2.1. INTRODUCTION

Lectins occurring in plants differ in their molecular structure, carbohydrate binding property and biochemical properties, hence are considered as complex and heterogeneous group of proteins (Van Damme et al. 1998a). Although vast majority of reported lectins are from seeds, equal number are known from every vegetative tissues, such as leaves, bark, stems, rhizomes, bulbs, and tubers (Etzler 1985; Peumans and Van Damme 1998 and 1999). In seeds, lectins constitute about 1-10 %, sometimes even higher (up to 50 %) part of the total seed proteins. Also in vegetative tissues (for example, bark, bulbs, tubers, rhizomes and corms) lectins constitute 1-20 % of total proteins (Peumans and Van Damme, 1998). These lectins both in seeds and other vegetative tissues are often developmentally regulated and exhibit seasonal variations in their concentrations (Peumans and Van Damme 1995). Considering the criterion of carbohydrate binding specificity, plant lectins are distinguished as mannose, mannose/glucose, mannose/maltose, Gal/GalNAc-, GlcNAc/(GlcNAc), fucose, and sialic acid binding lectins (Goldstein and Poretz 1986; Van Damme et al. 1998b). However, such a classification based on carbohydrate binding specificity was inappropriate with respect to the evolutionary relationships amongst this heterogeneous group of proteins. In addition, different carbohydrate binding motifs can recognize similar sugar structures.

Based on the sequence data obtained through molecular cloning and protein sequencing, plant lectins reported so far are classified into seven major
families of evolutionary and structurally related proteins; 1) The legume lectins, 2) The chitin binding lectins, 3) The type-2 ribosome inactivating proteins (RIP-2), 4) Monocot mannose-binding lectins (MMBL), 5) Jacalin related lectins, 6) Amaranthins and 7) The Cucurbitaceae phloem lectins. This classification of plant lectins proposed by comparison of sequence data is well accepted presently (Vandenborre et al. 2011).

In the recent past monocot mannose binding lectins have received greater attention, because of their interesting structural and biological properties (Van Damme et al. 1995; Barre et al. 1996). Galanthus nivalis agglutinin (GNA) was the first monocot mannose specific lectin reported, purified from the bulbs of snowdrop (Van Damme et al. 1987). Since then several structurally and evolutionarily related MMBLs were reported from different monocot families, viz., Amaryllidaceae (Van Damme et al. 1987; Kaku et al. 1990), Alliaceae (Van Damme et al. 1993), Araceae (Van Damme, et al. 1993; Mo et al. 1999), Orchidaceae (Van Damme et al. 1994), and Liliaceae (Van Damme et al. 1996). These lectins contain a conserved motif of five amino acids, Q-x-D-x-N-x-V-x-Y in their carbohydrate binding site and exhibit exclusive specificity for mannose, hence they are referred to as monocot mannose binding lectins (MMBL). In the recent they are also referred as GNA related lectins (Van Damme et al. 2007). Many of them exhibit weak affinity towards mannose, but strong affinity towards oligomannosides and high mannose containing N-glycans (Van Damme 2008). MMBL reported so far consist of highly homologous sequences with similar tertiary structures.
suggesting these proteins comprise a single super family of mannose binding proteins (Barre et al. 1996). Although, all monocot mannose binding lectins are very similar at the protein level, there are important differences in the processing and post translational modifications. In general they are synthesized as single polypeptide precursors and are post translationally cleaved into polypeptides of nearly equal size (Van Damme et al. 2007).

The present chapter describes the purification and physicochemical characterization of a mannose binding lectin from the tubers of edible epiphyte *Remusatia vivipara*, a member of monocot family. Also, the molecular cloning of the lectin gene, binding site analysis and multiple comparison of the sequence with other closely related MMBLs are presented. Finally, this chapter describes the protocol developed for the crystallization of the lectin and is the first lectin to be crystallized from the Araceae family, giving an opportunity to determine the 3D structure by X-ray crystallographic analyses.

### 2.2. MATERIALS

#### 2.2.1. Materials

The tubers of *R. vivipara* were collected from the forest around Kumta, Uttara Kannada District, Karnataka (Western Ghat region of Southern India) during October-December. Tubers were cleaned under running water, briefly air dried and stored in airtight polythene bags at -20 °C till further use.
Mucin (porcine stomach, type III), fetuin (from fetal calf serum), thyroglobulin (bovine thyroid) and various fine sugars used in hapten inhibition studies were from Sigma Chemical Co., St. Louis, USA. Sepharose 4B was obtained from Pharmacia Fine Chemicals, Uppsala, Sweden. Acrylamide and N’, N-methylene diacrylamide were from Koch-Light Laboratories, England and N, N, N’, N’-tetramethylene diamine was obtained from BDH Chemicals Ltd., Poole, England. Asialofetuin-Sepharose4B, affinity matrix used for the purification of the lectin was prepared by coupling asialofetuin to cyanogen bromide activated Sepharose 4B as described by March et al. (1974). Standard molecular weight marker proteins (Medium range) was procured from Sigma chemical Co. St. Louis, USA. Rabbit blood was collected from the healthy rabbits maintained in the local animal house. All other chemicals used were of analytical reagent grade and the reagents were prepared in double distilled water.
2.3. METHODS

2.3.1. Isolation and purification of *Remusatia vivipara* lectin (RVL)

Stored frozen tubers of *R. vivipara* after thawing (100 g) were cut into small pieces, homogenized in a waring blender and the lectin was extracted with 2000 ml 50 mM sodium phosphate buffer containing 154 mM NaCl, pH, 7.2 (PBS), by stirring magnetically overnight at 4 °C. The extract was filtered by passing through muslin cloth and the filtrate was clarified by centrifugation (4000 g for 20 min). Resulting clear supernatant was heated to 50 °C in a water bath for 20 min, cooled and the precipitate formed was removed by centrifugation (4500 g for 20 min). Clear supernatant was extensively dialyzed against PBS, subsequently with water, further the dialyzate was lyophilized. This crude lectin preparation was stored at 4°C for further purification.

2.3.2. Purification of RVL by affinity chromatography using asialofetuin-Sepharose 4B column

Freeze dried crude lectin powder (30 mg), suspended in 5 ml of PBS was passed through asialofetuin-Sepharose 4B affinity column (1.5 x 10 cm) equilibrated in PBS, at a flow rate of 15 ml/h. 3.0 ml fractions eluting from the column were collected on fraction collector (FRAC-100; Pharmacia). Unbound proteins from the column were eluted by washing the column with PBS until the absorbance of the eluting fractions read zero at 280 nm. Affinity bound lectin was eluted using 0.1M glycine-HCl buffer containing 500 mM NaCl, pH, 2.0. The column operations were carried out in a cold room at 4 °C. Lectin activity in the eluting fractions was determined by hemagglutination assay using trypsinized rabbit erythrocytes and the lectin peak fractions were pooled,
dialyzed against PBS followed by water and freeze dried. Homogeneity of the affinity purified lectin preparation was confirmed by SDS-PAGE on 15 % gel according to Laemmli et al. (1970).

2.3.3. Hemagglutinating activity and carbohydrate specificity

Hemagglutinating activity of the lectin was routinely determined at various stages of purification by two fold serial dilution technique in 96 well microtitre assay plates using trypsinized rabbit erythrocytes. The highest dilution of the extract causing visible agglutination was arbitrarily considered as the "titre" and the minimum concentration of the protein required for agglutination (MCA) as “one unit of hemagglutinating activity”. The specific hemagglutinating activity is expressed as activity unit mg\(^{-1}\) protein. Carbohydrate binding specificity of the purified lectin was determined by hapten inhibition assay.

Inhibition assays were carried out by incubating the lectin sample (diluted to give titre: 8) with serially diluted sugar (200 mM) /glycoprotein (1mg/ml) in a total volume of 50 µl, prior to the addition of erythrocytes and the hemagglutination was visually observed. The lowest concentration of the sugar/glycoprotein, which inhibited the agglutination, was taken as the inhibitory titre of the hapten.

2.3.4. Protein estimation

Protein concentration was determined by the method of Lowry O. H. et al. (1951), using bovine serum albumin as standard.
2.3.5. Sugar estimation

Total sugar content of glycoproteins was estimated by phenol-sulfuric acid method (Dubois et al. 1956) using glucose as standard sugar. In brief, to 0.5 ml of sugar solution (5-50 µg), 0.5 ml of 5% phenol solution was added and mixed. Then 2.5 ml of conc. H$_2$SO$_4$ was added directly using a plunger. The contents were vortexed, allowed to stand for 20 min at room temperature and the absorbance was measured at 490 nm against a suitable blank. Sugar concentration was estimated from a calibration curve prepared using glucose as standard sugar.

2.3.6. Determination of molecular weight by gel filtration chromatography

The native molecular mass of purified RVL was estimated by gel filtration chromatography on a Superdex G-75 (Amersham Biosciences) column (80 x 1.5 cm) of equilibrated in 25 mM Tris-HCl containing 154 mM NaCl, pH 7.2. Fractions of 3.0 ml were collected at a flow rate of 18 ml/h, and the elution of the protein was monitored by measuring the absorbance at 280 nm. The column was earlier calibrated using standard molecular weight marker proteins viz., lysozyme (14.3 kDa), lactoglobulin (18.4 kDa), chymotrypsinogen (25.6 kDa), pepsin (34.7 kDa), ovalbumin (45.0 kDa) and bovine serum albumin (BSA, 66.0 kDa). A calibration curve was obtained by plotting the K-average against log$_{10}$ molecular weights.
2.3.7. SDS-PAGE

Purified RVL was subjected to sodium dodecyl sulphate polyacrylamide gel electrophoresis in 15 % (w/v) acrylamide gel of 0.75 mm thickness using Mighty Small II SE 250 (Hoefer Scientific Instruments, San Francisco, USA) unit according to the method described by Laemmli et al. (1970), for determining homogeneity and the subunit molecular mass of the lectin. Subunit molecular mass of RVL was estimated from the calibration curve obtained using standard protein markers.

2.3.8. MALDI-TOF-MS

Molecular weight of the purified RVL was determined by MALDI-TOF-MS analysis on UltraflexTOF/TOF (Bruker Daltonics, Germany) mass spectrometer, equipped with a UV nitrogen laser of 337 nm. Purified lectin (1 µl) in water (0.5 µg/µl) mixed with equal volume of matrix solution (saturated solution of sinapinic acid in acetonitrile in 0.1 % aqueous trifluoroacetic acid) was deposited on the probe plate. The mass spectra were recorded in the reflectron positive ion mode after evaporation of the solvent and spectrum was acquired and analyzed by Bruker Daltonics FLEX control software.

2.3.9. Effect of temperature and pH

Thermal stability of the purified RVL was evaluated by determining the residual activity of the lectin after heat treatment at different temperatures. Purified RVL in PBS (50 µg/ml) was heated on a water bath maintained at different temperatures (20-80 °C) for 20 min. At each temperature an aliquot was drawn and cooled to room temperature and the residual lectin activity was assayed by hemagglutination assay as described earlier.
Stability of the purified lectin at different pH was determined by the hemagglutination activity of the lectin samples incubated in buffers of different pH, 2.0 (100 mM glycine-HCl buffer), pH, 4.3 (25 mM sodium acetate buffer), pH, 7.2 (50 mM sodium phosphate buffer) and pH, 8.5 (50 mM Tris-HCl 25 mM) and pH, 9.3 (50 mM sodium carbonate buffer). Incubation was carried out overnight at 4°C and the pH of the samples was adjusted to pH, 7.2 by adding 0.1 N NaOH before determining the lectin activity by hemagglutination assay.

2.3.10. Determination of N-terminal sequence

N-terminal sequence analysis of the purified RVL was determined from the PVDF electro blots (semi dry elecroblotting at 75 mA for 3h) after SDS-PAGE in 15 % gel. The sequence was determined on automated Edman degradation protein sequencer, PROCISE Protein Sequencing System (Applied Biosystems). The Procise Protein Sequencing System sequentially cleaves N-terminal amino acids from protein/peptide chains and analyzes the resulting phenylthiohydantoin (PTH) amino acid residues.

2.3.11. Dynamic light scattering (DLS)

In order to investigate the oligomeric properties of the purified RVL, dynamic light scattering (DLS) studies were carried out on DYNAPRO instrument (USA). Polydispersity (pd) and the hydrodynamic radius (R_H) of the molecule were estimated by DLS. The data was collected for purified RVL (0.1 mg/ml) suspended in water at room temperature and the correlation function was analyzed using Dynamics V6 software.
2.3.12. Cloning full length cDNA of RVL

Total mRNA from the leaves of *R. Vivipara* was extracted using RNA isolation kit (Fermentas, USA) as per the description provided by the manufacturer. PolyA+ mRNA (0.4 ng/µl) was used to synthesize the first strand cDNA (3’-ready cDNA) using SMART™ RACE cDNA amplification kit (Conetech), and 3’ RACE PCR was carried out using forward primer RBP44_RACE_F:5’-ATGCAG(C/G/T)A(G/C)GAC(T/G)GCAACCTG-3’ designed based on the conserved regions of previously reported mannose-binding monocot lectins. The amplified product was purified and cloned into pTZ57R/T, T/A cloning vector (InsT/A clone™, PCR product cloning kit from MBI Fermentas, USA), and transformed into *Escherichia coli* DH5α. The positive clones were picked up and sequenced. After the synthesis of first strand cDNA (5’-readycDNA), 5’ RACE PCR was carried out using a reverse primer RBP44_RVL_RACE_R:5’-GGGCCCAAGACGACGACGACGACCAGC-3’. The product was purified and cloned into the pTZ57R/T. Full-length cDNA sequence of RVL was deduced by assembling the sequences of the 3’ RACE and 5’ RACE products. Full length sequence of RVL was amplified using specific forward primer RVL_F:5’TTGGCCATGGCCAAGCTCCTC3’ and reverse primer RVL_R:5’GGCGAATTCTACGACGAGCAAA3’, designed containing restriction enzyme site (NcoI and EcoRI). Complete coding sequence of RVL was subsequently amplified by genomic DNA also using same specific primers. PCR amplification was performed under following conditions: cDNA denatured at 94 °C for 5 min, followed by 30 cycles of amplification (94 °C for 1 min, 57 °C for 30 s, and 72 °C for 2 min) and
The same program was used for performing 3' and 5' RACE and also for genomic DNA amplification mentioned earlier. The amplified full-length sequence of RVL was cloned into pTZ57R/T, and transferred to *E. coli* DH5α. The clone was confirmed by restriction digestion, RVL-specific PCR and sequencing.

2.3.13. **Crystallization of *Remusatia vivipara* lectin**

Crystallization conditions for purified RVL were standardized using crystallization kits from Hampton Research and Jena Biosciences. These conditions were screened using the sitting-drop method under oil, at 289 K, where the 2 µl drop typically contained protein at 10 mg/ml concentration and 2 µl of the precipitating reagent. Several crystallization conditions yielded poor quality thin needle shaped crystals having length less than 0.05 mm. The condition optimized with 30 % polyethylene glycol 4K in 100 mM sodium citrate buffer pH, 5.6 and 200 mM ammonium acetate yielded rod-shaped crystals of good diffraction quality.

2.4. **RESULTS**

2.4.1. **Isolation and purification of mannose lectin from the tubers of *Remusatia vivipara***

Crude extracts of *R. vivipara* tubers showed 100 fold higher hemagglutination activity towards trypsinized rabbit erythrocytes as compared to untrypsinized cells. However no activity was found with either trypsinized or untrypsinized human erythrocytes of A, B and O groups.
Preliminary fractionation methods using ammonium sulphate precipitation and methanol precipitation to fractionate the lectin and to eliminate the associated polysaccharides and other interfering substances were unsuccessful. However the heat treatment of the crude extract at 50 °C for 20 min not only reduced the viscosity of the extract, but also effectively removed 7.63 % of unwanted proteins while retaining the 99.95 % lectin activity resulting in overall 1.08 fold purification of the lectin (Table 2.1).

**Table 2.1.** Fold purification of RVL

<table>
<thead>
<tr>
<th>Purification steps of RVL</th>
<th>Hemagglutination Activity</th>
<th>Total volume of buffer in ml</th>
<th>Protein concentration in mg/ml</th>
<th>Total protein in mg</th>
<th>Total activity</th>
<th>Specific Activity/mg of lectin</th>
<th>Fold purification</th>
<th>Recovery of RVL (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>RVL Crude Extract</td>
<td>256 (1:100 dilution)</td>
<td>2000ml</td>
<td>1.5mg/ml</td>
<td>3000.0mg</td>
<td>10.23 X10³</td>
<td>3.41X10³</td>
<td>1.0</td>
<td>100%</td>
</tr>
<tr>
<td>After heat treatment of RVL</td>
<td>256 (1:100 dilution)</td>
<td>2000ml</td>
<td>1.38mg/ml</td>
<td>2771.01mg</td>
<td>10.225X10³</td>
<td>3.69X10³</td>
<td>1.08</td>
<td>99.95%</td>
</tr>
<tr>
<td>RVL after affinity purification</td>
<td>512 (1:250 dilution)</td>
<td>400ml</td>
<td>1.01mg/ml</td>
<td>401.0mg</td>
<td>10.165X10³</td>
<td>25.35X10³</td>
<td>7.43</td>
<td>99.36%</td>
</tr>
</tbody>
</table>

*This table represents the purification of RVL starting from 100g of tuber; results are the average of three sets.

Final purification of the lectin was achieved by single step affinity chromatography, on the asialofetuin-Sepharose 4B column. The elution profile of the lectin on the affinity column is presented in Fig. 2.2. Unbound proteins (79 %) were removed by washing with PBS and the affinity bound lectin was eluted using 100 mM glycine-HCl buffer pH 2.0 containing 500 mM NaCl. Affinity purification step resulted in 7.43 fold purification with an excellent
recovery of 99.36 % activity. This protocol gave 401 mg of the purified lectin starting from 100 gm of the tubers. The homogeneity of the purified lectin was established by SDS-PAGE in 15 % gel and the lectin was homogeneous, showed single band (Fig. 2.3).

Fig. 2.2. Purification of *Remusatia vivipara* lectin by affinity chromatography using asialofetuin Sepharose-4B column (1.5x 10 cm), equilibrated in PBS and the affinity bound lectin was eluted with 100 mM Glycine-HCl buffer, pH 2.0 containing 500 mM NaCl. Fractions of 3.0 ml were collected at a flow rate of 15 ml /h. ●● Absorbance at 280 nm; ▲▲ Hemagglutinating activity.

Fig. 2.3. SDS-PAGE of affinity purified RVL in 15% gel. Lane 1 and lane 2 contain standard molecular weight markers and purified RVL (20 µg) respectively.
2.4.2. Hapten inhibition study:

Carbohydrate specificity of the purified lectin was determined by hapten inhibition assays using simple sugars, sugar derivatives and glycoproteins. The results of the hapten inhibition studies are presented in Table 2.2. None of the simple sugars or the sugar derivatives inhibited the hemagglutination activity of the lectin. However the hemagglutinating activity of RVL was strongly inhibited by mucin, asialomucin, asialofetuin and thyroglobulin, but not by fetuin.

Table 2.2. Hapten inhibition studies of RVL

<table>
<thead>
<tr>
<th>Haptens tested for Inhibition</th>
<th>Minimum concentration required for inhibition in µg (MIC)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Simple sugars tested (200 mM):</td>
<td></td>
</tr>
<tr>
<td>1. D-Glucose</td>
<td>NI</td>
</tr>
<tr>
<td>2. D-Galactose</td>
<td>NI</td>
</tr>
<tr>
<td>3. D-Mannose</td>
<td>NI</td>
</tr>
<tr>
<td>4. L-Fucose</td>
<td>NI</td>
</tr>
<tr>
<td>5. D-Lactose</td>
<td>NI</td>
</tr>
<tr>
<td>Sugar derivatives tested (200 mM):</td>
<td></td>
</tr>
<tr>
<td>1. Glucosamine</td>
<td>NI</td>
</tr>
<tr>
<td>2. N-Acetyl D-galactosamine</td>
<td>NI</td>
</tr>
<tr>
<td>3. N-Acetyl D-mannosamine</td>
<td>NI</td>
</tr>
<tr>
<td>4. Sialic Acid</td>
<td>NI</td>
</tr>
<tr>
<td>5. Methyl α-D-Manno Pyranoside</td>
<td>NI</td>
</tr>
<tr>
<td>6. Gum Guar</td>
<td>NI</td>
</tr>
<tr>
<td>Glycoproteins tested (1 mg/ml):</td>
<td></td>
</tr>
<tr>
<td>1. Transferrin</td>
<td>NI</td>
</tr>
<tr>
<td>2. Ovalbumin</td>
<td>NI</td>
</tr>
<tr>
<td>3. Quail Albumin</td>
<td>NI</td>
</tr>
<tr>
<td>4. Fetuin</td>
<td>NI</td>
</tr>
<tr>
<td>5. Asialofetuin</td>
<td>3.125</td>
</tr>
<tr>
<td>6. Mucin</td>
<td>1.56</td>
</tr>
<tr>
<td>7. Asialomucin</td>
<td>3.125</td>
</tr>
<tr>
<td>8. Thyroglobulin</td>
<td>3.125</td>
</tr>
</tbody>
</table>
2.4.3. Molecular weight by gel filtration, SDS-PAGE and MALDI-TOF-MS

Purified RVL eluted as a single symmetrical peak (Fig. 2.4) on Superdex G-75 column and the molecular mass was estimated to be 49.54 ± 1 kDa from the calibration curve (Fig. 2.5). Whereas the subunit molecular mass of 11 ± 1 kDa was obtained by SDS-PAGE (Fig. 2.2). The molecular weight determined by MALDI-MS gave two distinct peaks (Fig. 2.6) corresponding to 12.02 and 12.76 kDa. Indicating that RVL has two non-identical subunits, which were neither detected by gel filtration nor by SDS-PAGE.

**Fig.2.4.** Molecular mass determination of RVL by gel filtration chromatography on Superdex-75 column (1.5 x 80 cm) equilibrated in PBS. Purified lectin (2 mg) in 2 ml of PBS was applied and eluted from the column at a flow rate of 18 ml/h and the fractions of 3 ml were collected; the column was precalibrated using molecular weight marker proteins; Bovine serum albumin (66 kDa), ovalbumin (45 kDa), pepsin (34.7 kDa), chymotrypsinogen (25.6 kDa), lysozyme (14.3kDa). (●) Absorbance at 280 nm; (▲) Hemagglutinating activity.
Fig. 2.5. Calibration curve for the estimation of molecular weight of RVL: Log molecular weight plotted against K average, marker proteins; BSA (66 kDa), ovalbumin (45 kDa), pepsin (34.7 kDa), chymotrypsinogen (25.6 kDa), lysozyme (14.3 kDa).

Fig. 2.6. Molecular weight determination by MALDI analysis showing two prominent peaks at 12.02 and 12.77 kDa respectively.
2.4.4. Determination of N-terminal sequence

Because of the ambiguities raised with respect to estimated molecular masses on SDS-PAGE, Gel filtration column and MALDI, it was necessary to substantiate the occurrence of subunits. Hence N-terminal analysis of the lectin was carried out. N-terminal sequence was determined from the lectin blots on the PVDF membrane after SDS-PAGE by automated Edman degradation protein sequencer and the PTH derivatives were identified. Interestingly sequence analyses gave two amino acid sequences of eight residues each, sequence a) LGTNYLLS and sequence b) NIPFTNNL. These results indicated the presence of two non identical subunits observed from the MALDI data.

2.4.5. Dynamic light scattering

Dynamic light scattering (DLS) were carried out to understand the oligomeric behaviour of the purified RVL. DLS studies did not indicate the presence of subunits shown by MALDI analysis. Hydrodynamic radius ($R_H$) of RVL obtained by DLS (Fig. 2.7) was 45 Å and showed 18.6 % polydispersity indicating the existence of single species.

![Fig. 2.7. Hydrodynamic radius of purified RVL determined by dynamic light scattering studies, showing the plot of % intensity Vs. hydrodynamic radius, using concentration of 0.1 mg/ml of RVL in PBS.](image-url)
2.4.6. Cloning of full length cDNA of RVL and sequence analysis

Contig assembly using sequences of 3’ and 5’ RACE-PCR products gave an assembly of 771 bp coding sequence for RVL and the sequence analysis showed RVL gene was intron less. Hence the full length RVL gene was amplified from the genomic DNA of *R. vivipara* using specific forward primer sequence RVL_F:5’TTGGCCATGGCCAAGCTGCTCCTC3’ containing NcoI site and reverse primer sequence RVL_R:5’GGCGAATTCTACGACGCAGCAA3’ containing EcoRI site, which allowed the expression in *E. coli*. From the full length coding sequence of 771 bp, coding for polypeptide chain corresponding to 256 amino acids (Fig. 2.8 A-B) accounting to molecular mass of 28.2 kDa.

Data base retrieval with BLASTp (NCBI-blast www.ncbi.nlm.nih.gov/BLAST) showed very high homology with many of the reported monocot mannose binding lectins. Multiple sequence alignment of the deduced amino acid sequence of RVL is compared with sequences of lectins from *G. nivalis*, *Z. mays*, *P. pedatisecta*, *A. macrorrhizos*, *A. amurense* and *A. lobatum* (Fig. 2.9).
Fig. 2.8 A. Full length DNA sequence and deduced amino acid sequence of RVL gene: Start and stop codons are underlined (shaded in grey), arrows indicating the cleavage sites and conserved amino acid regions are shown in bold (shaded in red). Two N-terminal sequences are shown in shaded yellow.
Fig. 2.8 B. The schematic diagram during the post translation cleavage of RVL into two mature peptides.

Fig. 2.9. Predicted amino acid sequence of R. vivipara compared with the sequences of GNA, GNAmaize, and the lectins from Araceae family; P. pedatisecta, A. macrorrhizos, A. amurense, A. lobatum. Amino acid residues at the carbohydrate binding sites are shown in grey box. (*) indicate the identical residues in all the seven lectins, (+) the identical residues in six of seven lectins and identical in five are shown by minus (-).
2.4.7. Temperature and pH stability

Purified RVL was stable in neutral pH up to 80°C for 20 min, also the lectin exhibited tolerance under wide ranging pH of 2.0 to 9.3.

2.4.8. Crystallization of RVL

On optimization of crystallizing conditions for RVL, rod-shaped crystals of RVL with good diffraction quality were obtained in 30 % polyethylene glycol 4 K, 100 mM sodium citrate buffer pH, 5.6 and 200 mM ammonium acetate. The crystals grew to dimensions of 0.15mm x 0.2mm x 1.4mm within a period of 30 days. Finally the crystal structure of RVL is determined at 2.4Å resolution (Fig. 2.10).

![Rod shaped crystals of RVL](image)

**Fig. 2.10.** The rod shaped crystals of RVL. The crustal structure of RVL is determined at 2.4 Å resolution.
2.5. DISCUSSION

Very few mannose binding lectins are characterized from Araceae species as compared to several lectins reported from other monocot families (Van Damme et al. 1995; Mo et al. 1999; Shangary et al. 1995).

The crude extract of proteins was highly viscous because of high polysaccharide content leading to problems in handling. Heat treatment of the crude extract aided in overcoming this problem. Subsequent single step affinity purification on asialofetuin-Sepharose 4B column resulted gave an excellent yield of 401.0 mg of high purity from 100 g tubers with overall 7.43 fold purification. Total soluble protein content of *R. vivipara* tubers was estimated to be 3.0 % and lectin constitute approximately 13.36 % of the total protein content. Similar reports on the lectin content in the tubers of Araceae family are previously reported from *Arum maculatum*, *Arisaema flavum*, *Arisaema consanguineum*, *Sauromatum venosum* (Singh et al. 2004; Bains et al. 2005).

RVL agglutinates rabbit erythrocytes, showed 100 fold higher hemagglutination activity towards trypsinized erythrocytes as compared to untrypsinized cells. However RVL doesn’t recognize human erythrocytes of A, B and O groups. Several other mannose binding lectins reported earlier (Van Damme et al. 2007; Sandu et al. 1990) also exhibit similar hemagglutination activity, agglutinating rabbit erythrocytes but not human RBCs of any blood group.

Hapten inhibition studies to determine the carbohydrate binding specificity indicated RVL has complex sugar specificity, since the
hemagglutinating activity of RVL was strongly inhibited by mucin, asialomucin, asialofetuin and thyroglobulin, but not by fetuin and any simple sugars. The complex carbohydrate specificity exhibited by RVL is similar to other monocot mannose binding lectins like *Colocasia esculenta* lectin, *Xanthosoma sagittifolium* lectin, *Arum maculatum* lectin and *Typhonium divaricatum* lectin (Mo et al. 1999; Van Damme et al. 2007; Luo et al. 2007). Considering the exclusive carbohydrate specificity of these lectins towards high mannose type N-glycans commonly occurring in animal glycoproteins, it is speculated that these plant proteins are evolved for defense purpose (Van Damme et al. 2007).

Also the lectin exhibited tolerance under wide ranging pH and temperature. Observed heat stability of RVL and stability under wide ranging pH is similar to *G. nivalis* agglutinin (GNA), *Aspidistra elatior* lectin and some lectins from *Araceae* family (Van Damme et al. 2007; Shangary et al. 1995; Xu et al. 2007).

Molecular weight estimation by gel filtration, SDS-PAGE and MALDI-TOF-MS has revealed that RVL has two non-identical subunits, which were not detected by either gel filtration or SDS-PAGE. Further confirmation for the occurrence of two non-identical subunits was obtained from the molecular cloning studies. Further these results were not surprising as many reported monocot mannose binding lectins composed of two polypeptides of nearly equal size which are cleaved products of single precursor poly peptide chain (Van Damme et al. 1995, 2007).
N-terminal sequence analysis gave two amino acid sequences of eight residues each, sequence a) LGTYLLS and sequence b) NIPFTNNL. MALDI analysis supported the presence of two different subunits in RVL. This property was similar with lectins from *Arum maculatum* and *Colocasia esculenta* (Van Damme et al. 1995).

The dynamic light scattering studies were carried out to understand the purity and oligomeric behaviour of the purified RVL, and it is considered as preliminary experiment for the crystallizing the lectin because it gave the basic information about the homogeneity of protein and its hydrodynamic radius ($R_H$). DLS experiment has shown that RVL has got the $R_H$ of 45 Å and showed 18.6% polydispersity indicating the existence of single dimer species. DLS studies carried out for peanut agglutinin (Dev et al. 2006) and soyabean lectin (Fasina et al. 2003) also indicated similar properties. From these results, it was concluded that the native RVL is a dimeric molecule without indication for heterogeneity. Hence it was crystallized successfully.

The recombinant RVL encodes for 256 amino acids. The N-terminal sequence a) LGTYLLS of RVL obtained by Edman sequencing coincided with deduced amino acid sequence of the cDNA between L$_{24}$ to S$_{31}$ whereas the sequence b) NIPFTNNL coincided between N$_{140}$ to L$_{147}$. Considering the rules for predicting the signal peptide (Von Heijine 1986), it may be concluded that RVL has signal sequence peptide of 23 amino acids accounting for the molecular mass of 2.5 kDa, which upon cleavage between A$_{23}$ and L$_{24}$ residues results in 233 residue polypeptide representing the native RVL. In other words
the polypeptide of 233 residues with molecular mass of 25.7 kDa is derived from the precursor molecule of 256 amino acid residues. The eight residue N-terminal sequence of b- peptide deduced, align from 140-147 residues evident that the lectin is synthesized as a single large precursor (233 residues), which undergo post translational cleavage between residues R$_{139}$ and N$_{140}$ and developed into two mature peptides.

These findings also indicated the presence of two putative domains in RVL, each containing one mannose binding site as reported for other monocot mannose binding lectins like Arum maculatum lectin, Colocasia esculenta lectin, Xanthosoma sagittifolium lectin (Van Damme, E.J.M., et al., 1995), Pinellia pedatisecta lectin (Lin et al. 2007; Van Damme et al. 1998; Fei et al. 2003) and Typhonium divaricatum lectin (Luo et al. 2007).

Mass spectrometric data of RVL showed two consistant peaks of masses 12.020 kDa and 12.760 kDa. Calculated mass of 12.715 kDa for the polypeptide chain of 116 amino acids (24-139) is in agreement with MALDI mass of 12.760 kDa. However, the mass of 13.00 kDa calculated for the polypeptide chain of 117 residues (140-256) was not in accordance with the MALDI estimate of 12.02 kDa. This discrepancy could be attributed to post translational processing of propeptide on the C-terminal side of the polypeptide chain (140-256) as suggested for other GNA related lectins (Van Damme et al. 2007).

It is evident from the sequences retrieved from NCBI database, RVL showed 80-90% homology with several reported monocot mannose binding
lectins of Araceae family. However the extent of sequence homology of RVL with GNA and GNA\textsubscript{maize} is only 41 % and 38 % respectively. Indeed RVL shared common conserved motif of \textbf{Q-x-D-x-N-x-V-x-Y} at the carbohydrate binding site. Any changes occurring due to insertion or deletion of key amino acids in this conserved motif is determinant in conferring exclusive specificity for these lectins towards high mannose N-glycans (Van Damme et al. 1995; Hirai et al. 1993). Thus RVL is closely related to other monocot mannose binding lectins of Araceae family, but is different from GNA with respect to mannose binding property and the number of binding domains.

Finally to summarize briefly about the findings of this chapter, RVL is purified to homogeneity by single step affinity chromatography from the tubers of edible epiphyte \textit{Remusatia vivipara}. RVL specifically agglutinated the rabbit erythrocytes but not, any of the human blood group antigens. The hapten inhibition studies have demonstrated that, RVL having affinity towards the complex sugars but not towards simple sugars and their derivatives. RVL is a tetramer of 49.54 kDa with subunits of 12.0 kDa as determined by gel filtration chromatography and SDS-PAGE. However the MALDI analysis and the N-terminal sequencing by Edman degradation have shown the presence of two different subunits in RVL. But, the Dynamic Light Scattering study has demonstrated that RVL preparation is a single species with a hydrodynamic radius of 4.5 nm. Further the molecular cloning and the sequence analysis indicated that RVL is synthesized from single gene, as polypeptide precursor, which is later cleaved into two mature peptides by post-translational
modification. Finally, the RVL is successfully crystallized, and the structure is determined at 2.4 Å resolution (Shetty et al. 2012).

Considering the sugar specificity and unique structural design of RVL, it can have immense application in medicine and agriculture. Being a predominant storage protein in the vegetative tissues, RVL has thought to be involved in the defense mechanism of the plant against various attackers like insect pests.