Chapter 2

Materials and Methods

2.1 Introduction

The present chapter deals with materials and methods used for purification and biochemical and biophysical characterization of *Spatholobus parviflorus* lectin (SPL). The purification protocol included ammonium sulphate precipitation, ion exchange chromatography and affinity chromatography.

2.2 Materials

Seeds of *Spatholobus parviflorus* were the source for the isolation of lectin.The seeds were obtained from campus of Pariyaram medical college in kannur district, kerala. They were washed, dried in sunlight and stored in refrigerator at 4^{o}C. Various chemicals and equipments used have been specified by manufactures/suppliers, make, and model etc appropriately.

Chemicals

The following chemicals have been purchased from Sigma Chemicals Co. St .Louis, USA: Sephadex C 50, BSA, N –acetyl D-galactosamine, polyethylene glycol, and molecular weight markers for electrophoresis.

The following chemicals were obtaine d from Hi media laboratories Mumbai: D-fucose, D–xylose, D-ribose, lactose, sucrose, starch, D-mannitol, arabinose, maltose, D- glucose, mannose, D-galactose, bromophenol blue, trypsin, agar, guar gum and anthrone reagent.

The following chemicals were procured from BDH Chemicals, England: B-mercaptoethanol, acrylamide and bis-acrylamide.

The following chemicals were purchased from MERCK, Mumbai:

Sodium dodecyle sulphate, ammonium per sulphate, N,N,N,N, Tetramethyl ethylene diamine (TEMED) sodium chloride, sodium citrate, ctric acid, disodium hydrogen phosphate, potassium dihydrogen phosphate, ammonium sulphate, tris base, petroleum ether, sodium hydroxide, 2-Methyl, 2-4 pentane diol (MPD) and epichloro hydrin. Coomassie brilliant blue R-250 purchased from Fluka AG Chem, sephadex C-50, polyethylene glycol 8000(PEG 8K), and molecular weight markers were purchased from GeNei Bangalore.

2.3 Red blood cells

Human blood samples were collected from healthy volunteers of the school of life sciences, Kannur University. All the blood samples were stored at 4^{o}C with anticoagulant-acid dextrose until required.

2.3.1 Trypsinised erythrocytes

Erythrocytes were collected from the stock blood suspension by centrifugation at 5000 rpm for 5 min. The RBCs were washed three times in PBS and resuspended in PBS to get a suspension of 4% (v/v). To 100 ml of this suspension, 1ml of 1% trypsin was added and the mixture was incubated at 37^{o}C for 1 hr.The erythrocytes were then washed three times in PBS to remove the trypsin and finally resuspended in PBS to give a final concentration of 2% erythrocyte suspension.

2.4 Purification of SPL

All the isolation procedures were carried out at 4^{o}C. The dry seeds were ground well with mortar and pestle and sonicated.The sonicated slurry was stirred at 4^{o}C overnight with 20 mM phosphate buffered saline, pH 7.4. The slurry was filtered through cheese cloth and centrifuged at 15,000g for 15

minutes. The supernatant was collected and filtered through the Whatman No 1 filter paper. The clear pale yellow filtrate was used for the isolation of lectin.

2.5 Ammonium sulphate fractionation

One of the initial steps involves purifying proteins based on their solubilties in varying concentrations of ammonium sulfate. The solubility of a protein is sensitive to the concentrations of dissolved salts. The solubility of a protein at low ionic strength generally increases with the salt concentration (salting in). At high ionic strength, the solubility of proteins decreases (salting out). Many unwanted proteins can be eliminated by adjusting the salt concentration in a solution containing the crude extract to just below the precipitation point of the protein to be purified. In this case the protein to be purified remains in solution, while many others are precipitated. Likewise, unwanted proteins can be eliminated by adjusting the salt concentration to just over the precipitation point of the protein of interest. In this case, the protein of interest will be precipitated while many others will remain in solution. The extract was brought to 50% saturation with ammonium sulphate by the gradual addition of the salt. The pH was adjusted to 7.4 and the mixture was stirred for one hour, and then centrifuged at 20,000g for 15 minutes. The precipitate was discarded and the clear supernatant was collected. The supernatant is then brought to 70% saturation of ammonium sulphate by the gradual addition of salt and adjusted the pH to 7.4. Then the slurry was stirred and centrifuged at 20,000g for 15 minutes. The supernatant was discarded and the precipitate was collected and dissolved in minimum amount of phosphate buffered saline (PBS) at pH 7.4.

2.6 Dialysis of ammonium sulphate fraction

Dialysis is a protein purification method that separates proteins from other small molecules, such as salt, by using a semipermeable membrane. This membrane contains micro pores through which the small molecules will escape. Therefore, protein molecules having dimensions significantly greater than the pore diameter are retained inside the dialysis bag. The small molecules and salt will diffuse out through the membrane and into the dialysate outside of the bag. After the initial step of ammonium sulphate precipitation, the precipitate was dialysed extensively against several changes of PBS. Afterwards, the insoluble materials in the protein solution were removed by centrifugation at 15,000g for 10 minutes. And, the clear supernatant solution was subjected to chromatographic sepation.

2.7 Ion exchange chromatography

2.7.1 Standardization of ion exchange chromatography

2.7.1.1 Standardization of pH

Different buffers of varying pH were prepared (2-10.5), 1 gm of Sephadex C 50 were swelled in 50 ml deionized water and kept overnight at 4^{o}C and made the slurry. 1 ml slurry was taken in boiling test tubes and adds 2 ml of buffer of varying pH in different test tubes. To each test tube 0.5 ml of protein sample was added and kept overnight for incubation. After incubation, the supernatatnt was taken and hemagglutination test was conducted. The test tube at which maximum activity was taken as the optimum pH.

2.7.1.2 Standardization of elution buffer

The elution was done with the same buffer containing a linear gradient of NaCl (0-5 M). Different molar concentration of NaCl (0-5 M) was added 1 ml each to boiling tube containing resin and buffer of optimum pH. In each tube 0.5 ml of protein sample were added, mixed it thoroughly and kept it overnight for incubation. Hemagglutination test was conducted by using supernatant.

Ion exchange chromatography allows molecules to be separated based upon their charge. If the protein has more positive charges it is saidto be a basic protein. If the negative charges are greater, the protein is acidic. A basic protein, which is positively charged, can bind to a support, which is negatively charged. An acidic protein which is negatively charged, will bind to positively support. Ion exchange chromatography can be subdivided into cation exchange chromatography, in which positively chaged ion binds to a negatively charged resin. And anion exchange chromatography, in which binding ions are negative and the immobalised functional group is positive. So, ion exchange chromatography relies to charge–charge interactions between the proteins in the sample and the charges immobilized on the resin. Ion exchange chromatography uses pellicular polymeric resins that are compatible with a wide spectrum of pH.

The dialysed ammonium sulphate fraction (50-70) was loaded on a sephadex C 50 column; pre equilibrated with the dialysis buffer. The resin bound protein was eluted with a salt gradient prepared in buffers containing 0.1M NaCl to 0.5 M NaCl. Elution of the column was at flow rate of 50 ml/h and the elution profile was monitored by absorbance at 280nm. 4 ml fractions were collected and all fractions were checked for heamagglutination

activity.The fractions showing heamagglutination activity were pooled and it was dialysed extensively against PBS and it was subjected to further purification (Silva *et al*., 2007).

2.8 Purification by affinity chromatography

2.8.1 Activation of Guar Gum

Guar Gum was weighed (16 gm) and then mixed with an emulsion of 30 ml 3 N NaOH and 3 ml epichlorohydrine at 40^0C for 24 hours in a waterbath and then at 70^0C in a hot air oven for 12 hours. The activated guar gum was then washed several times with distilled water until neutralized. Then, it is filled with column of 3×60 cm dimensions and was equilibrated with PBS, pH 7.4.

2.8.2 Affinity chromatography

The clear protein solution was applied to the column at a flow rate of 50 ml/h. It was then allowed to stay for 2 hours. After loading the sample, the column was washed by the same buffer of three times of bed volume to remove all the unbound proteins. The bound protein was then decoupled with 0.2 M D-galactose in PBS, pH 7.4. The elution was carried out at a flow rate of 40 ml/h. The eluted protein was collected as fractions of 3 ml. The fractions containing pure protein only were pooled and dialyzed extensively against PBS, pH 7.4, with several changes and finally against distilled water. Then the dialyzed protein solution was concentrated with polyethylene glycol 20000 and lyophilized. All the above process was carried out at 4^oC.

2. 9 Iso electric focusing (IEF)

The pI of the SPL was determined by using the technique of isoelectric focusing. 40 ml of the lectin extract mixed with 1.0 ml of ampholyte (pH 3/10, BioRad) was loaded into Rotofor (Preparative Iso-Electric Focusing cell (IEF); BioRad) under controlled manner using a 50 ml syringe and the IEF chamber was filled completely with the sample. A power pack was connected to the system that delivered 12 W to the chamber. The resultant voltage from the system was noted during an interval of every 30 min. A gradual increase of the voltage (from 300 V to 500 V) for three hours was noted. The whole unit was connected to a cooling system and was maintained at 4°C during the entire run. The concentrated and separated samples were vacum harvested as (2.0 ml x 20) different fractions from the IEF chamber using a harvesting kit. The pH of each of the fractions obtained was read. Subsequently, the fractions were assayed for hemagglutination activity as well as the total soluble protein content.

2.10 Sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE)

2.10.1 Reagents for SDS-PAGE

SDS-PAGE was carried out according to the procedure of Weber and Osborn (1969) at pH 7.2 and also according to the procedure of Laemmli (1970). Laemmli system SDS-PAGE was discontinuous. The system consists of two types of gels: a) stacking gel, and b) resolving gel.

Reagents

1.	Acrylamide	: 30.0%
	N,N′ methylene bis acrylamide	: 0.8%
	Distilled water to make 100 ml	
2.	Resolving gel buffer	
	Tris-HCl, pH 8.8	: 3.0 M
3.	Stacking gel buffer	
	Tris-HCl, pH 6.8	: 0.5 M
4.	Ammonium per sulphate (APS) freshly prepared	: 100 mg
	Distilled water to make 1 ml	
5.	Sodium dodecyl sulphate	: 10 g
	Distilled water to make 100 ml	
6.	Running or reservoir buffer, pH 8.3	
	Tris base	: 3.03 g
	Glycine	: 14.4 g
	Sodium dodecyl sulphate	: 1.0 g
	Distilled water to make 1000 ml	
7.	**Sample buffer**	
	Glycerol	: 1.6 ml
	Tris-HClpH 6.8	: 1.0 ml
	2-mercapto ethanol	: 0.4 ml

Sodium dodecyl sulphate	: 1.6 ml
Bromophenol blue	: 0.4 ml
Distilled water	: 3.0 ml

8. **Staining solution**:

Coomassie brilliant blue R-250	: .5%(v/v)
Distilled water	: 100 ml

9. **Destaining solution:**

Glacial acetic acid	: 7.5 %
Distilled water to make 100 ml	

1. **Resolving gel composition:** (12.5%)

Acrylamide-bis acrylamide	: 12.5 ml
Resolving gel buffer stock	: 3.75 ml
Sodium dodecyl sulphate	: 0.3 ml
Ammonium per sulphate	: 1.5 ml
Water	: 14.45 ml
TEMED	: 0.020 ml

2. **Stacking gel composition**:

Acrylamide-bis acrylamide	: 2.5ml
Stacking gel buffer	: 5 ml
Sodium dodecyl sulphate	: 0.2 ml
Ammonium per sulphate	: 1 ml
Water	: 11.3 ml
TEMED	: 0.020 ml

Marker-GeNei PMWM widerange marker was used which

contains the following proteins

Protein	**Molecular weight (Daltons)**
Phosphorylase b	:97,400
Albumin, bovine serum	: 66,000
Ovalbumin,	: 43,000
Carbonic anhydrase, bovine erythrocytes	: 29,000
Trypsin inhibitor, soybean	: 20,100
lysozyme	: 14,300

Stock solutions

Reconstituted each vial with 100 µl of deionized water and kept in freezer.

2.10.2 Procedure

12% gel was prepared and allowed to pre run for 1 hour at 80 V. Protein concentrations of the samples were made equal. Sample and sample buffer were mixed in the ratio 1:1 and kept in a boiling water bath for 5 minutes and cooled. 20 µl was loaded in to each well. Marker mixed with sample buffer (in 1:1 ratio) was loaded into one well. It was allowed to run at 70 V. After the completion of running, gel was stained for 30 minutes and then destained in 7.5% acetic acid.

2.11 Matrix-Assisted Laser Desorption/Ionization mass Spectroscopy

The SPL was analysed on a BRUKER MALDI ToF Mass spectrometer. Mass spectrum was recorded on an Ultraflex TOF/TOF MALDI mass spectrometer (Bruker Daltonics) using Cinapinic acid as matrix.The nitrogen laser used 337 nm. Total 500 scans were averaged for final spectrum.

2.12 Protein estimation

Protein concentration was determined by the method of Lowry et *al.* (Lowry et *al*., 1951) using bovine serum albumin as standard

2.12.1 Lowrys Method

Reagents

1. Sodium hydroxide solution 0.1 N
2. Folins and Ciocalteus phenol reagent
3. Alkaline copper solution

2.12.2 Preparation of alkaline copper

Prepared 0.5% $CuSO_4$, 1% sodium potassium tartarate and 2% Na_2CO_3 in 0.1 N sodium hydroxide Mixed 0.5 ml of copper sulphate, 0.5 ml of sodium potassium tartarate and 50 ml of Na_2CO_3 in 0.1 N sodium hydroxide.

2.12.3 Procedure

To 0.2 ml of protein solution or test solution, added 5 ml of alkaline copper and 1.8 ml of 0.1 N sodium hydroxide. Incubated the mixture for 15 minutes at room temperature. Then added 0.5 ml of folins-phenolic reagent to the above mixture and incubated again for 30 minutes at room temperature. Blue colour developed, indicated the presence of protein.

2.13 Hemagglutination Tests

The human blood samples were collected from healthy persons. The blood samples were collected aseptically and stored in acid citrate dextrose containinig bottle (acid: blood ratio 0.4:5 ml). When required, erythrocytes were washed with phosphate buffered saline, pH 7.4. The hemeagglutination assay was done in a microtitre plate, using washed RBC of A, B, O and AB phenotypes, of human erythrocytes. The RBCs was washed twice in PBS, pH 7.4 to get a cell suspension of 2% (v/v). The hemagglutination activity of SPL was assayed by serial two-fold dilution method. 100 µl of the plant extract was serially diluted in PBS and each dilution was mixed with equal volume of 2% RBC suspension and incubated for one hour at room temperature. The end point of assay was visually estimated after an hour. The titre was taken as the reciprocal of highest dilution of lectin giving visible agglutination which is a hemagglutination unit (HAU).

2.14 Biochemical and biophysical characterization of SPL

2.14.1 Estimation of total sugar content

Total neutral sugar content of the purified lectin preparation was estimated by anthrone method using D-glucose as standard. This is a simple calorimetric method with relative insensitivity to interferences from the other cellular components.

2.14.1.1 Materials

Anthrone reagent: Dissolved anthrone in 100 ml of ice cold 95% $H2SO_4$, Prepared fresh before use. Standared glucose: stock-Dissolved 100 mg in 100

ml water, working standared- 10 ml of stock diluted to 100 ml with distilled water.

2.14.1.2 Procedure

The first step in total carbohydrate measure was to hydrolyse the polysaccharides and to dehydrate the monomers. The 5 carbon and 6 carbon sugers were converted to furfural and hydroxymethyl furfural, respectively. When anthrone was added it reacted with digestion product to give coloured compound. The amount of total carbohydrate in the sample is then estimated via reading the absorbance of the resulting solution against a glucose standard curve. Prepare the anthrone solution freshly and also prepared the standared glucose solution. Final concentration of the glucose was 100mg glucose per litre. Pipetted out working standards as 0.2 ml, 0.4 ml, 0.6 ml, 0.8 ml, and 1 ml, and test solution, diluted with distilled water. Added 4 ml of anthrone reagent. Placed the tubes on a heating block. Removed the tube and cooled down to the room temperature, record it at 578 nm.

2.14.2 Sugar inhibition studies

The carbohydrate binding specificity of purified SPL was studied by carrying out sugar inhibition assays described under methods. The hemagglutination inhibition study of the lectin was done by different sugars. D-glucose, D–galactose, Nacetyl-D–galactosamine, D-fructose, lactose, sucrose, D–mannose, D–arabinose, D-fucose, D-xylose, maltose, N-acetyle-D-glucosamine, melibiose, D-ribose, D-sorbitol, raffinose, were the sugars used. The assay was performed as follows: 45µl of different sugar solution (0.2 M) was placed in the microtitre plate and serially diluted. Then 10 µl of the purified lectin (2.5 mg/ml) was added to each well. After an hour of incubation at 37^{o}C 45 µl of 4% human RBC was added and the mixture was

allowed to stand for one hour at room temperature. The degree of hemagglutination was examined and maximum dilution of sugar solution showing hemagglutination inhibition was recorded.

2.14.3 Temperature stability studies of SPL

Temperature is an important factor which can affect the protein structure, folding and stability. Gentle heating above a critical temperature can loosen the linkage. Increase in temperature, in certain cases results increase in chemical potential energy and it is great enough to breake some of the weak bonds that determine the the three dimentional shape of the active proteins. This could lead to thermal denaturation of the protein and thus inactivate the proteins. Inactivation by heat denaturation has a profound effect on the activity. Thermal stability of purified lectin was studied by incubating samples at various temperatures. Lectin solution (2.5 mg/ml) was taken in a screw cap tube. Each sample was heated for 15 minutes at 30, 40, 50, 60, 70, 80, 90, 100^{o}C. 50 µl aliquots were drawn from each treatment and cooled to the room temperature and assayed for hemagglutinating activity with human erythrocytes checked for residual hemagglutination activity.

2.14.4 pH dependant studies of SPL

The pH dependence of SPL was examined by using buffers ranging from pH 2-12. The buffers used were glycine-HCl for pH 2 and 3, prepared by using 0.2 M glycine and 0.2 M HCl, acetate buffer for pH 4 and 5. Citrate buffer was used for buffer with pH 6. Phosphate buffer was ued as buffer for pH 7 and 8, and for pH 9, 10 and 11 glycine–NaOH used. A volume of 50 µl lectin solution was incubated with 50 µl of buffer for 1 hour at room temperature. Hemagglutination assay was conducted with 4% of human erythrocytes.

2.14.5 Effect of denaturing agents

Denaturation of proteins involves the disruption and possible destruction of both the secondary and tertiary structures. Since denaturation reactions are not strong enough to break the peptide bonds, the primary structure (sequence of amino acids) remains the same after a denaturation process. Denaturation disrupts the normal alpha-helix and beta sheets in a protein and uncoils it into a random shape.

Denaturation occurs because the bonding interactions responsible for the secondary (hydrogen bonds to amides) and tertiary structures are disrupted. In tertiary structure there are four types of bonding interactions between side chains including hydrogen bonding, salt bridges, disulfide bonds, and non-polar hydrophobic interactions that may be disrupted. Therefore, a variety of reagents and conditions can cause denaturation. The most common observation in the denaturation process is the precipitation or coagulation of the proteins. In quarternary structure denaturation, protein sub-units are dissociated and/or the spatial arrangements of protein subunits are disrupted. Tertiary denaturation involves the disruption of covalent interactions between amino acid side chains (such as disulfide bridges between cysteine groups), non covalent dipole-dipole interactions between polar amino acid side chains and the (surrounding solvent) van der Waals (induced dipole) interactions between nonpolar amino acid side chains. In secondary structure denaturation, proteins lose all regular repeating patterns such as alpha-helices and beta-pleated sheets, and adopt a random coil configuration. Primary structure, such as the sequence of amino acids held together by covalent peptide bonds, is not disrupted by denaturation. Most proteins lose their biological function when denatured. For example, enzymes lose their activity, because the substrates can no longer bind to the active site,

since amino acid residues involved in stabilizing substrates in transition states are no longer positioned to be able to do so. In many proteins denaturation is reversible (the proteins can regain their native state when the denaturing influence is removed). This was important historically, as it led to the notion that all the information needed for proteins to assume their native state was encoded in the primary structure of the protein, and hence in the DNA that codes for the protein.

The effect of denaturing agents such as urea, thiourea, and guanidine hydrochloride was examined on SPL activity. In this experiment, a wide range of concentrations of denaturants from. 0.5 to 8 M at 0.5 M interval was used. The 50 µl of each test solution was incubated with 50 µl of each lectin solution (2 mg/ml) in a microtitre plate at 37^{o}C for 1 hour and the hemagglutination activity was checked for untreated and treated samples.

2.14.6 Metal ion requirement

2.14.6.1 Atomic absorption spectroscopy for metal ion detection

The absorption of energy by ground state atoms in the gaseous state forms the atomic absorption spectroscopy. When a solution containing metallic species is introduced into a flame, the vapour of metallic species will be obtained. When a light of specific wave length is allowed to pass through a flame having atoms of the metallic species, part of that light will be absorbed and the absorption will be proportional to the density of the atoms in the flame.

Atomic absorption (AAS) was employed for the accurate determination of Mn and Ca. It was conducted in AAS Varian AA 240 FS flame mode spectrophotometer (Varian Inc. USA). Air, acetylene flame for,

Mn, Ca was used. The wavelengths used for Mn was 279.5 nm and Ca 422.7 nm. The protein sample was aspirated and mixed as an aerosol with combustible gasses (acetylene and air or acetylene and nitrous oxide) in flame atomic absorption spectroscopy. The mixture was ignited in a flame of temperature ranging from 2100°C to 2800°C (depending on the fuel-gas used). During combustion, atoms of the element of interest in the sample were reduced to the atomic state. A light beam from a lamp, whose cathode made of the element being determined, was passed through the flame into a monochronometer and detector. Free, unexcited ground state atoms of the element absorb light at characteristic wavelengths. This reduction of the light energy at the analytical wavelength was a measure of the amount of the element in the sample. In general, after choosing the proper hollow cathode lamp for the analysis, the lamp should be allowed to warm up for a minimum of 15 minutes, unless operated in a double beam f-y mode. During this period, aligned the instrument, positioned the monochromator at the correct wavelength, selected the proper monochromator slit width, and adjusted the hollow cathode current according to the manufacturer's recommendation.

2.14.6.2 Metal ion requirement of SPL

The metal ion requirement for lectin activity was examined by demetallizing the sample and then treating with different metal ions (Kawagishi *et al* 1990). Lectin solution (100 µl of 2.5 mg/ml) as taken in an eppendorf tube and incubated with 400 µl of 10 mM EDTA at pH 5.0 for 20 h at 4°C. The sample was then dialyzed against 20 mM PBS pH 7.4 and 50 µl aliquots were transferred to eppendorf tubes containing 50 µl of l mM $CaCl_2$,

$MnCl_2$ and $MgCl_2$ and incubated for 2 hrs. Activity of the samples was then examined by agglutination.

2.14.7 Antifungal property of SPL

The assay of the lectin for antifungal activity toward fungal many species were carried out in 100 x 15 mm. petri plates containing 10 ml of potato dextrose agar (PDA) (Wang *et al*., 2004). The fungal mycelia were inoculated at the centre of the plate. After the mycelia colony had developed, sterile blank paper disks (0.625 cm in diameter) were placed at a distance of 0.5 cm away from the rim of the mycelial colony. An aliquot (15 µl) of the lectin with 10 µg, 5 µg, 2.5 µg, 1.25 µg and 0.625 µg/ml were added to a disk while 15 µl of the buffer served as control. The plates were incubated at 25°C for 72 hrs until mycelia growth had enveloped peripheral disks containing the control and had formed crescents of inhibition around disks containing samples with antifungal activity. The organisms used for this test were *Aspergillus niger, Aspergillus flavus, Fuscarium oxysporium, Aspergillus fumigatus, Penicillium sp., Rhizopus sp., Trichoderme sp., Neurospora sp. and Cladospermum sp.*

2.15 Fluorescence studies of SPL

2.15.1 Introduction

Lectins of leguminous plants have predominantly β-pleated sheet structures, as is evident from their CD spectra in the 190-250 nm regions (Sharon & Lis, 1990). Since β-sheet CD spectra are relatively weak, contributions from transitions of side chain chromophores in this region will be significant. There is no readily apparent correlation between these CD assignments to groups and the carbohydrate specificities of the lectin. Spectroscopic methods are traditionally been used for directly monitoring

protein-unfolding transitions. Fluorescence, circular dichroism, differential UV/visible spectroscopy etc. are the common solution phase spectroscopic methods.

Among the more common and convenient spectroscopic methods for tracking the unfolding transitions are, the measurements of circular dichroism in the deep UV-region where the signal responds to change in secondary structures and fluorescent spectroscopy that depends on aromatic residues where, for instance, changes in the fluorescent intensities of tryptophan reflect changes in the microenvironment of these residues. The UV-spectrophotometric data can also provide changes in secondary and tertiary state of proteins. (Yanari & Bovey, 1960).

By measuring the protein secondary structure elements, experimental verification of a predicted folding motif may be obtained. Circular dichroism and fluorescence spectroscopy are highly sensitive optical measurements by which the fraction peptide bonds in α-helical, β-pleated sheets and a periodic conformation may be estimated. Moreover, the quantitative measurement of the secondary structure provides significant insight into the structural features critical to biological functions. Many problems associated with molecular biology, such as protein folding, protein-protein interactions, etc. can be cleared by these sensitive spectroscopic techniques. Spectral effects of conformational changes that accompany substrate and inhibitor binding, subunit assembly and effects of agents like heat, urea, and guanidine hydrochloride may also be observed as they provide valuable information on protein folding and especially changes on site directed mutagenesis and conformation of expressed proteins. In addition, spectroscopic measurements of secondary structures are a valuable tool for assessing protein aggregation and stability.

Quenchers are substances that reduce the fluorescence. The best known collisional quencher is oxygen (Kautsky, 1939). Aromatic and aliphatic amines are efficient quenchers of fluorescence of most unsubstituted aromatic hydrocarbons (Knibbe *et al*., 1968). Acrylamide is an effective collisional quencher that targets the amino acid tryptophan (Eftink & Ghiron, 1976). Quenching with larger halogens like iodide may be a result of intersystem crossing to an excited triplet state, promoted by the spin-orbit coupling of the excited fluorophore and halogen (Kaska, 1952). Iodide is negatively charged and hydrated. Its quenching activity is assumed to be limited only to the surface tryptophan residues (Lackowicz *et al*., 1983). Fluorescence quenching studies were usually done to look into the changes in tryptophan environment of a protein. Acrylamide is a neutral efficient collisional quencher. Similarly, iodide is also a negatively charged hydrated quencher and its action is normally restricted to the surface tryptophan residues (Lackowicz *et al*., 1983). These molecules do not readily penetrate the hydrophobic interior of proteins, and only those tryptophan residues on the surface of proteins are quenched.

2.15. 2 Unfolding studies

In order to verify the changes in protein folding under various environments, fluorescence spectra of SPL were recorded on a Perkin Elmer LS 55-spectrofluorimeter using excitation wavelength of 280 nm. Protein concentration was 4 mg/ml and emission slit width of 2.5 nm were used. Protein samples were incubated for one hour at each denaturant (urea, GuHCl) in 20 mM Tris HCl buffer, pH 8.0 before recording the spectra. The scan range was between 300 and 400 nm and the scan speed was 120 nm/min. The response width was 31nm. Fluorescence intensity was measured in a.u. (arbitary units).

2.15.2.1 Guanidine hydrochloride induced denaturation

4 mg/ml protein sample was incubated for one hour with guanidine hydrochloride (GuHCl) of concentrations ranging from 0.5 to 6 M and the fluorescence spectrum is obtained. The scanning range was 300 nm- 400 nm. The excitation was done at 280 nm.

2.15.2.2 Urea induced denaturation

The changes in fluorescence at varying concentrations of urea are monitored 4 mg/ml lectin preparation was incubated for one hour with urea concentration is from 0.5 M to 6 M. The excitation wavelength is 280 nm and scanning range was 300- 400 nm.

2.15.2.3 Quenching studies.

Quenching of intrinsic fluorescence by acrylamide (Eftink & Ghiron, 1976) was performed for the native protein. Aliquots from a 4 M stock solution of acrylamide were added to protein samples in a thermostated cuvette at 25°C. Emission intensities at 280 nm were monitored and the bandwidths (excitation and emission) were 5 nm. Quenching of fluorescence was also performed with iodide.

Protein concentration used for quenching studies was 4 mg/ml. Quenching fluorescence spectra of SPL for various acrylamide concentrations were recorded. The concentrations of acrylamide used ranges from 0.02 M to 0.18 M. Similarly potassium iodide concentrations of the range from 0.02 M to 0.2 M were used for iodide quenching studies.

2.16 Circular dichroism studies of SPL

2.16.1 Circular dichroism

Circular dichroism (CD) spectroscopy measures differences in the absorption of left-handed polarized light versus right-handed polarized light which arise due to structural asymmetry. The absence of regular structure results in zero CD intensity, while an ordered structure results in a spectrum which can contain both positive and negative signals. This is an excellent tool for rapid determination of the secondary structure and folding properties of proteins that have been obtained using recombinanat techniques or purified from tissues. Circular dichroism spectroscopy is particularly good for: determining whether a protein is folded, and if so characterizing its secondary structure, tertiary structure, and the structural family to which it belongs. Comparing the structures of a protein obtained from different sources (e.g. species or expression systems) or comparing structures of different mutants of the same protein demonstrating comparability of solution conformation after changes in manufacturing processes or formulation comparability protocols, studying the conformational stability of a protein under stress -thermal stability, pH stability and stability to denaturants and how the stability is altered by buffer composition or addition of stabilizers and excipients. CD is excellent for finding solvent conditions that increase the melting temperature and/or the reversibility of thermal unfolding, conditions which generally enhance shelf life and determining whether protein-protein interactions alter the conformation of protein components. lf there are any conformational changes, this will result in a spectrum which will differ from the sum of the individual components.

2.16.2 Determination of protein secondary structure by Circular Dichroism

Secondary structure can be determined by CD spectroscopy in the 'far-UV'spectral region (190-250 nm). At this wavelength, the chromophore is the peptide bond and the signal arises when it is located in a regular, folded environment. Alpha-helix, beta-sheet and random coil structures will generate characteristic CD spectra. Secondary structure elements present in any protein can thus be determined by analyzing its 'far-UV' CD spectrum as a sum of fractional multiples of such reference spectra for each structural type. Like all spectroscopic techniques, the CD signal reflects an average of the entire molecular population. Thus, while CD can determine that a protein contains about 50% alpha-helix, it cannot determine which specific residues are involved in the alpha-helical portion. 'Far-UV' CD spectra require 20 to 200, µl of solution containing 1 mg/ml to 50 µg/ml protein, in any buffer which does not have a high absorbance in the region of the spectrum. (High concentrations of histidine, or imidazole, for example, cannot be used in the 'far-UV' region.)

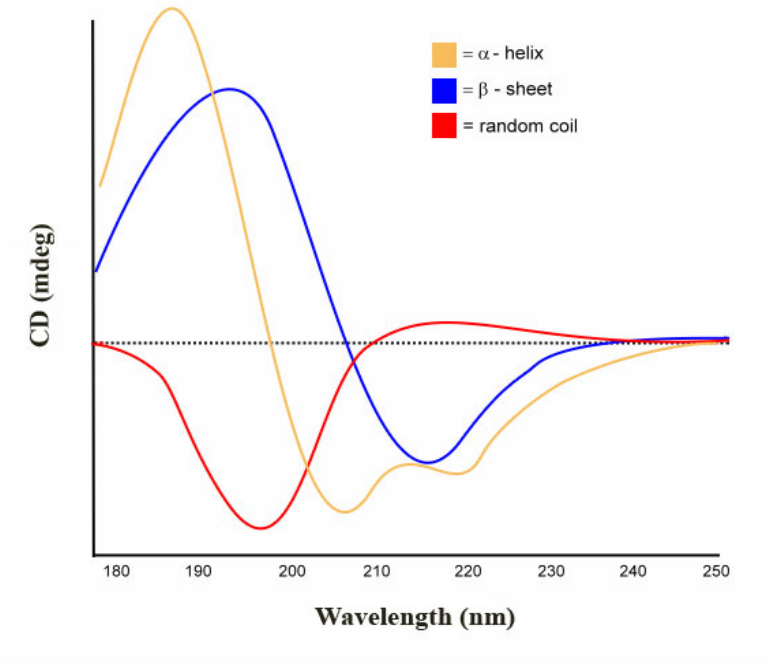

Figure 2.1: CD spectrum of secondary structure protein

(www.Proteinchemist.com)

2.16 .3 Information about protein tertiary structure from Circular Dichroism

The CD spectrum of a protein in the 'near-UV' spectral region (250-350 nm) can be sensitive to certain aspects of tertiary structure. At these wavelengths the chromophores are the aromatic amino acids and disulfide bonds and the CD

signals they produce are sensitive to the overall tertiary structure of the protein. Signals in the region from 250-270 nm are attributable to phenylalanine residues, signals from 270-290 nm are attributable to tyrosine and those from 280-300 nm are attributable to tryptophan. Disulfide bonds give rise to broad weak signals throughout the 'near-UV' spectrum. If a protein retains secondary structure but no defined three-dimensional structure (e.g. an incorrectly folded or "molten-globule" structure), the signals in the 'near-UV' region will be nearly zero. On the other hand, the presence of significant 'near-UV' signals is a good indication that the protein is folded into a well-defined structure. The 'near-UV' CD spectrum can be sensitive to small changes in tertiary structure due to protein-protein interactions and/or changes in solvent conditions.The signal strength in the near-UV CD region is much weaker than that in the 'far-UV' CD region. 'Near-UV' CD spectra require about 1 ml of protein solution with an OD at 280 nm of 0.5 to 1 (which corresponds to 0.25 to 2 mg/ml for most proteins).

2.16. 4 Demonstrating comparability of conformation

Often it is necessary to demonstrate that different lots of a protein have equivalent conformations. For example after a scale-up in the purification process or to qualify a new manufacturing site the CD can be used for testing/ensuring quality. Thermal stability is assessed using CD by following changes in the spectrum with increasing temperature. In some cases, the entire spectrum in the 'far'or 'near-UV' CD region can be followed at a number of temperatures. Alternatively, a single wavelength can be chosen which monitors some specific feature of the protein structure and the signal at that wavelength is recorded continuously as the temperature is raised. CD is also often used to

assess the degree to which pH of solution, buffers, and additives such as sugars, amino acids or salts alter the thermal stability.

2.16.5 Experimental procedure

Circular dichroism spectroscopic measurements were carried out by JASCO J-715 Spectropolarimetre. Cuvetts were washed with ethanol and water before starting the experiments. Baseline correction was done by using 10 Mm PBS, pH 7.4. SPL concentration of 1 mg/ml was taken in a quartz cuvett. Measured the spectrum in 'far UV' (250-190 nm) with band width 2.0 nm, response 8 sec and scanning speed was 50 nm/min at room temperature, three scans were averaged for each spectra. The percentage of helix, sheet and random coil strand was calculated by K2D.

2.17 Crystallization and X- ray structure determination

2.17.1 Crystallization

The first step considered crucial in protein structure determination is the growth of diffraction quality single crystals. In the absence of any single concrete theory behind the mechanism of crystallization, we have treated the protein crystallization as a trial and error procedure invoking experience and crystallization reports as guiding principles. It is accepted that the presence of impurities, ionic strength, pH, temperature, precipitating agent and several unspecified factors play role in crystallization process. In crystallization experiments carried out by us, a precipitant such as, PEG, salt, or an organic solvent is diffused into a pure protein solution maintained at particular pH and temperature such that diffraction quality crystals grow from the solution (Haas

& Drenth, 1998). Among various crystallization techniques known, hanging drop vapor diffusion method is widely adopted and has produced more crystallized proteins than all other methods combined (Chayen, 1998). This method is simple, consumes less protein and it is easy to monitor the progress of crystallization. In a typical experimental set up using multiwell trays, 3 µl of protein solution was placed on a siliconized cover slip, mixed it with 3 µl of the precipitant solution and allowed to slowly equilibrate against 0.5 ml reservoir solution of the precipitant. For setting up crystallization concentrated solutions of pure lectin, in buffers were used. The concentrations of proteins were estimated as described by Lowry *et al.*, (1951). Initial crystallization trials of lectin did not yield quality crystals. Different precipitants such as ammonium sulfate, PEGs of different molecular weights and 2-methylpentane-2, 4-diol (MPD) were tried for crystallization. We also tried sitting drop and micro batch technique with SPL.

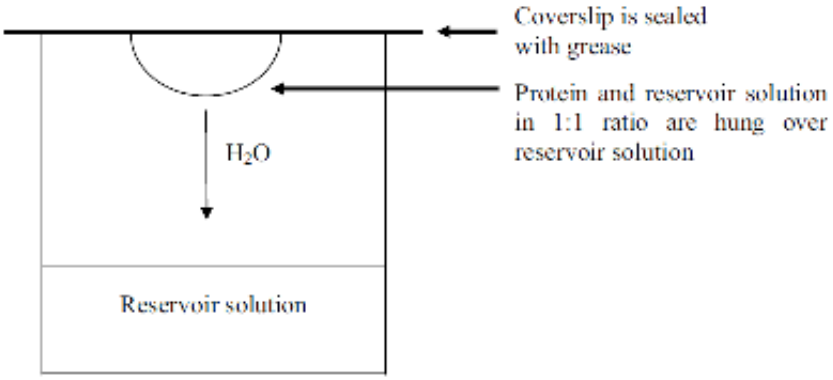

Figure 2.2: Thehanging-drop vapor diffusion method for protein crystallization (Hampton Research, 2001).

2.17.2 Data collection

After growing protein crystals of suitable size and quality, the immediate step in an X-ray structure determination is to measure the intensities of Bragg reflections. Data collection is best performed as a highly interactive process. The materials required are a crystal, an X-ray source with shutter, an area detector, and goniometer to orient and rotate the crystal. Protein crystals diffract X-rays much less than do the small molecule crystals; hence the diffraction intensity is weak. Thus protein data collection requires high intense X-ray source and high sensitivity area detector. However, in the last decade the combination of powerful tools such as, X-ray synchrotron sources, image plate and area detectors and software's used for collecting and processing the data, have transformed protein crystallography into a powerful structural tool in the area of structural biology and genomics (Beauchamp & Isaacs, 1999; Blundell *et al*., 2002). The X-ray storage-phosphor image plate (IP) is considered to be the most suitable detector for acquiring protein data with a home source, as well as at synchrotron radiation facility. IP is a very sensitive detector, with wider dynamic range; it has high spatial resolution and high count rate capacity, which are the fundamental requirements of an X-ray area detector (Amemiya, 1997). The IP has enabled protein crystallographer to obtain very accurate data sets with reduced X-ray dosage and exposure time. The plate can be erased by exposure to intense white radiation and can be used repeatedly. The IP has a radius of 350 mm mounted on a solid base and an INDY computer can control its movements through a controller box. Each collected image has a size of 2 MB, which can be stored in the computer. The crystal alignment was done through the CCD camera and TV monitor assembly. The processes

of exposure, data collection, readout and storage of data are carried out automatically via a VME-based microprocessor system.

Diffraction data were collected using MAR 345 image plate at the X –ray facility for Structural Biology at the Molecular Biophysics Unit, IISc, Bangalore, India. The X-ray beam with wavelength, 1.5418 Å CuKα was generated from rotating anode (Bruker Micro star) X-ray generator operating at 50 kv and 100 mA. Now it is routine to collect the macromolecular data at cryogenic temperatures. This technique of flash cooling protein crystals at liquid nitrogen temperatures and collecting data offers several benefits. Some of them are: reduces radiation damage of the crystal on exposure to X-rays, invariably improves the limit of resolution, decreases thermal parameters, allows storage and reuse of crystals, and helps to overcome the scaling problem by enabling the completion of entire data collection using only one crystal (Garman & Schneider, 1997). The SPL crystal was frozen under the cryostream of liquid nitrogen (100 K). During data acquisition the crystals were oscillated about an axis perpendicular to the X-ray beam, with a chosen, relatively small angle of oscillation of 1° per frame. Crystal to detector distance chosen was 200 mm, based on the longest unit cell dimension, mosaic spread, etc., so that the intensity spots are well resolved, is approximately equal to the longest crystal cell dimension. The exposure time depends on the quality of crystal and oscillation range; larger the oscillation range, longer the exposure time required. However, the situation where the intensities crossing the limit of the dynamic range of image plate has been avoided. One way to minimize such errors is to apply more than one oscillation of phi (φ) per image for exposure times greater than 600 sec. In our data collection, an exposure time was used 180 sec per frame and oscillation 0.5^{o} per frame.

2.17.3 Data Processing

Indexing, processing, scaling and merging of data were carried out by MOSFLM and SCALA of CCP4 (Collaborative Computational Project, Number 4, 1994) suit program. MOSFLM can process diffraction image from a wide range of detectors and produces output as an MTZ files of reflections indicates with their intensities and standard deviation. This MTZ files is then pass in to other programs of the CCP4 program suit for further data reduction. Auto indexing will be carried out by MOSFLM using spots from selected images.

SCALA merge multiple observations of reflections and produce a file that contain averaged intensities of each reflection. SCALA require an input MTZ files containing merged intensities, such as that produced by MOSFLM. This program is used to analysis scales and B factor, R merge, completeness, multiplicity, correlation within and between data set can be done. The program TRUNCATE in the CCP4 suit programs(CCP4, 1994) used to convert a file of averaged intensities to a file containing structure factor amplitude (F) and the original intensities using the procedure of French and Wilson. The structure factor amplitude thus calculated were all positive and follows Wilson's distribution.

2.17.4 Mathew's number

Once the space group and unit cell dimensions of the crystal are known it is possible to estimate the number of molecules in the crystallographic asymmetric unit and the solvent content of the protein crystals with the knowledge of the molecular weight of protein. The following equations are used (Matthews, 1968).

Vm = Unit cell volume / (Mol.Wt.X nX z)

Vsolv = 1- (1.23/ Vm)

Where Vm is the Mathew's number, n is the number of molecules per asymmetric unit and z is the Avogadro's number; (Vsolv) is the solvent content of protein crystals. The Matthew's number and the solvent content were calculated for the SPL crystal form. The solvent content of the SPL crystal was calculated by assuming a dimer molecule in the asymmetric unit.

2.17.5 Structure solution

Multi-wavelength Anomalous Dispersion (MAD), Multiple Isomorphous Replacement (MIR) and Molecular Replacement (MR) are the three methods widely used in protein crystallography to solve structures. MAD technique is currently most popular owing to the recent technical advances made in the field of synchrotron radiation, image plate and CCD detectors for data collection and molecular biology techniques that provide selenium derivative for any protein (Ealick, 2000). MIR technique requires more than one heavy atom derivative.

2.17.6 Molecular replacement (MR)

This technique is the simplest of all and can be used when a homologous protein with structural similarity is available in the database. The pioneering studies of Rossmann & Blow, 1962) laid the foundation of Molecular Replacement method (Rossmann, 2001). Owing to the rapid expansion of Protein Data Bank (PDB; Berman *et al*., 2000) with the increase in the number of models available, MR method is now routinely used in protein structure determinations.

The success of molecular replacement is critically dependent on the quality of the search model. Several model-preparation procedures are integrated in the molecular-replacement program MOLREP (Vagin &

Teplyakov, 1997). These include model modification on the basis of amino-acid sequence alignment and model correction based on analysis of the solvent-accessibility of the atoms. MOLREP is an automated program for MR (Vagin & Teplyakov, 1997) where, along with the default protocol, there are various search strategies as options. The program is part of the CCP4 package. AMoRe (Navaza, 1994), Phaser (Read, 2001), the MR implementation in CNS (Brünger *et al.*, 1998) and MOLREP together cover more than 95% of structures solved by MR. A special feature of MOLREP is that it offers several built-in model-preparation functionalities. The integration of model-preparation and Patterson function techniques in one program has several advantages. Apart from convenience, such integration allows specific adjustment of the model-modification parameters for an efficient Patterson search. Moreover, the weighting parameters for the rotation function (RF) and translation function (TF) functions are more reliable if they are derived from the original sequence and atomic coordinates of the homologous protein. Such an integrated approach has proven to be efficient and has recently been implemented in several MR pipelines including BALBES (Long *et al.*, 2008), MrBUMP (Keegan & Winn, 2008) and JSCG (Schwarzenbacher *et al.*, 2008), in which MOLREP itself is used as a component. Currently, MOLREP is being updated to fulfill the requirements raised by BALBES development and benefits from its training. n this subsection, it is assumed that the user has provided an input file containing the sequence of the target protein. In this case, the first stage of model preparation in MOLREP includes alignment of the sequence derived from the search model to the target sequence and, provisionally, deletion of residues and atoms of the search model that do not map on to the target sequence. This is a conservative approach to model correction, as no new atoms are added

and the coordinates of preserved atoms are not changed. Besides, the sequence identity derived from the alignment of the two sequences is further used for weighting of the X-ray data.

The sequence alignment implemented in MOLREP is a modified version of the dynamic alignment algorithm (Needleman & Wunsch, 1970), which takes into account the known three-dimensional structure of the search model. Thus, buried residues contribute to the total alignment score more than residues at the surface. In addition, it is assumed that gaps and insertions are impossible within sequence segments corresponding to helices and strands. Producing a search model is only one aspect of model preparation. Another aspect is to define the weighting scheme that is most suitable for a given model. During model preparation, MOLREP estimates a number of parameters for the search model. Two of them, the radius of gyration of the model and its sequence identity with the target protein, are translated into the parameters of Gaussian low-pass and high-pass filters (see, for example, Gonzalez & Woods, 2002), which define defining structural details of finer scale than the r.m.s.d. between the search and target molecules. The exact value of this r.m.s.d. is the treatment of translational NCS, otherwise known as pseudo-translation, is a special case of model modification because it is applied to the TF but not to the RF. In MOLREP, this modification is applied implicitly in reciprocal space and therefore can be considered as either model modification or as weighting of the X-ray data. Translational NCS is detected and the NCS translation vector is derived using the experimental Patterson function. MOLREP assumes that translational NCS is present if there is a non-origin peak in the Patterson function with a height of 1/8 or more of the origin peak height. An additional requirement is that this peak is sufficiently distant from the origin (three-quarters of the

diameter of the search model) to ensure that it is not caused, for example, by regular structural patterns in the target molecule. Such an approach is simple, works in most cases and is therefore used by default. However, neither false positives nor false negatives can be excluded.

2.18 Structure refinement

2. 18.1 Refinement and model building

An initial model built into an experimental map of approximately phased molecular replacement solution, will usually contain many errors. Generating the correct structure from structure factors and initial phases is composed of two parts. The first part is the computational refinement in which the difference in observed and calculated structure factors (R *factor*) and differences between observed and optimal stereo chemical, electrostatic and van der Waals contacts are minimized. The second part is manual intervention. Computational refinement tends to converge in a local minimum of the total energy landscape. Manual intervention allows for changes to the structures that are outside the radius of convergence, r .Refinement of a protein structure involves minimizing a function involving observed (Fo) and calculated structure factors (Fc) and along with applying restraints such as minimizing the difference between the refined parameters and ideal parameters of the stereo chemistry (Drenth, 1994). Refinement techniques generally fall into two categories, depending upon whether the calculations are performed either in the real space or in the reciprocal space. Refinement methods based on reciprocal space is preferred over real space because the former ones are computationally less expensive. To prevent the model from going into local minimum, interactive graphics was used for checking the fit of the model to the electron density. To monitor the progress

of refinement the crystallographic parameter called R factor is used. This is defined as,

$$R = \frac{\sum_{hkl} ||Fo(hkl)| - |Fc(hkl)||}{\sum_{hkl} |Fo(hkl)|}$$

Where h, k, l, are the Miller indices of Bragg reflections, and the summation is over all the reflections.

2.18.2 CCP4 program suite

The CCP4 program suite is the program package most widely used by X-ray crystallographers in structure determination and analysis of a mcromolecules. The CCP4 suite is an integrated set of programs for protein crystallography developed by close collaboration of crystallographers under an initiative by the UK Biotechnology and Biological Sciences Research Council (formerly the ERC). Some software developed elsewhere is also included. The CCP4 suite contains programs for all aspects of protein crystallography, including data processing, data scaling, Patterson search and refinement, isomorphism and molecular replacement, structure refinement, phase improvement and density modification, and presentation of results.

2.8.3 Rigid body refinements

Rigid body refinement is the first step in a protein structure refinement procedure that fixes the gross features of a molecule for further refinement (Head-Gordon & Brooks, 1991). The adjustment of the model consists of refining the three positional parameters and one temperature factor of all the atoms in the structure except hydrogen atoms. This procedure minimizes the R factor value by refining three rotational and three translational degrees of freedom of the subunits or groups. It is possible to regard the entire molecule as a rigid entity and refine its position and orientation in the unit cell. Each molecule or subunit is treated as a continuous mass distribution located at the center of mass position defined by,

$$R_J = (1/M_J) \sum_i m_i r_i$$

Where $M_J = \sum_i m_i$

The m_i's are the atomic masses and J labels the rigid bodies. Positional refinement was done using constraints and restraints. The ratio of the number of parameters to be determined to the number of observations is very high in a protein structure determination. This ratio could be improved in two ways, one is by reducing the number of parameters by the use of constraints, and the other by increasing the number of observations by the addition of restraints. The stereo chemical information such as bond lengths, bond angles, conformational angles, planarity etc., obtained from small molecule structures, are used for applying restraints and constraints. Restraints are considered when a specific parameter has restricted freedom, limited to a range of values, whereas, the parameter is constraint, when it can assume only a specific value.

2.18.4 Restraint refinement

During the course of refinement of macromolecules some groups of atoms may have to be constrained or restrained to improve the ratio of observables to parameters. As rigid bodies, or, restrain or constrain the bond lengths, bond angles, non crystallographic symmetry (NCS) and atomic positions to a desired value by use of appropriate force constants. Restraints are given when limited freedom can be given for a parameter. When a parameter has to be held to an exact value, then it is constrained. In practice, a constraint is a restraint with infinite force constant. In NCS symmetry restraints,the molecules in the asymmetric unit are superposed by least squares superposition and the average coordinates (x_{av}) of individual atoms are computed. If x represents the coordinates of individual atoms, then each atom can be restrained according to the mathematical term:

$$E_{NCS} = w(x-x_{av})^2$$

The corresponding B-factor restraints are given by:

$$B_{NCS} = (b-b_{av})^2/\sigma^2_{NCS}$$

Where w is a weight function, b and b_{av} are the respective individual and average temperature factors of NCS related atoms and σ_{NCS} is the target deviation for B-factor restraints.

2.18.5 Omit maps

Composite 2Fo-Fc omit map using OMIT program of CCP4 program suite was computed to remove the model bias during each iteration of model building process. The composite omit map of the entire model in the crystallographic asymmetric unit was calculated by omitting 20 amino acid residues of the current model successively at a time and then building the rest of the model in the omit density. Composite omit maps were thus calculated for the entire macromolecule under study and the model was built into the omit density.

2. 18. 6 Refinement by maximum-likelihood method

The initial models of SPL from MR calculations were refined using the program REFMAC (implemented in CCP4) which makes use of maximum-likelihood equations. REFMAC (Murshudov *et al*., 1997, 1999) is a macromolecular refinement program which has been integrated into the CCP4 suite (Section 25.1.2.4). Read (1990) and Bricogne (1991) have suggested a maximum-likelihood target that should be based on various probability distributions. One of the expected advantages of maximum likelihood refinement is a decrease in refinement bias, as the calculated structure-factor amplitudes will not be forced to match the observed amplitudes (Read, 1997). Use of appropriate likelihood targets through the incorporation of the effect of measurement of error and the use of cross-validation data to estimate the σ (sigma) values are the key ingredients in the likelihood refinement. Verification tests have shown that for refinement, maximum likelihood method is more than twice as effective compared to least-squares method, in

improving the model (Pannu & Read, 1996). REFMAC program can carry out rigid body restrained or unrestrained refinement using X-ray data (Murshudov *et al*., 1997) or idealization of a macromolecular structure. There are options to use different minimization methods. The program minimizes the coordinate parameters to satisfy a maximum-likelihood or least squares residual. REFMAC also produces an output file with extension MTZ (named after three of its progenitors, McLaughlin, Terry and Zelinka) containing weighted coefficients for s a weighted mFo-DFc and 2mFo-DFc maps. About 5% of the reflections were kept aside during refinement to calculate R free for cross validation (Brunger, 1992). NCS averaging is effectively done for an asymmetric unit composed of N similar objects related by non-crystallographic symmetry (NCS). Rossman and Blow (1963) proposed this method, by which the current phases of reflections can be improved by averaging over the electron densities of NCS related objects. NCS averaging requires an accurate estimate of NCS operators and exact information on the position and shape of the objects whose density has to be averaged (Vellieux & Read, 1997). Since lectin structures had more than one molecule in their asymmetric units, their initial models were refined taking advantage of the presence of non-crystallographic symmetry. The NCS restraint was removed and subunits refined independently in the final cycles of refinement.

2. 18.7 Graphics, visualization, electron density maps and model building

After phasing the reflections obtained from X-ray diffraction, using one of the methods discussed above, an electron density map is calculated using

Fourier transform. The formula for the Fourier summation to calculate an electron density map is:

$$\rho(xyz) = (1/V_c) \sum_{h}\sum_{k}\sum_{l} | F | \cos [2\pi (hx+ky+lz) - \alpha\, h\, k\, l$$

Where x, y, z are the fractional coordinates of each point in the unit cell, F (hkl) is the structure factor, V is the unit cell volume and α (hkl) is the phase angle. The h, k, l, are Miller indices. The difference maps such as 2Fo-Fc and Fo-Fc for electron density are used to identify errors in the model structure and to refine the positional and displacement parameters. Generally the deviations of the model from the reference molecule could be detected in the Fo-Fc difference map. The maps were contoured at 1s for 2Fo-Fc and 3.0 σ for Fo-Fc. Here s refers to r.m.s. deviation in the mean density in electrons/$Å^3$. Poorly defined regions of the map were examined with the maps contoured at lower levels. Difference electron density maps are important for locating bound ligands in protein structures (Glusker *et al.*, 1994). The graphic software program COOT was used for displaying and examining the electron density maps, for displaying atoms, for interactive fitting and optimizing the geometry and also for solvating the structures. The displayed map and the molecule could be rotated and viewed from any direction. Each residue starting from N-terminal to the C terminal was examined for their optimum fitting in the electron density maps and corrects geometry. The deviating ones were corrected using Real Space Refinement. A difference Fourier (Fo-Fc) map was calculated to identify the deviations of the protein from its search model. Both the aligned sequences of the respective families and the observed difference densities were used to ascertain the identity of dissimilar residues. After every cycle of visual fitting of the model to the calculated electron density, it was subjected to several cycles of refinement using REFMAC. After the refinement the changed residues were carefully checked in the new map.

The fitted residue was retained if no difference in density (Fo-Fc) was observed and R factor and R free improved. This procedure was repeated a number of times till all the observed difference densities in the vicinity of protein atoms were accounted and both R factor and R free values converged. The solvent molecules were placed wherever the Fo-Fc density was observed above 3 σ level and when the water molecules made reasonable hydrogen bonds. To add water molecules, difference map peak analysis was carried out and the map above 5 σ were observed and manual water addition were carried out for better fitting of water molecule. It is further analyzed by CHECK WATER in the validate panel and the water molecule with close contact less than 2.3 σ and greater than 3.5 σ and map level 1 σ were either removed or arranged in space to get the exact fitting. Coot water was used for adding water molecules to each of the structures through interactive mode. The addition of water molecules was started only when the refinement of the model reached an R factor below 25%. The water search was limited to distance of 5.0 Å from protein, and the distance between any two waters not less than 2.5 Å. Initially water molecules were added at a higher sigma level, later when the structures were refined; this progressively was reduced to lower levels, up to 3.5 σ.

2. 18.8 Analysis and comparison of structures

The program PROCHECK was used for checking the stereochemistry and quality of the model. This program is written by Laskowski (Laskowski *et al*., 1993) and forms part of the CCP4 suite of programs. Once the model has improved, after every cycle of model fitting and run of REFMAC it was subjected to stereo chemical quality check. The program compares and assesses the quality of the model with the available structures of similar or

better resolution than the reference structure. The output contains a comprehensive residue by residue listing of the parameters and their graphical representation. The program highlights the regions of the structure a criterion, the angle C-O….O greater than 90 was considered are unusual. The r.m.s deviations in Cα positions on superpositioning of SPL with different galactose specific legume lectins were conducted. The lectins used for the superpositioing were *Dolichos biflorus* seed lectin with adenine (1BJQ), *Dolichos biflorus* vegetative lectin (1G8W), Soyabean agglutinin 2, 6 pentasacchrides (2SBA), *Phaseolus vulgaris* (1FAT), and *Robinia psuedoacacia* bark lectin (1FNY).

2.18.9 Hydrogen bonds

Hydrogen bonds were calculated using the CONTACT program of CCP4 suite. Hydrogen bonded contacts were considered for distances less than 3.5 σ between donor and acceptor and angle N-H….O (type) greater than 120^{o}. Similarly for O.…O contacts, in addition to the distance criteria the angle C-O…O greater 90º was considered. The NACCESS program (Hubbard & Thornton, 1993) was used to calculate the atomic accessible surface are defined by rolling a probe of given size around a van der Waals surface. This program is an implementation of the method of Lee and Richards (1971). The buried surface area calculation due to oligomerization had also been done with the same program. The solvent molecule in the primary solvent shell was calculated accoeding to C. Sadasivan *et al., (*1998).

2.18.10 Figures

Most of the figures in the thesis were prepared using MOLSOFT ICM Pro 3.4-9 and surface representation of SPL tetramer was prepared using YASARA (Yet Another Scientific Artificial Reality Application, Elmar KrIeger 2003).