. Reeke, G. N., Becker, J. W. and Edelman, G. M., The covalent and three dimensional structure of concanavalin A. IV. Atomic coordinates, hydrogen bonding, and quaternary structure. *J. Biol. Chem.*, 1975, **250**, 1525–1547.
7. Hardman, K. D. and Ainsworth, C. F., Structure of concanavalin A at 2.4-Ang resolution. *Biochemistry*, 1972, **11**, 4910–4919.
8. Wright, C. S., The crystal structure of wheat germ agglutinin at 2.2 Å resolution. *J. Mol. Biol.*, 1977, **111**, 439–457.
9. Salunke, D. M., Khan, M. I., Surolia, A. and Vijayan, M., Crystallization and preliminary X-ray studies of the anti-T lectin from peanut (*Arachis hypogaea*). *J. Mol. Biol.*, 1982, **154**, 177–178.
10. Mande, S. C. *et al.*, Structural studies on peanut lectin. *Indian J. Biochem. Biophys.*, 1988, **25**, 166–171.
11. Salunke, D. M., Khan, M. I., Surolia, A. and Vijayan, M., Preparation and preliminary X-ray studies of three acidic pH crystal forms of the anti-T lectin from peanut (*Arachis hypogaea*). *FEBS Lett.*, 1983, **l56**, 127–129.
12. Dhanaraj, V., Patanjali, S. R., Surolia, A. and Vijayan, M., Preparation and preliminary X-ray studies of two crystal forms of the anti-T lectin from jackfruit (*Artocarpus integrifolia*). *J. Mol. Biol.*, 1988, **203**, 1135–1136.
13. Banerjee, R., Dhanaraj, V., Mahanta, S. K., Surolia, A. and Vijayan, M., Preparation and X-ray characterization of four new crystal forms of jacalin, a lectin from *Artocarpus integrifolia*. *J. Mol. Biol.*, 1991, **221**, 773–776.
14. Sankaranarayanan, R., Puri, K. D., Ganesh, V., Banerjee, R., Surolia, A. and Vijayan, M., Crystallization and preliminary X-ray studies of the basic lectin from winged bean (*Psophocarpus tetragonolobus*). *J. Mol. Biol.*, 1993, **229**, 558–560.
15. Banerjee, R. *et al.*, Crystal structure of peanut lectin, a protein with an unusual quaternary structure. *Proc. Natl. Acad. Sci. USA*, 1994, **91**, 227–231.
16. Sankaranarayanan, R., Sekar, K., Banerjee, R., Sharma, V., Surolia, A. and Vijayan, M., A novel mode of carbohydrate recognition in jacalin, a Moraceae plant lectin with a β-prism fold. *Nature Struct. Biol.*, 1996, **3**, 596–603.
17. Montfort, W. *et al.*, The three-dimensional structure of ricin at 2.8 A. *J. Biol. Chem.*, 1987, **262**, 5398–5403.
18. Hester, G., Kaku, H., Goldstein, I. J. and Wright, C. S., Structure of mannose-specific snowdrop (*Galanthus nivalis*) lectin is representative of a new plant lectin family. *Nature Struct. Biol.*, 1995, **2**, 472–479.
19. Goel, M. *et al.*, Crystal structures of the PNA-porphyrin complex in the presence and absence of lactose: mapping the conformational changes on lactose binding, interacting surfaces, and supramolecular aggregations. *Biochemistry*, 2005, **44**, 5588–5596.
20. Goel, M., Anuradha, P., Kaur, K. J., Maiya, B. G., Swamy, M. J. and Salunke, D. M., Porphyrin binding to jacalin is facilitated by the inherent plasticity of the carbohydrate-binding site: novel mode of lectin–ligand interaction. *Acta Crystallogr., Sect. D*, 2004, **60**, 281–288.
21. Jain, D., Kaur, K. J. and Salunke, D. M., Enhanced binding of a rationally designed peptide ligand of concanavalin A arises from improved geometrical complementarity. *Biochemistry*, 2001, **40**, 12059–12066.
22. Goel, M. *et al.*, Functional equality in the absence of structural similarity: an added dimension to molecular mimicry. *J. Biol. Chem.*, 2001, **276**, 39277–39281.
23. Jain, D., Kaur, K. J. and Salunke, D. M., Plasticity in protein–peptide recognition: crystal structures of two different peptides bound to concanavalin A. *Biophys. J.*, 2001, **80**, 2912–2921.
24. Jain, D., Kaur, K., Sundaravadivel, B. and Salunke, D. M., Structural and functional consequences of peptide–carbohydrate mimicry: crystal structure of a carbohydrate-mimicking peptide bound to concanavalin A. *J. Biol. Chem.*, 2000, **275**, 16098–16102.
25. Mikeska, R. *et al.*, Mistletoe lectin I in complex with galactose and lactose reveals distinct sugar-binding properties. *Acta Crystallogr., Sect. F*, 2005, **1**, 17–25.
26. Mishra, V. *et al.*, Crystal structure of Himalayan mistletoe ribosome-inactivating protein reveals the presence of a natural inhibitor and a new functionally active sugar-binding site. *J. Biol. Chem.*, 2005, **280**, 20712–20721.
27. Mishra, V. *et al.*, Structure of a novel ribosome-inactivating protein from a hemi-parasitic plant inhabiting the northwestern Himalayas. *Acta Crystallogr., Sect. D*, 2004, **60**, 2295–2304.
28. Bagaria, A., Surendranath, K., Ramagopal, U. A., Ramakumar, S. and Karande, A. A., Structure–function analysis and insights into the reduced toxicity of *Abrus precatorius* agglutinin I in relation to abrin. *J. Biol. Chem.*, 2006, **281**, 34465–34474.
29. Sharma, U., Katre, U. V. and Suresh, C. G., Crystal structure of a plant albumin from *Cicer arietinum* (chickpea) possessing hemopexin fold and hemagglutination activity. *Planta*, 2015, **241**, 1061–1073.
30. Rao, K. N., Suresh, C. G., Katre, U. V., Gaikwad, S. M. and Khan, M. I., Two orthorhombic crystal structures of a galactose-specific lectin from *Artocarpus hirsuta* in complex with methyl-α-D-galactose. *Acta Crystallogr., Sect. D*, 2004, **60**, 1404–1412.
31. Surya, S., Geethanandan, K., Sadasivan, C. and Haridas, M., Gallic acid binding to *Spatholobus parviflorus* lectin provides insight to its quaternary structure forming. *Int. J. Biol. Macromol.*, 2016, **91**, 696–702.
32. Surya, S., Abhilash, J., Geethanandan, K., Sadasivan, C. and Haridas, M., A profile of protein–protein interaction: crystal structure of a lectin–lectin complex. *Int. J. Biol. Macromol.*, 2016, **87**, 529–536.
33. Abhilash, J., Geethanandan, K., Bharath, S. R., Sabu, A., Sadasivan, C. and Haridas, M., The crystal structure of a lectin from *Butea monosperma*: insight into its glycosylation and binding of ligands. *Int. J. Biol. Macromol.*, 2015, **72**, 1376–1383.
34. Geethanandan, K., Abhilash, J., Bharath, S. R., Sadasivan, C. and Haridas, M., X-ray structure of a galactose-specific lectin from *Spatholobous parviflorous*. *Int. J. Biol. Macromol.*, 2011, **49**, 992–998.
35. Salunke, D. M., Swamy, M. J., Khan, M. I., Mande, S. C., Surolia, A. and Vijayan, M., Arrangement of subunits in peanut lectin. Rotation function and chemical cross-linking studies. *J. Biol. Chem.*, 1985, **260**, 13576–13579.
36. Banerjee, R., Das, K., Ravishankar, R., Suguna, K., Surolia, A. and Vijayan, M., Conformation, protein–carbohydrate interactions and a novel subunit association in the refined structure of peanut lectin–lactose complex. *J. Mol. Biol.*, 1996, **259**, 281–296.

37. Sharma, V., Vijayan, M. and Surolia, A., Imparting exquisite specificity to peanut agglutinin for the tumor-associated Thomsen–Friedenreich antigen by redesign of its combining site. *J. Biol. Chem.*, 1996, **271**, 21209–21213.
38. Ravishankar, R., Ravindran, M., Suguna, K., Surolia, A. and Vijayan, M., Crystal structure of the peanut lectin–T–antigen complex. Carbohydrate specificity generated by water bridges. *Curr. Sci.*, 1997, **72**, 855–861.
39. Sharma, V., Srinivas, V. R., Adhikari, P., Vijayan, M. and Surolia, A., Molecular basis of recognition by Gal/GalNAc specific legume lectins: influence of Glu 129 on the specificity of peanut agglutinin (PNA) towards C2-substituents of galactose. *Glycobiology*, 1998, **10**, 1007–1012.
40. Ravishankar, R., Surolia, A., Vijayan, M., Lim, S. and Kishi, Y., Preferred conformation of C-lactose at the free and peanut lectin bound states. *J. Am. Chem. Soc.*, 1998, **120**, 11297–11303.
41. Ravishankar, R., Suguna, K., Surolia, A. and Vijayan, M., Structures of the complexes of peanut lectin with methyl-β-galactose and *N*-acetyllactosamine and a comparative study of carbohydrate binding in Gal/GalNAc-specific legume lectins. *Acta Crystallogr., Sect. D*, 1999, **55**, 1375–1382.
42. Ravishankar, R., Thomas, C. J., Suguna, K., Surolia, A. and Vijayan, M., Crystal structures of the peanut lectin-lactose complex at acidic pH: retention of unusual quaternary structure, empty and carbohydrate bound combining sites, molecular mimicry and crystal packing directed by interactions at the combining site. *Proteins Struct. Funct. Genet.*, 2001, **43**, 260–270.
43. Adhikari, P. *et al.*, Mutational analysis at Asn-41 in peanut agglutinin. A residue critical for the binding of the tumor-associated Thomsen-Friedenreich antigen. *J. Biol. Chem.*, 2001, **276**, 40734–40739.
44. Pratap, J. V. *et al.*, The combination of molecular dynamics with crystallography for elucidating protein–ligand interactions: a case study involving peanut lectin complexes with T-antigen and lactose. *Acta Crystallogr., Sect. D,* 2001, **57**, 1584–1594.
45. Natchiar, S. K. *et al.*, Structural plasticity of peanut lectin: an X-ray analysis involving variation in pH, ligand binding and crystal structure. *Acta Crystallogr., Sect. D*, 2004, **60**, 211–219.
46. Natchiar, S. K. *et al.*, Multivalency in lectins – a crystallographic, modelling and light-scattering study involving peanut lectin and a bivalent ligand. *Curr. Sci.*, 2006, **90**, 1230–1237.
47. Natchiar, S. K., Srinivas, O., Mitra, N., Surolia, A., Jayaraman, N. and Vijayan, M., Structural studies on peanut lectin complexed with disaccharides involving different linkages: further insights into the structure and interactions of the lectin. *Acta Crystallogr., Sect. D*, 2006, **62**, 1413–1421.
48. Natchiar, S. K., Suguna, K., Surolia, A. and Vijayan, M., Peanut agglutinin, a lectin with an unusual quaternary structure and interesting ligand binding properties. *Crystallogr. Rev.*, 2007, **13**, 1–26.
49. Prabu, M. M. *et al.*, Carbohydrate specificity and quaternary association in basic winged bean lectin: X-ray analysis of the lectin at 2.5 Å resolution. *J. Mol. Biol.*, 1998, **276**, 787–796.
50. Kulkarni, K. A., Sinha, S., Katiyar, S., Surolia, A., Vijayan, M. and Suguna, K., Structural basis for the specificity of basic winged bean lectin for the Tn-antigen: a crystallographic, thermodynamic and modelling study. *FEBS Lett.*, 2005, **579**, 6775–6780.
51. Kulkarni, K. A., Katiyar, S., Surolia, A., Vijayan, M. and Suguna, K., Structural basis for the carbohydrate-specificity of basic winged-bean lectin and its differential affinity for Gal and GalNAc. *Acta Crystallogr., Sect. D*, 2006, **62**, 1319–1324.
52. Kulkarni, K. A., Katiyar, S., Surolia, A., Vijayan, M. and Suguna, K., Generation of blood group specificity: new insights from structural studies on the complexes of A- and B-reactive saccharides with basic winged bean agglutinin. *Proteins Struct. Funct. Genet.*, 2007, **68**, 762–769.
53. Kulkarni, K. A., Katiyar, S., Surolia, A., Vijayan, M. and Suguna, K., Structure and sugar-specificity of basic winged-bean lectin: structures of new disaccharide complexes and a comparative study with other known disaccharide complexes of the lectin. *Acta Crystallogr., Sect. D*, 2008, **64**, 730–737.
54. Manoj, N., Srinivas, V. R., Surolia, A., Vijayan, M. and Suguna, K., Carbohydrate specificity and salt-bridge mediated conformational change in acidic winged bean agglutinin. *J. Mol. Biol.*, 2000, **302**, 1129–1137.
55. Jeyaprakash, A. A., Srivastav, A., Surolia, A. and Vijayan, M., Structural basis for the carbohydrate specificities of artocarpin: variation in the length of a loop as a strategy for generating ligand specificity. *J. Mol. Biol.*, 2004, **338**, 757–770.
56. Thamotharan, S. *et al.*, Modification of the sugar specificity of a plant lectin: structural studies on a point mutant of *Erythrina corallodendron* lectin. *Acta Crystallogr., Sect. D*, 2011, **67**, 218–227.
57. Shetty, K. N., Latha, V. L., Rao, R. N., Nadimpalli, S. K. and Suguna, K., Affinity of a galactose-specific legume lectin from *Dolichos lablab* to adenine revealed by X-ray crystallography. *IUBMB Life*, 2013, **65**, 633–644.
58. Jeyaprakash, A. A. *et al.*, Crystal structure of the jacalin–T-antigen complex and a comparative study of lectin–T-antigen complexes. *J. Mol. Biol.*, 2002, **321**, 637–645.
59. Jeyaprakash, A. A., Katiyar, S., Swaminathan, C. P., Sekar, K., Surolia, A. and Vijayan, M., Structural basis of the carbohydrate specificities of jacalin: an X-ray and modeling study. *J. Mol. Biol.*, 2003, **332**, 217–228.
60. Jeyaprakash, A. A. *et al.*, Structural basis for the energetics of jacalin–sugar interactions: promiscuity versus specificity. *J. Mol. Biol.*, 2005, **347**, 181–188.
61. Sharma, A., Sekar, K. and Vijayan, M., Structure, dynamics, and interactions of jacalin. Insights from molecular dynamics simulations examined in conjunction with results of X-ray studies. *Proteins Struct. Funct. Bioinformat.*, 2009, **77**, 760–777.
62. Abhinav, K. V., Sharma, K., Swaminathan, C. P., Surolia, A. and Vijayan, M., Jacalin-carbohydrate interactions: distortion of the ligand molecule as a determinant of affinity. *Acta Crystallogr., Sect. D*, 2015, **71**, 324–331.
63. Abhinav, K. V., Sharma, K., Surolia, A. and Vijayan, M., Effect of linkage on the location of reducing and nonreducing sugars bound to jacalin. *IUBMB Life*, 2016, **68**, 971–979.
64. Abhinav, K. V., Sharma, K., Surolia, A. and Vijayan, M., Distortion of the ligand molecule as a strategy for modulating binding affinity: further studies involving complexes of jacalin with β-substituted disaccharides. *IUBMB Life*, 2017, **69**, 72–78.
65. Pratap, J. V., Jeyaprakash, A. A., Rani, P. G., Sekar, K., Surolia, A. and Vijayan, M., Crystal structures of artocarpin, a Moraceae lectin with mannose specificity, and its complex with methyl-α-D-mannose: implications to the generation of carbohydrate specificity. *J. Mol. Biol.*, 2002, **317**, 237–247.
66. Jeyaprakash, A. A., Srivastav, A., Surolia, A. and Vijayan, M., Structural basis for the carbohydrate specificities of artocarpin: variation in the length of a loop as a strategy for generating ligand specificity. *J. Mol. Biol.*, 2004, **338**, 757–770.
67. Singh, D. D., Saikrishnan, K., Kumar, P., Surolia, A., Sekar, K. and Vijayan, M., Unusual sugar specificity of banana lectin from *Musa paradisiaca* and its probable evolutionary origin. Crystallographic and modelling studies. *Glycobiology*, 2005, **15**, 1025–1032.
68. Sharma, A. and Vijayan, M., Influence of glycosidic linkage on the nature of carbohydrate binding in β-prism I fold lectins: an X-ray and molecular dynamics investigation on banana lectin-carbohydrate complexes. *Glycobiology*, 2011, **21**, 23–33.
69. Chandra, N. R., Ramachandraiah, G., Bachhawat, K., Dam, T. K., Surolia, A. and Vijayan, M., Crystal structure of a dimeric mannose-specific agglutinin from garlic: quaternary association and carbohydrate specificity. *J. Mol. Biol.*, 1999, **285**, 1157–1168.
70. Ramachandraiah, G., Chandra, N. R., Surolia, A. and Vijayan, M., Re-refinement using reprocessed data to improve the quality of the

structure: a case study involving garlic lectin. *Acta Crystallogr., Sect. D*, 2002, **58**, 414–420.
71. Ramachandraiah, G., Chandra, N. R., Surolia, A. and Vijayan, M., Computational analysis of multivalency in lectins: structures of garlic lectin–oligosaccharide complexes and their aggregates. *Glycobiology*, 2003, **13**, 765–775.
72. Shetty, K. N., Bhat, G. G., Inamdar, S. R., Swamy, B. M. and Suguna, K., Crystal structure of a-prism II lectin from *Remusatia vivipara*. *Glycobiology*, 2012, **22**, 56–69.
73. Sharma, A. *et al.*, The sequence and structure of snake gourd (*Trichosanthes anguina*) seed lectin, a three-chain nontoxic homologue of type II RIPs. *Acta Crystallogr., Sect. D*, 2013, **69**, 1493–1503.
74. Chandran, T., Sharma, A. and Vijayan, M., Structural studies on a non-toxic homologue of type II RIPs from bitter gourd: molecular basis of non-toxicity, conformational selection and glycan structure. *J. Biosci.*, 2015, **40**, 929–941.
75. Sinha, S., Gupta, G., Vijayan, M. and Surolia, A., Subunit assembly of plant lectins. *Curr. Opin. Struct. Biol.*, 2007, **17**, 498–505.
76. Abhinav, K. V. and Vijayan, M., Structural diversity and ligand specificity of lectins. The Bangalore effort. *J. Macromol. Sci., Part A*, 2014, **86**, 1335–1355.
77. Chandra, N. R., Prabu, M. M., Suguna, K. and Vijayan, M., Structural similarity and functional diversity in proteins containing the legume lectin fold. *Protein Eng.*, 2001, **14**, 857–866.
78. Chandran, T., Sharma, A. and Vijayan, M., Generation of ligand specificity and modes of oligomerization in beta-prism I fold lectins. *Adv. Proteins Chem. Struct. Biol.*, 2013, **92**, 135–178.
79. Prabu, M. M., Suguna, K. and Vijayan, M., Variability in quaternary association of proteins with the same tertiary fold: a case study and rationalization involving legume lectins. *Proteins Struct. Funct. Genet.*, 1999, **35**, 58–69.
80. Delbaere, L. T. J. *et al.*, Molecular recognition of a human blood group determinant by a plant lectin. *Can. J. Chem.*, 1990, **68**, 1116–1121.
81. Shaanan, B., Lis, H. and Sharon, N., Structure of a legume lectin with an ordered *N*-linked carbohydrate in complex with lactose. *Science*, 1991, **254**, 862–866.
82. Schwarz, F. P., Puri, K. and Surolia, A., Thermodynamics of the binding of galactopyranoside derivatives to the basic lectin from winged bean (*Psophocarpus tetrogonolobus*). *J. Biol. Chem.*, 1991, **266**, 24344–24350.
83. Surolia, A., Sharon, N. and Schwarz, F. P., Thermodynamics of monosaccharide and disaccharide binding to *Erythrina corallodendron* lectin. *J. Biol. Chem.*, 1996, **27**, 17697–17703.
84. Sharma, A. and Vijayan, M., Quaternary association in β-prism I fold plant lectins: insights from X-ray crystallography, modelling and molecular dynamics. *J. Biosci.*, 2011, **36**, 793–808.
85. Liu, W., Yang, N., Ding, J., Huang, R. H., Hu, Z. and Wang, D. C., Structural mechanism governing the quaternary organization of monocot mannose-binding lectin revealed by the novel monomeric structure of an orchid lectin. *J. Biol. Chem.*, 2005, **280**, 14865–14876.
86. Sharma, V. and Surolia, A., Analyses of carbohydrate recognition by legume lectins: size of the combining site loops and their primary specificity. *J. Mol. Biol.*, 1997, **267**, 433–445.
87. Sharma, A., Chandran, D., Singh, D. D. and Vijayan, M., Multiplicity of carbohydrate-binding sites in beta-prism fold lectins: occurrence and possible evolutionary implications. *J. Biosci.*, 2007, **32**, 1089–1110.
88. Abhinav, K. V., Sharma, A. and Vijayan, M., Identification of mycobacterial lectins from genomic data. *Proteins Struct. Funct. Bioinformat.*, 2013, **81**, 644–657.
89. Patra, D., Srikalaivani, R., Misra, A., Singh, D. D., Selvaraj, M. and Vijayan, M., Cloning, expression, purification, crystallization and preliminary X-ray studies of a secreted lectin (Rv1419) from *Mycobacterium tuberculosis. Acta Crystallogr., Sect. F*, 2010, **66**, 1662–1665.
90. Patra, D., Sharma, A., Chandran, D. and Vijayan, M., Cloning, expression, purification, crystallization and preliminary X-ray studies of the mannose-binding lectin domain of MSMEG-3662 from *Mycobacterium smegmatis*. *Acta Crystallogr., Sect. F*, 2011, **67**, 596–599.
91. Patra, D., Mishra, P., Vijayan, M. and Surolia, A., Negative cooperativity and high affinity in chitooligosaccharide binding by a *Mycobacterium smegmatis* protein containing LysM and Lectin domains. *Biochemistry*, 2016, **55**, 49–61.
92. Patra, D., Mishra, P., Surolia, A. and Vijayan, M., Structure, interactions and evolutionary implications of a domain-swapped lectin dimer from *Mycobacterium smegmatis*. *Glycobiology*, 2014, **24**, 956–965.
93. Abhinav, K. V., Samuel, E. and Vijayan, M., Archeal lectins: an identification through a genomic search. *Proteins Struct. Funct. Bioinformat.*, 2016, **84**, 21–30.
94. Sivaji, N., Abhinav, K. V. and Vijayan, M., Crystallization and biochemical characterization of an archaeal lectin from *Methanococcus voltae* A3. *Acta Crystallogr., Sect. F*, 2017, **F73**, 300–304.

ACKNOWLEDGEMENTS. Scores of scientists have worked in the lectin and other programmes mentioned in this article. The contributions of some of the structural biologists who are currently working in India have been referred to in the text. Others, mostly now abroad, who contributed to the efforts on lectins include S. Raghunandan, Kalyan Das, late V. Dhanaraj, A. Stephen Suresh, A. A. Jeyaprakash, S. Kundhavai Natchiar, Moses M. Prabu, Karthika. N. Shetty, Alok Sharma, D. Patra, K. V. Abhinav and Thyageswar Chandran. The programme was made possible on account of strong biochemical support. Late Mohammed Islam Khan, M. J. Swamy, S. R. Patanjali, V. Anantharam, M. V. Krishna Sastry, S. K. Mahanta, K. D. Puri, Shreeta Acharya, Vivek Sharma, Mili Kapoor, Sandra Misquith, Kiran Bachhawat, C. P. Swaminathan, V. Bhanuprakash Reddy, Anita Patil, Dipti Gupta and Padmanabh Mishra, are prominent among those who made significant biochemical contributions for the success of the structural biological studies on lectins at Bengaluru. The lectin programme and the X-ray facility used for pursuing it have been continuously supported by the Department of Science and Technology, Government of India.

Received 16 January 2019; accepted 7 February 2019

doi: 10.18520/cs/v116/i9/1490-1505