CHAPTER 3

Biophysical, Structural characterization and exploring the potential binding site of a lectin (PotHg) from tubers of Solanum tuberosum

3.1 Summary

The secondary structure analysis of PotHg was done by circular dichroism studies. Analysis of far-UV CD spectrum by the software CDpro showed secondary structure composition as 1.73% α-helical content (regular and distorted), 43.93% β-strand content (regular and distorted), 19% turns and 33%unordered structure. CD-melt analysis showed the stability of PotHg up to 75°C, which is consistent with temperature stability studies. Near-UV CD spectrum of PotHg was showed peaks around 285 and 291 nm which implies a compact and rigid structure. Fluorescence spectroscopy was done to analyze the sugar binding activity of PotHg. The intrinsic fluorescence of purified protein was quenched upon titrating with mannose and galactose although these sugars did not inhibit the hemagglutination activity of PotHg. To study the environment of surface exposed tryptophan in PotHg, solute quenching studies were carried out at pH 5.0, 7.0 and 10.0 using quenchers such as acrylamide, potassium iodide, and cesium chloride. The results indicated that the surface exposed tryptophan environment is electropositive in nature, and change in pH affected the environment of a tryptophan residue. Analysis of far-UV CD spectrum of PotHg was showing a positive peak around 225 nm which is common among many protease inhibitors and lectins. The addition of sugar change the intensity of CD spectra, but it did not change the peak around 225 nm. Preliminary crystallization and X-ray diffraction studies are also reported for PotHg in this study.

3.2 Introduction

Three-dimensional structure of protein, which is responsible for its functions is the sum of many factors such as hydrogen bonds, ionic, hydrophobic, Van der Waal's interactions, covalent interactions and disulfide linkages. Changes in native conformations of proteins upon ligand binding or any other external environments changes such as pH or temperature can be followed by using fluorescence and circular dichroism spectroscopy. Single molecule X-Ray crystallography is used to obtain the three-dimensional structure of a protein and its interaction with ligands. In this chapter, we report the analysis of CD spectroscopy study results for investigating the secondary structure elements of PotHg. Also, crystallization and diffraction of obtained crystals have been described in this chapter.

Studies reported in Chapter 2 of this thesis showed the presence of a Kunitz-type protease inhibitor with lectin-like activity in tubers of potato. This is the first report of such dual functional protein from potato. PotHg showed hemagglutination activity in pH range 4-9 and temperature up to 80 C and studies on inhibition of hemagglutination activity suggested that it has an affinity for fetuin. Chemical modification of tryptophan residue by NBS rendered purified PotHg inactive, suggesting a direct or indirect role of tryptophan residues in carbohydrate recognition and binding. Molecular docking studies using Autodock tools and simple sugars such as mannose and galactose also showed the role of tryptophan in carbohydrate binding. The studies on the conformational stability of this protein in various conditions by making use of fluorescence spectroscopy and circular dichroism have been described in this chapter. Saccharide binding, as well as the accessibility of tryptophan residue of the lectin, has also been studied. Since hemagglutination assay is qualitative in nature and has many limitations, these can be overcome through many accurate and quantitative methods to calculate the exact thermodynamic parameters involved in sugar-lectin interactions.

The aromatic amino acids (Tryptophan, Tyrosine, and Phenylalanine) in protein molecules get excited upon absorbance of energy. Since the residues in higher energy states are unstable, they come down to its ground stated which is accompanied by the release of excess energy through fluorescence. All the three aromatic amino acids present in a protein sequence have specific spectral properties. Tryptophan residue can be specifically irradiated at 295 nm, which allows study involving fluorescence alone from tryptophan residues. The intensity and the wavelength of maximum fluorescence emission of tryptophan are highly dependent on the polarity of the environment surrounding it. Hence by studying the tryptophan fluorescence, the conformational changes in the protein can be monitored.

Side chains of tryptophan residues are very sensitive to their local environments, and they can be monitored through fluorescence spectroscopy studies to obtain information about protein structure and conformation. Tryptophan residue can be specifically irradiated at 295 nm, which allows study involving fluorescence spectroscopy. In chapter 2 of this thesis, we showed the role of tryptophan molecule in sugar recognition and interaction by modifying it and to confirm this further we have carried out carbohydrate binding properties of PotHg through fluorescence spectroscopy. We also tried to understand the tryptophan environment at different pH using steady state quenching of PotHg.

Fluorescence quenching studies can be used to monitor the environment around aromatic residues in proteins. The interaction between a fluorophore and a molecule induces perturbation or modification in the fluorescence parameters like intensity, quantum yield and lifetime. Two types of fluorescence quenching can be observed when the interaction takes place between the fluorophore and the quencher molecule. Collisional quenching occurs when the fluorophore and another molecule diffuse in the solution and collide with each other. In this case, the two molecules do not form a complex. In static quenching, on the other hand, two molecules bind one to the other forming a complex (Albani, 2004). The most commonly used quenching molecules are acrylamide, succinimide (neutral), iodide (cationic) and cesium (anionic) and acrylamide, being a small uncharged molecule, can diffuse within a protein and can quench the fluorescence of buried tryptophans as well. Iodide and cesium ions, on the other hand, quench the fluorescence of tryptophans present at or near the surface of the protein. Being charged ions, their quenching efficiency also depends on the charge surrounding the tryptophan (Albani, 2004).

Circular dichroism is a very sensitive method to obtain the protein secondary structure. Circular dichroism spectroscopy measures the differences in absorption of the left and right-handed polarized light by molecules which emerge due to structural asymmetry. This property is characteristic of all the optically active molecules such as carbohydrates and amino acids. Using CD spectra of proteins in the far-ultraviolet region (180-250 nm), the secondary structure elements present in the protein can be estimated. Secondary structure elements in proteins (α helix, β sheet, double helix, etc.) also give rise to characteristic CD spectra. Far-UV CD spectra can also be used to follow the conformational changes in the protein. At these wavelengths, peptide bonds in the protein act as chromophores and the signal produced by them is characteristic of the secondary structure. In the near UV region (250- 300) nm disulfide bonds and the aromatic amino acids acts as chromophores. Tryptophan gives a characteristic signal in the region from 280-300. While phenylalanine gives a signal in the region from 250-270 nm and tyrosine gives a signal in the region

from 270-290 nm. The signal arises in the near-UV region are sensitive to the overall tertiary structure of the protein (Kelly and Price, 2000).

3.3 Crystallization and X-Ray diffraction

The three-dimensional structure of a protein provides a detailed and often unique information regarding protein functions and crystallization of protein molecule is the first step in solving the three-dimensional structure by X-Ray diffraction method. To date > 80% of protein data bank entries are macromolecule structure that was defined using X-ray diffraction. One of the foremost conditions for obtaining the macromolecule crystals to carry out the X-Ray diffraction studies is protein homogeneity, and crystallization experiments need the protein to attain the supersaturation state without causing aggregation or precipitation. It represents a reversible equilibrium phenomenon due to the minimization of the free energy of the system (Weber, 1991). In the supersaturated conditions, the system is no longer at equilibrium, and it will be thermodynamically driven toward a new equilibrium state with a corresponding new free energy minimum. The supersaturated molecules lose rotational and translational freedom by forming many new stable noncovalent chemical bonds, thus minimizing the free energy of the system. The four main steps of protein crystallization are: (1) The protein should be pure (2) it should be dissolved in a suitable solvent which helps the protein to be precipitated by a salt or an organic compound (3) the solution is brought to supersaturation state (4) After the formation of nuclei, the actual crystal growth can begin.

In a past number of techniques has been used to bring the protein into a supersaturation tests. The three main method of carrying out crystallization are vapor-diffusion method, micro batch method and dialysis method (Ducruix and Giege, 1992). Vapor diffusion method is one of the most successful methods for protein crystallization. In this technique evaporation and diffusion of water between solutions of different concentration and the protein molecules achieves the supersaturation state and can be carried out in two ways: hanging drop method & sitting drop method. In this study, we have carried out the crystallization in micro-batch plate under oil method. In this method the protein sample and precipitant solution are dispensed into the well of a plate and the well is covered with paraffin oil or other oils to minimize the evaporation rate.

3.3.1 Data Collection and Processing

For data collection on X-ray beam at room temperature, the crystal is kept in a thin-walled capillary made up of a glass of quartz. Data collection at room temperature causes the radiation

damage to the crystal due to high energy X-Ray. To prevent the radiation damage data can be collected at a very low temperature of 100-120 K. This temperature is obtained by the continued stream of cold nitrogen gas and complete data set can be obtained without damage to the crystals. One of the drawbacks of data collection at this temperature is the formation of ice ring which interference in the data collection and subsequent data processing. X-ray diffraction data are collected using Cu-Kα X-rays mostly, and this particular X-ray source is chosen due to its wavelength of 1.54 Å, which is similar to the bond length between the C-N and C-O. The crystals are oscillated through an angle of 0.5 to 1° about an axis perpendicular to the direction of the X-ray beam, and the diffraction data is collected on image plate and further processed. There are seven major steps present in the analysis and reduction of single crystal data diffraction. Visualization and preliminary analysis of the unprocessed detector data, indexing of the diffraction pattern, refinement of the crystal and detector parameters, integration of the diffraction maxima, finding the relative scale factors between measurements, precise refinement of crystal parameters using the whole data set, and merging and statistical analysis of the measurements related by space-group symmetry

The first three steps are carried out with various programs like MOSFLM, DENZO, and SCALEPACK present in CCP4 suite (Battye et al., 2011; Otwinowski, Z and Minor W., 2006). MOSFLM is a program for integrating single crystal diffraction data from area detectors, and it assumes that the experiment is conducted using the Arndt-Wonacott oscillation method. Mosflm can process diffraction images from a wide range of detectors and run all the images into a data processing server. It is designed to simplify data processing by dividing the process into a series of steps such as indexing, strategy calculation, cell refinement, integration, and history, which are normally carried out sequentially. It produces, as output, an MTZ file of reflection indices with their intensities and standard deviations.

3.4 Material

Acrylamide, potassium iodide, cesium chloride, sodium thiosulfate, galactose, mannose, were purchased from Sigma, USA. All other chemicals used were of analytical grade. Quencher solutions (acrylamide, potassium iodide, and cesium chloride) were prepared as 5M stocks in deionised water. Potassium iodide solution was made in 200μM sodium thiosulfate solution to prevent the formation of triiodide (I-3) ions. Sugar solutions were prepared as 1000mM stocks. Buffers were prepared as 1M stock solution and pH was adjusted at 25^0C. Protein solutions were

centrifuged at 8000 x *g* for 5 minutes before use. The software Microcal Origin 6.1 was used to plot the graphs and do the related calculations

3.5 Methods

3.5.1 Purification of PotHg

PotHg was purified as explained in section 2.8.1 of chapter 2 of this thesis. The concentration of protein was determined by taking the absorbance reading at 280nm.

3.5.2 Fluorescence measurements

The intrinsic fluorescence of the protein was measured using a Hitachi F-7000 FL fluorescence spectrophotometer. The samples were kept in a quartz cuvette, at a room temperature. To eliminate the background emission, the signal produced by either buffer solution, or buffer containing the appropriate quantity of carbohydrate was subtracted from final sample readings. The protein solution (~0.1mg/ml in 20 mM Tris-HCl pH 8.0) was excited at 295nm, and the emission was recorded in the range of wavelengths 300-400nm. Each spectrum was an average of 3 scans, and both the excitation and emission spectra were obtained setting the slit width at 5 nm, and speed 1200 nm min^{-1}.

3.5.2.1 Carbohydrate binding studies

The binding of sugars to PotHg was studied by titrating the protein solution against sugar solutions and monitoring the fluorescence using fluorescence spectroscopy. The temperature was maintained at 30°C throughout the experiment. Titrations were performed by adding aliquots (4µl at each time) of test sugars from stock (250mM) in Tris-HCl (pH 7.5) to 1ml of purified protein (0.1mg/ml) in the same buffer. Samples were mixed thoroughly with pipetting, making sure that no bubble formation occurred, Corrections were also made to compensate the dilution effect upon addition of sugar solution. At the highest concentration of the saccharide, the volume change was less than 5% of the solution in the cuvette. Each data point was an average of three independent sets of experiments, and the following equation was used to determine the association constant (Ka) by the reported method (Chipman et al., 1967).

$$\log [C]f = -\log[Ka] + \log [(F0-FC)/(FC-F\infty)] \quad \text{(Eq. 3.1)}$$

Where C is Concentration of carbohydrate, F0 is Fluorescence intensity of protein alone; Fc is Fluorescence intensity of protein in the presence of carbohydrate and F∞ is Fluorescence intensity of protein at ∞ concentration of carbohydrate. From the ordinate intercept of the double reciprocal plot of F0/(F0-FC) versus 1/[C], where F0 and FC are the fluorescence intensities of the free protein and the protein plus sugar concentration [C], F∞, the fluorescence intensity upon saturation of all the sugar binding sites is obtained. In the plot of log[(F0-FC)/(FC-F∞)] versus log[C], the abscissa intercept yielded the Kd value (the dissociation constant) for the protein-sugar interactions, which is the reciprocal of Ka (the association constant). The free energy changes of the Association (ΔG) were calculated by using the equation:

$-\Delta G = RT \ln(Ka)$ (Eq. 3.2)

3.5.2.2 Solute quenching studies of PotHg

Fluorescence quenching experiments of PotHg were carried out at 25 0C. Purified protein samples in different buffers (acetate buffer pH 5.0, phosphate buffer pH 7.0 and glycine-NaOH buffer pH 9.0) were titrated with aliquots (10µl) of quenchers (acrylamide, cesium chloride, and potassium iodide) from a 5M stock solution in Tris-HCl pH 8.0 and fluorescence spectra were recorded after each addition and fluorescence intensities were corrected for dilution. To prevent the tri-iodide formation, sodium bisulfate (200µM) was added to potassium iodide stock solution. Protein excitation was done at 295nm and emission spectra were recorded between 300 and 400nm, at room temperatures.

Quenching data for all these quenchers were analyzed by the Stern-Volmer and modified Stern-Volmer equations (Lehrer 1971):

$Fo/F = 1+Ksv [C]$; (Eq. 3.3)

$Fo/\Delta F = fa^{-1} + 1/(Kafa[C])$ (Eq. 3.4)

Fo and F are fluorescence intensities in the absence and presence of quenchers

[C] is the quencher concentration

Ksv is the Stern-Volmer constant

$\Delta F = (Fo-F)$ is the change in fluorescence intensity at any point

fa is the fraction of fluorophore available to the quencher.

The slope of a plot of Fo/F versus [C] gave values for Ksv and the slope of a plot of Fo/ ΔF versus 1/[C] gave values for 1/Kafa and Y intercept of this slope gave values for 1/fa. From this values of Ksv, and % Fa was calculated. % Fa shows the availability of fluorophore to be quenched by the quencher.

3.5.3 Circular Dichroism Spectroscopy

To record the CD spectra, a Jasco J-810 (Jasco corp., Tokyo, Japan) spectropolarimeter equipped with Peltier thermostat was used. Two concentrations of purified protein (5μM and 25μM) in 10mM sodium phosphate buffer (pH 7.5) were placed in a 2mm path length rectangular quartz cell and spectra recorded at a scan speed of 20nm/min with a response time of 4s, a slit width of 2nm and the spectra recorded for the far-UV region (250 to 200 nm) and the near-UV region (300 to 250nm), respectively. Analysis of the CD spectrum was done using three different methods, namely CDSSTR, CONTINLL and SELCON3 (Sreerama et al., 1999; Johnson, 1999; Provencher et al., 1981) and employing the software routines available at CDPro. A basis set containing 43 proteins was used as a reference for fitting the experimental spectrum. To investigate the effect of carbohydrate binding on the secondary and tertiary structure, the spectra were recorded in the absence as well as in the presence of 1mM galactose. The structural changes that occurred in the protein during thermal unfolding of the protein were determined by recording the CD spectra between 25°C and 90°C, at a scan rate of 1°C/minute and by monitoring the changes in ellipticity (Θ) at 225 nm, corresponding to highest positive peak in the far-UV region. Ellipticity can define as a unit for CD data which is related to absorbance by a factor of 32.98. Ellipticity is usually reported in millidegrees. And is plotted against wavelength in general CD diagram.

3.5.4 Crystallization and data collection

3.5.4.1 Siliconization

The glassware and coverslips used in crystallization trials were siliconized before carrying out the crystallization. Siliconization helps in converting the glass surface into extremely hydrophobic in nature; this prevents interactions between surface and buffer or proteins. The glassware was first washed with liquid soap, washed with free flowing water and was air dried in hot air oven at 80°C. after drying the glassware, 10ml of carbon disulfide mixed with 1ml of trimethylchlorosilane was added to the glassware, so the mixture covers the entire surface of glassware. After sill iconization glassware were washed thoroughly with methanol and then with double distilled water and finally dried in hot air over at 80° C.

3.5.4.2 Crystallization trials of PotHg

The initial crystallization trials were carried out using Screen 1 (Crystal Screen TM-HR2-110) and Screen 2 (Crystal ScreenTM-HR2-112), Index and Index HT crystallization buffers obtained from Hampton Inc. Additionally, crystallization screens of conditions in which homologous proteins were crystallized were also screened. The different buffer conditions were set up in the micro batch 60 wells plate acquired from Griener biologics. The plates were filled with total of 8ml of mixture of paraffin and silicone oil in ratio of 1:1 and the protein (concn. in buffer from 5 mg/mL to 40 mg/mL) and different buffers were added in the wells in different ratio of (1:1,1:2, 1:3) generating a wide range of different conditions for crystallization. Plates were regularly monitored under a polarizing microscope and promising conditions further optimized to get better crystals.

3.5.4.3 Data Collection

The initial X-ray diffraction data were collected at cryotemperature (100K) using Cu-Kα radiations from Rigaku rotating anode X-ray generator, and the images were collected on an R-Axis IV++ image plate detector. The Crystal-to-detector distance was adjusted to 200mm, and each oscillation frame was recorded with a 0.5º oscillation per frame. The exposure time for collecting each diffraction data was of 5 minutes and the image frames manually checked for diffraction quality.

3.5.5 Surface Analysis and Molecular docking studies of PotHg

To explore the possible tryptophan involved in carbohydrate binding, surface analysis of PotHg was carried out in order to find the surface exposed tryptophan and the surrounding environment. Computed Atlas of Surface Topography of proteins (CASTp) available at (http://sts.bioe.uic.edu/castp/calculation.php) was used to generate this information (Dundas et al., 2006). Chimera was used for molecular visualization and superposing the structures (Pettersen et al., 2004). Molecular docking of the ligand, in the putative binding site, was done using AutoDock tool (ADT) (Morris et al., 2009). Intermediary steps, such as pdbqt files for protein and ligands preparation and grid box creation were completed using Graphical User Interface program ADT. ADT assigned polar hydrogens, united atom Kollman charges, solvation parameters and fragmental volumes to the protein. AutoGrid was used for the preparation of the grid map using a grid box. The grid size was set to 40 × 38 × 40 XYZ points, respectively, and grid center was designated at dimensions (x, y, and z): -7.358, 0.45 and 16.096, respectively. Autodock employed iterated local search global optimizer and during the docking procedure, both, the protein and ligands were

considered as rigid. The pose with the lowest energy of binding or binding affinity was extracted. Pymol was used for viewing hydrogen bonding of ligand with protein. The coordinates of galactose and mannose were downloaded from ZINC database (http://zinc.docking.org/)

3.6 Results

3.6.1 Carbohydrate binding studies

Upon excitation at 295nm PotHg sample gave an emission spectrum centered at 342nm (Figure 3.1 A). Titration of the protein with galactose and mannose resulted in a decrease in the emission intensity by 30% & 24% respectively with no change in maxima (Figure 3.1 and Figure 3.2), although these sugars were unable to inhibit the hemagglutination activity of PotHg. In figure 3.1A spectrum one corresponds to the protein alone and spectra 2-8 correspond to those recorded in the presence of increasing concentrations of galactose. Similarly, in figure 3.2B spectrum one corresponds to the protein alone and spectra 2-7 correspond to those recorded in the presence of increasing concentration of mannose. A plot of the change in fluorescence intensity, ΔF (= Fo – F) as a function of the added ligand concentration is shown in these figures, and this represents the binding curve for the titration.Double logarithmic plots (Figure 3.1 C and Figure 3.2 C) were near unity for the ligands used, indicating one binding site per monomer of lectin. The affinities of PotHg towards galactose and mannose are considerably low (Ka= 357 and 667M^{-1} respectively).

The corresponding ΔG values for binding of galactose and mannose to PotHg are -16.57 and -15 kJ mol^{-1}, respectively. These results indicated the spontaneous nature of these bindings.

Figure 3.1 Fluorescence spectra of PotHg in the absence and presence of galactose (A)Represents fluorescence quenching of PotHg on the addition of galactose to the protein solution. (B) The plot of (Fo-F) vs. [C] shows saturation of binding site with increasing conc. Of galactose (C) Double logarithmic plot for binding of galactose with PotHg

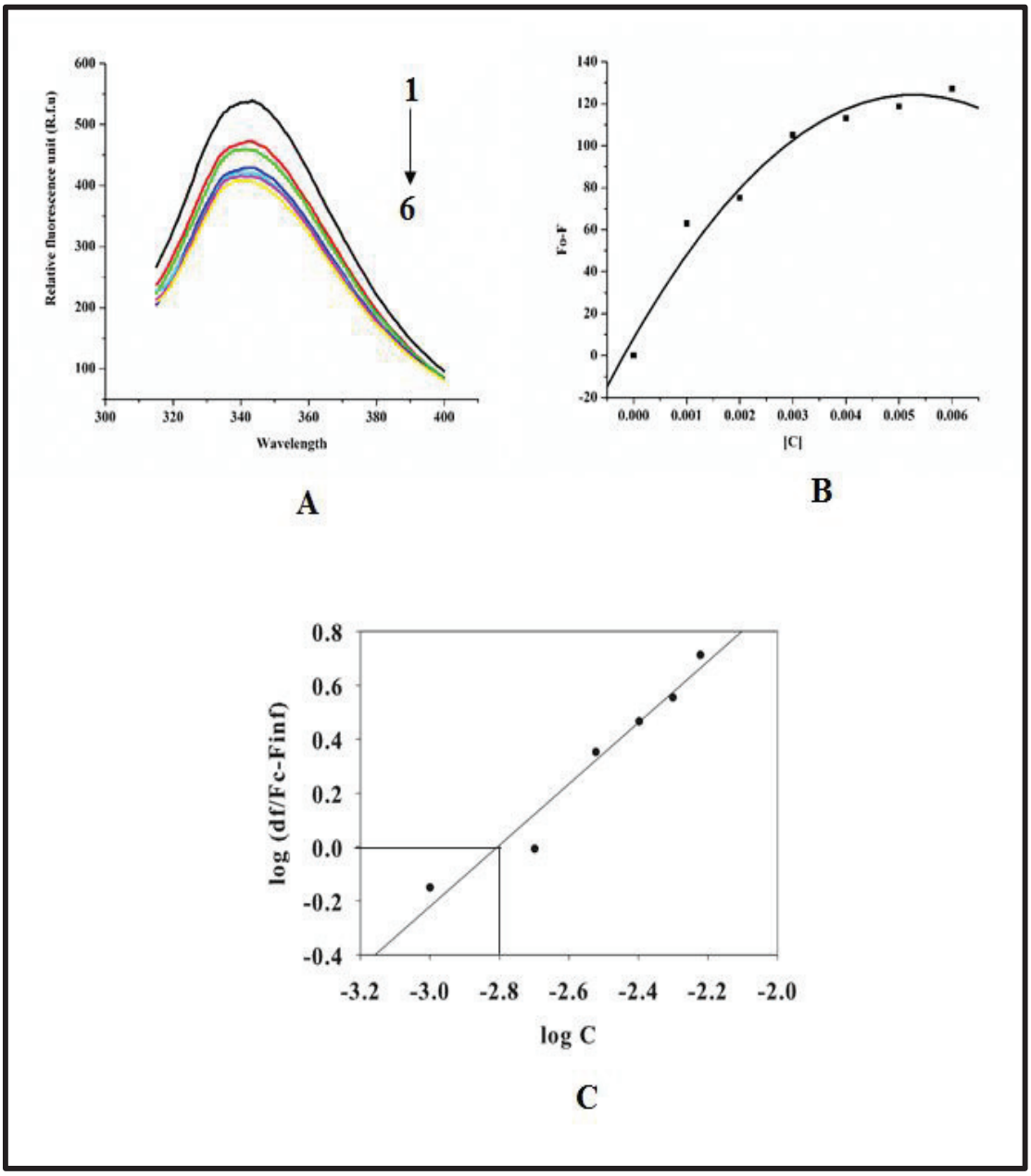


Figure 3.2 Fluorescence spectra of PotHg in the absence and presence of mannose (A) Represents fluorescence quenching of PotHg on the addition of mannose to the protein solution. (B) Plot of (Fo-F) vs. [C] shows saturation of binding site with increasing concentration of mannose (C) Double logarithmic plot for calculating binding affinity of mannose with PotHg

3.6.2 Solute quenching studies

Carbohydrate-binding studies using fluorescence spectroscopy and tryptophan modification by NBS confirmed the role of tryptophan in sugar binding. Sequence analysis of PotHg showed the presence of three tryptophan molecules in the protein. The extent of quenching with all three quenchers at a final concentration of 0.5M is shown (Table 3.1). Out of the three quenchers used, acrylamide was the best quencher and quenched 70% of the total fluorescence of PotHg solution. The charged quenchers I^- and Cs^+, which are unable to penetrate the protein matrix and access the buried tryptophan, quenched 35% and 22%, respectively, of the total fluorescence shown by PotHg. From the slope, Stern-Volmer plot of all the three quenchers at three different pH (Figure 3.3 A, B, C) Stern-Volmer constant Ksv, for each quencher, at each pH, was calculated (Table 3.1). From the modified Stern-Volmer plots obtained with all three quenchers at three different pH (Figure 3.4 D, E, F), the Y intercept was used to obtain values of fa, which is a fraction of fluorophore accessible to the quencher. From the collected data, it was observed that acrylamide had complete access to tryptophan at pH 5.0 and 10.0, whereas at pH 7.0, the accessibility was decreased to 75%. Negatively charged iodide ion had 80% access to fluorophore at pH 7.0 and 10.0 but at pH 5.0 the accessibility decreased to 29%. On the other hand, cesium ion had 28% access to fluorophore at pH 5.0, 7.0 and 10.0 (Table 3.1).

Table 3.1 Summary of parameters obtained from intrinsic fluorescence quenching of PotHg

Quencher	**Quenching (%)**			**Ksv**			**Fa (%)**		
	pH 5	pH 7	pH 10	pH 5	pH 7	pH 10	pH 5	pH 7	pH 10
Acrylamide	62	76	65	3.315	4.79	3.476	100	74.6	100
CsCl	20	25	22	0.49	0.573	0.539	27.7	27.7	28.0
KI	35	34	34	1.09	1.03	1.02	29.0	81.1	76.8

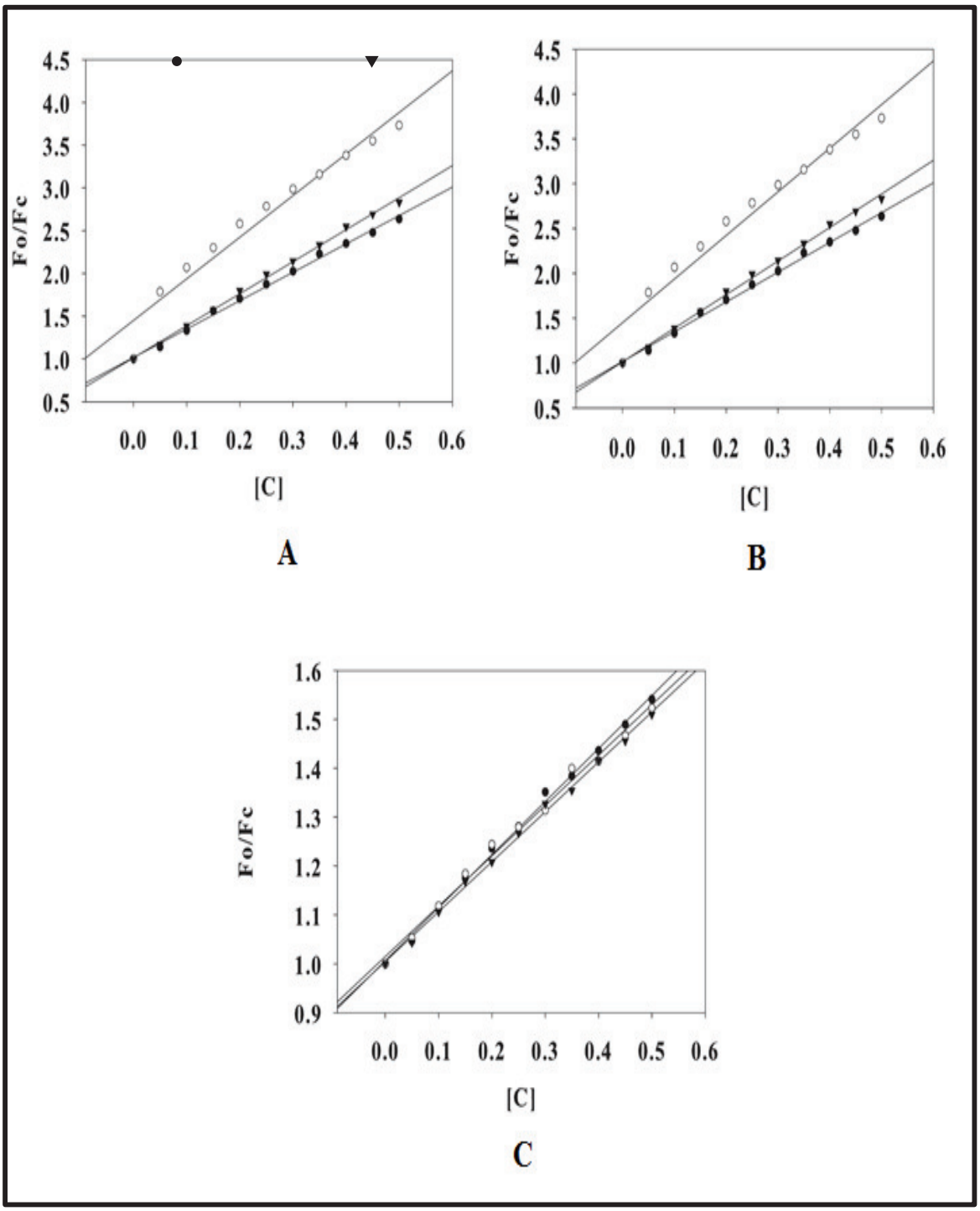


Figure 3.3 Stern-Volmer plots of fluorescence quenching of PotHg. The different quenchers used are (A) Acrylamide (B) chloride ion (CsCl) and (C) Iodide ion (KI). (○) Protein at pH 7, ◒ Protein at pH 5 and◓ Protein at pH 9.

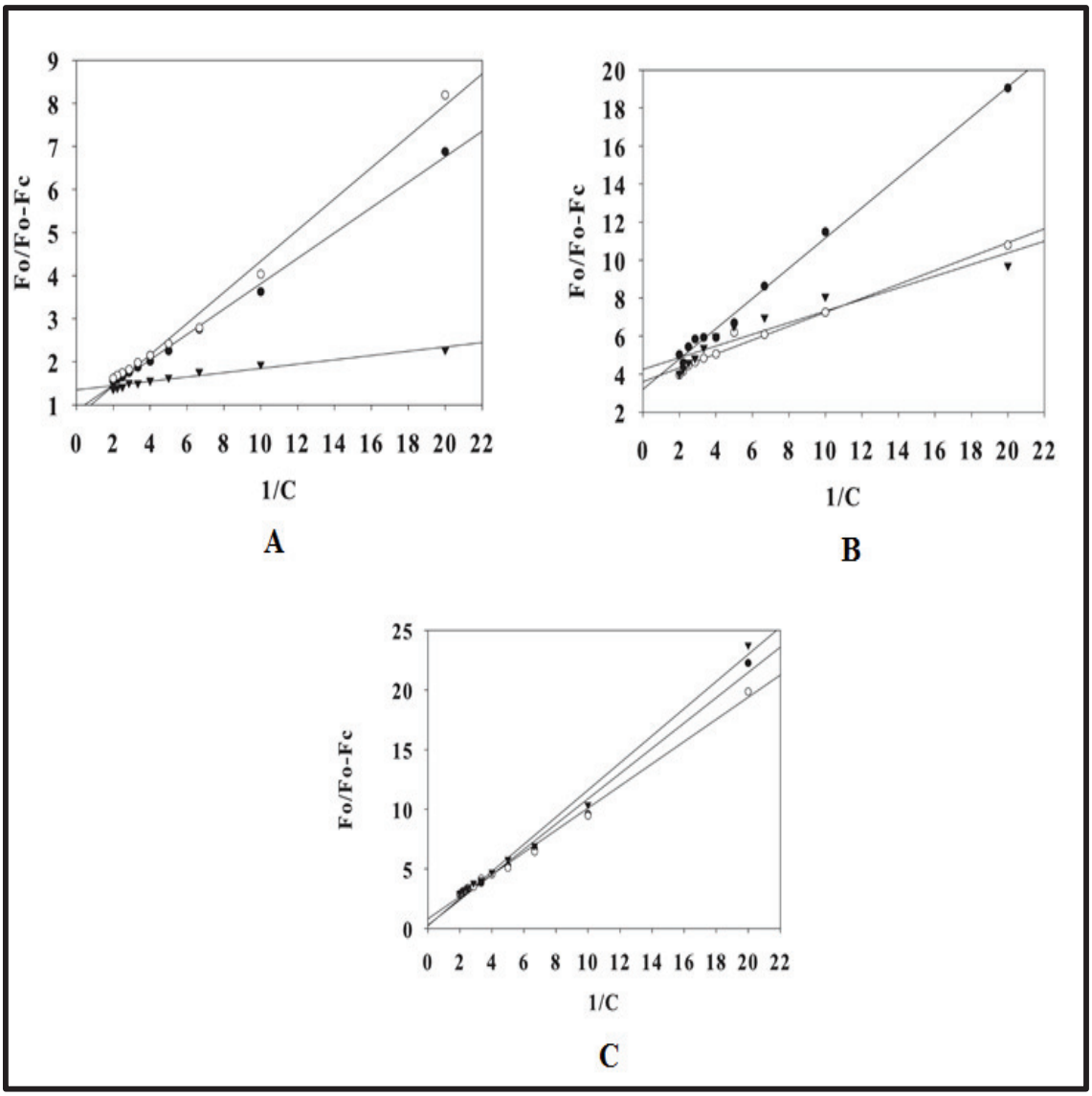


Figure 3.4 Modified Stern-Volmer plots of fluorescence quenching of PotHg. The different quenchers used are (A) Acrylamide (B) chloride ion (CsCl) and (C) Iodide ion (KI). (○) Protein at pH 7, (▼) Protein at pH 5 and (▼) Protein at pH 9.

3.6.3 Circular Dichroism spectroscopy

The far-UV Circular dichroism spectrum of native PotHg is characterized by a maximum near 225nm and a small minimum near 197nm with zero crossing near 215 nm (Figure 3.5 A). In the presence of galactose, the intensity near 225nm was increased, but there was no shift in Λ_{max}, indicating the absence of any change in secondary structure of the protein, after addition of galactose.

The analysis of near-UV CD spectra provided an understanding of the tertiary structure of a protein by giving insight into the interaction of aromatic amino acid side chains with peptide bonds. Analysis of near-UV CD spectra (Figure 3.5 B) of PotHg in the absence and presence of galactose showed Λ_{max} at 285nm and 292nm, respectively. The peak at 285nm signified the presence of tyrosine and the peak at 292nm represented the presence of tryptophan, thereby implying that the structure of PotHg is compact and rigid.

Data from CD melt studies indicated that the protein was stable until 75°C but beyond 75°C, it unfolded, as observed by a dip in the molar ellipticity values (Figure 3.5 C). These findings are by the observation of temperature stability assay; wherein PotHg showed hemagglutination activity till 80°C.

To derive information on the content of secondary structure elements of PotHg, the far-UV CD spectrum was analyzed by three different methods, namely CDSSTR, CONTINLL and SELCON3 (Table 3.2). The content of various secondary structures obtained from the CDSSTR analysis are: 0% regular α-helix, 3.6% distorted α-helix (given a total α-helical content of 3%), 26.9% regular β-sheet content and 13.6% distorted β-sheet (total of 41.5% of pleated sheet structure), 21.6% β-turns and 34.4% unordered structure. Along with these values, values obtained from CONTINLL and SELCON3 are listed in the table. Overall results obtained from analyzing far-UV CD spectrum data suggested PotHg is a predominately β-sheet protein with almost very little α-helical content.

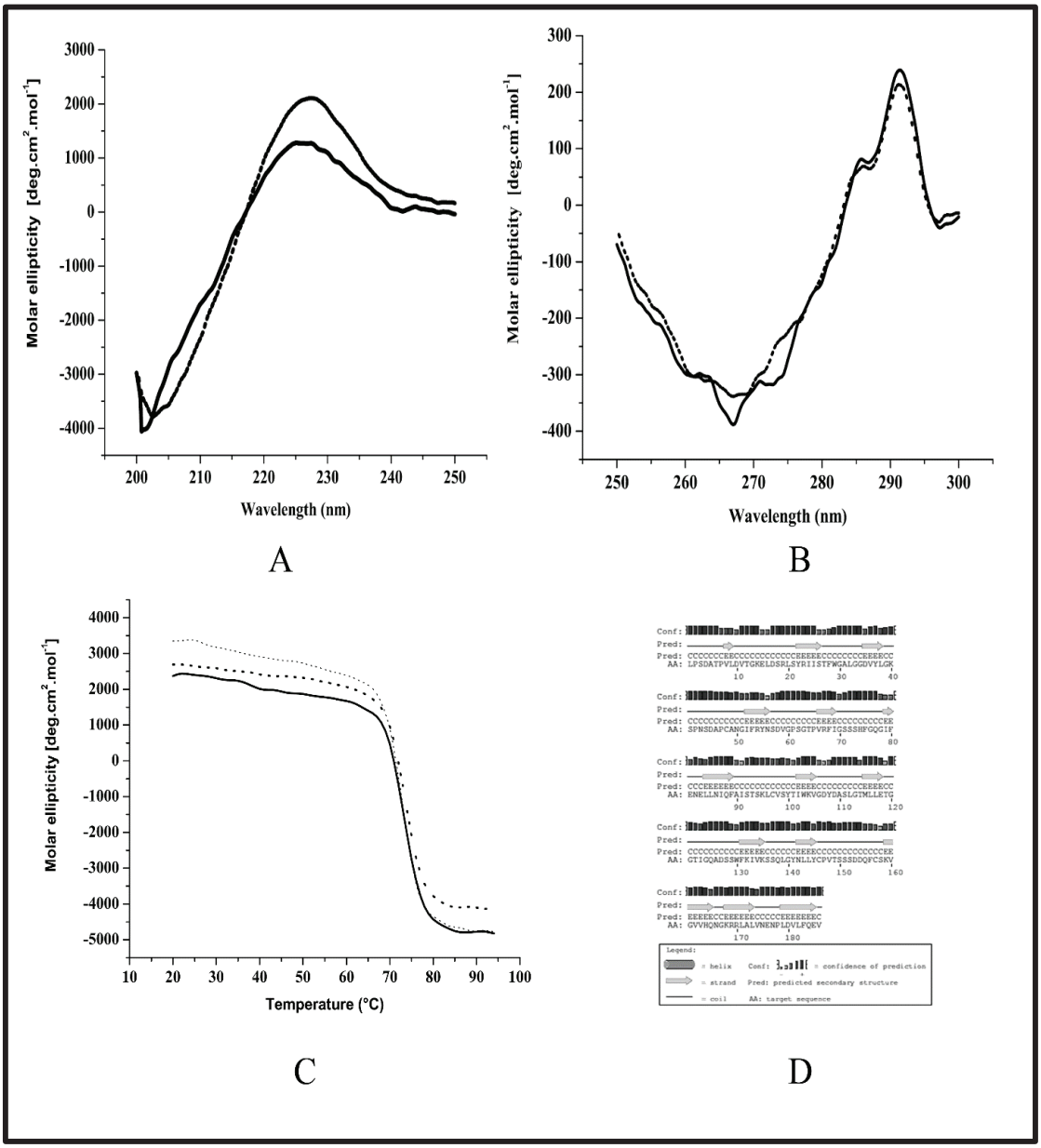


Figure 3.5 Circular dichroism (CD) spectra of PotHg in absence and presence of galactose (A) Far-UV spectra of PotHg (20 µM) (B) Near UV spectra of PotHg (80 µM). The solid line represents PotHg in the buffer, and the dotted line represents a protein in complex with galactose(C) CD melt analysis of PotHg and PotHg in complex with galactose and mannose. The ellipticity (Θ) in the far UV region of the native protein (—)(…..) lectin–mannose and (…..) lectin-galactose complex was monitored at 228 nm as a function of temperature. The sharp decrease in the ellipticity around 75 °C indicates the transition from the folded structure to the unfolded state. (D) Predicted secondary structure for PotHg from psipred protein analysis workbench

The results of secondary structure prediction based on the sequence of PotHg was carried out using PSIPRED (Figure 3.5 D). The results obtained from PSIPRED analysis are listed in the Table 3.2. The prediction results from PSIPRED indicates that protein is mainly consists of β-sheet. Since PSIPRED did not distinguish between turns and unordered structure rest of the structure was shown as turns. Secondary structure prediction by PSIPRED is comparable to experimental results obtained from circular dichroism spectroscopy.

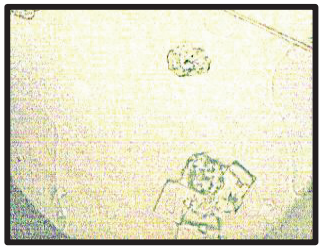
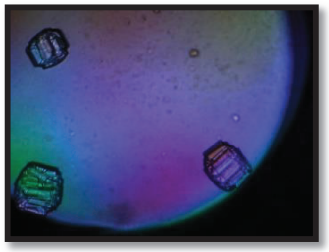
Table 3.2 Secondary structure elements of PotHg obtained from CD analysis

Method	α%	β%	Turns	Unordered
From CD Spectral Analysis				
CDSSTR	3.6	41.5	21.6	34.4
CONTINLL	1.7	45.4	18.2	34.3
SELCON3	0.0	44.9	19.4	32
Average	**1.76**	**43.93**	**19.73**	**33.56**
From Theoretical prediction				
PSIPRED	0.0	37.10	62.90	-

3.6.4 Crystallization & data collection

The purified protein PotHg (15 mg/mL in 20 mM Tris-HCl (pH 7.8) was first screened with different standard kits from Hampton Inc. Out of many conditions, three conditions gave some well shaped crystals after 21 days and promising for X-ray data collections (Table 3.3)

Table 3.3 Crystallization trials of PotHg

Name	Crystallization Condition	Image of Crystal
PotHg_C1	0.2 M Sodium acetate trihydrate, 0.1 M Sodium cacodylate trihydrate pH 6.5 & 30% v/v PEG 8,000	
Potg_C2	0.02 M Calcium chloride dihydrate, 0.1 M Sodium acetate trihydrate pH 4.6 & 30% v/v (+/-)-2-Methyl-2,4-pentanediol	
PotHg_C3	22% PEG8000 & 0.25M ammonium sulfate, pH 8.5	

The data collection of the micro batch crystal, PotHg_C1, (Table 3.3) was mounted for data collection under cryo conditions with oil as a cryoprotectant but did not show any diffraction. Data collection of crystals, PotHg_C2, also gave the similar results with no diffraction (Figure 3.6 A). Crystals, PotHg_C3, gave a diffraction pattern (Figure 3.6 B), but the diffraction spots were

overlapped due to poor crystal quality and could not be processed further. To improve the crystal quality, different parameters of crystallization condition were changed and the standardization such as a change in concentration of ammonium sulfate, change in PEG 8000 concentration and change in pH and temperature combined with a change in protein concentration. These efforts did not improve the crystal quality, and further trials are underway to obtain the good quality crystal for diffraction.

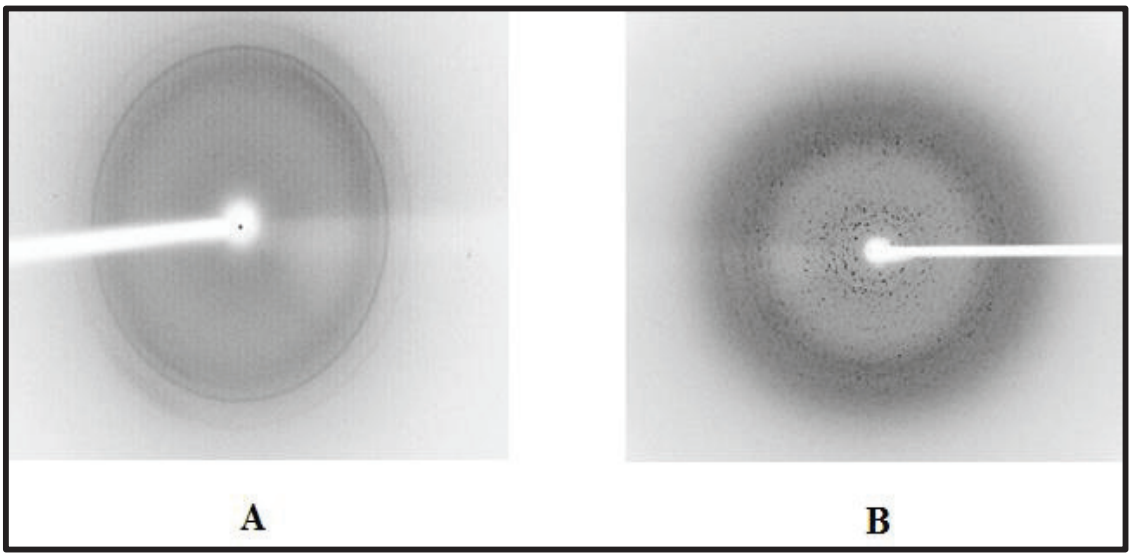


Figure 3.6 Diffraction images of PotHg. Reference image file from diffraction of PotHg in different conditions. (A) Diffraction pattern obtained for PotHg_C2. (B) Diffraction pattern obtained for PotHg_C3.

3.6.5 Molecular docking studies of PotHg

To find out which amino acids are involved in carbohydrate binding, docking studies of modeled PotHg with mannose and galactose was carried out, using ADT. Since NBS modification and fluorescence spectroscopy showed the role of tryptophan in sugar binding, we explored, which one out of three tryptophan present in the sequence is involved in sugar binding. To explore this first, we built the homology model of PotHg from I-TASSER server (Section 2.6.8.3). Once the homology model was built, CASTp server was used to find out the cavity present on the surface of PotHg, where sugar can go and bind with the protein. Results from CASTp analysis showed, out of three, two tryptophan molecules were present on the surface of the PotHg, and one tryptophan was buried inside (Figure 3.7 A). Also, surface charged analysis of these two exposed tryptophan

showed, one tryptophan (Trp 29) molecule has net positively charged environment (Figure 3.7 B) and another tryptophan (Trp 130) is having a net negatively charged environment (Figure 3.7 C).

Cavities surrounding two tryptophan which are surface exposed were used to define the binding site. Results from these studies showed that out of the two surface exposed tryptophan residues, only one (tryptophan at position 29) showed hydrogen bond formation with mannose. The binding energy for this interaction with mannose was -4.4Kcal/J (Figure 3.8 A). The other amino acids which were involved in this interaction were Asp 58, Val 59, Gly 60 and Pro 61. The second surface exposed tryptophan (position 130) revealed no hydrogen bond formation with mannose. Only one amino acid Gly 124 formed hydrogen bonds with mannose. The binding energy of this binding was -3.0 Kcal/J (Figure 3.8 B). Similar results were obtained in docking studies with galactose. Galactose showed hydrogen bond formation with one of the two surface exposed tryptophan (Tryptophan at position 29). The binding energy of this interaction was -4.8 Kcal/J (Figure 3.8 C). The other amino acids which were involved in this interaction were Val 58, Gly 60 and Pro 61. The second surface exposed tryptophan (position 130) followed the similar out come as in the case with mannose. There was no hydrogen bond formation between tryptophan and galactose (Figure 3.8 D).

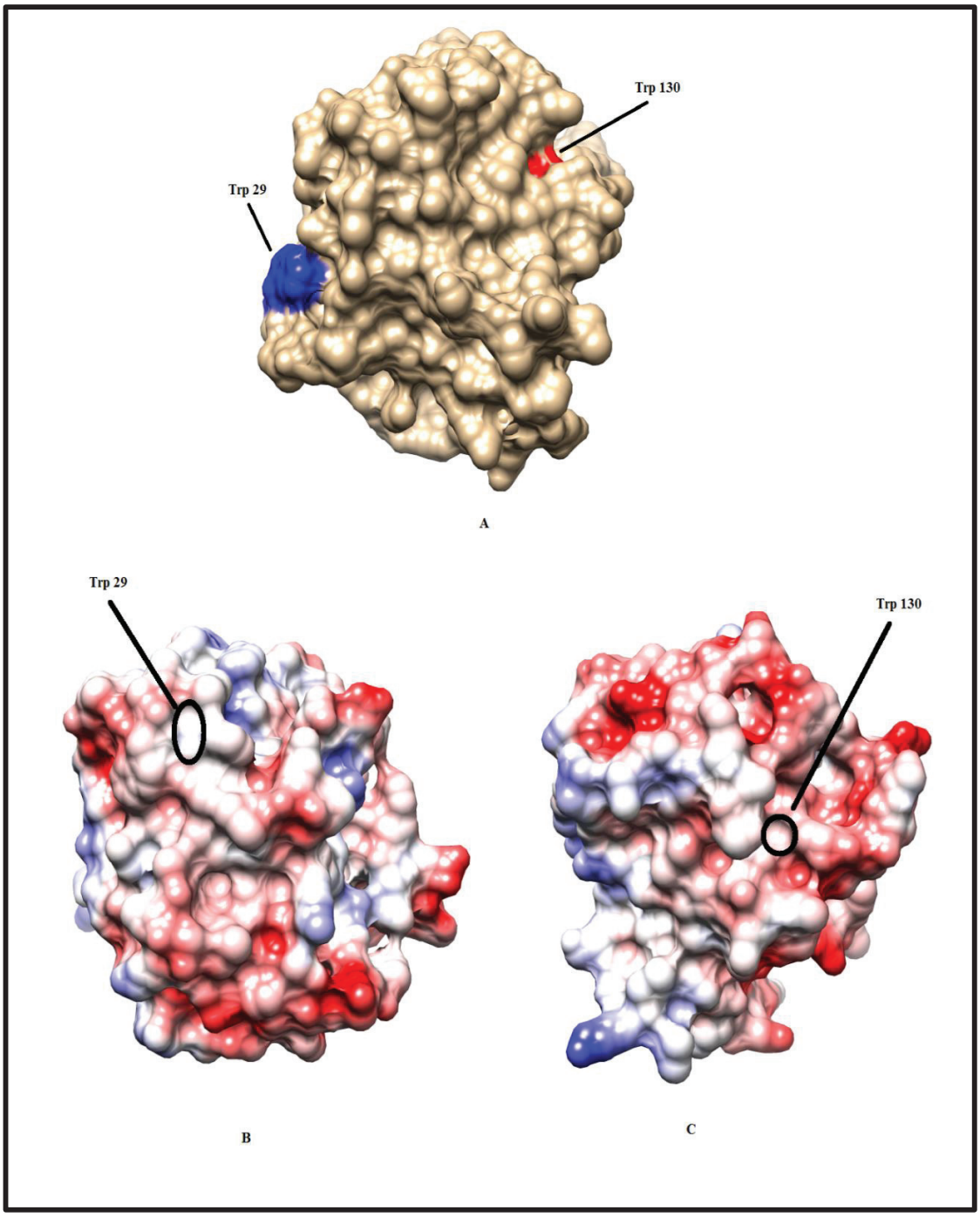


Figure 3.7 Analysis of Surface exposed Tryptophan in PotHg (A) Demonstration of surface exposed Tryptophan in PotHg by CASTp server. Results showed, Tryptophan 29 (Blue) and Tryptophan 130 (Red) are exposed on the surface. (B) Distribution of columbic charges on the surface of PotHg. Trp 29 is surrounded by the positively charged environment. (C) Distribution of columbic charges on the surface of PotHg. Trp 130 is surrounded by negatively charged environment.

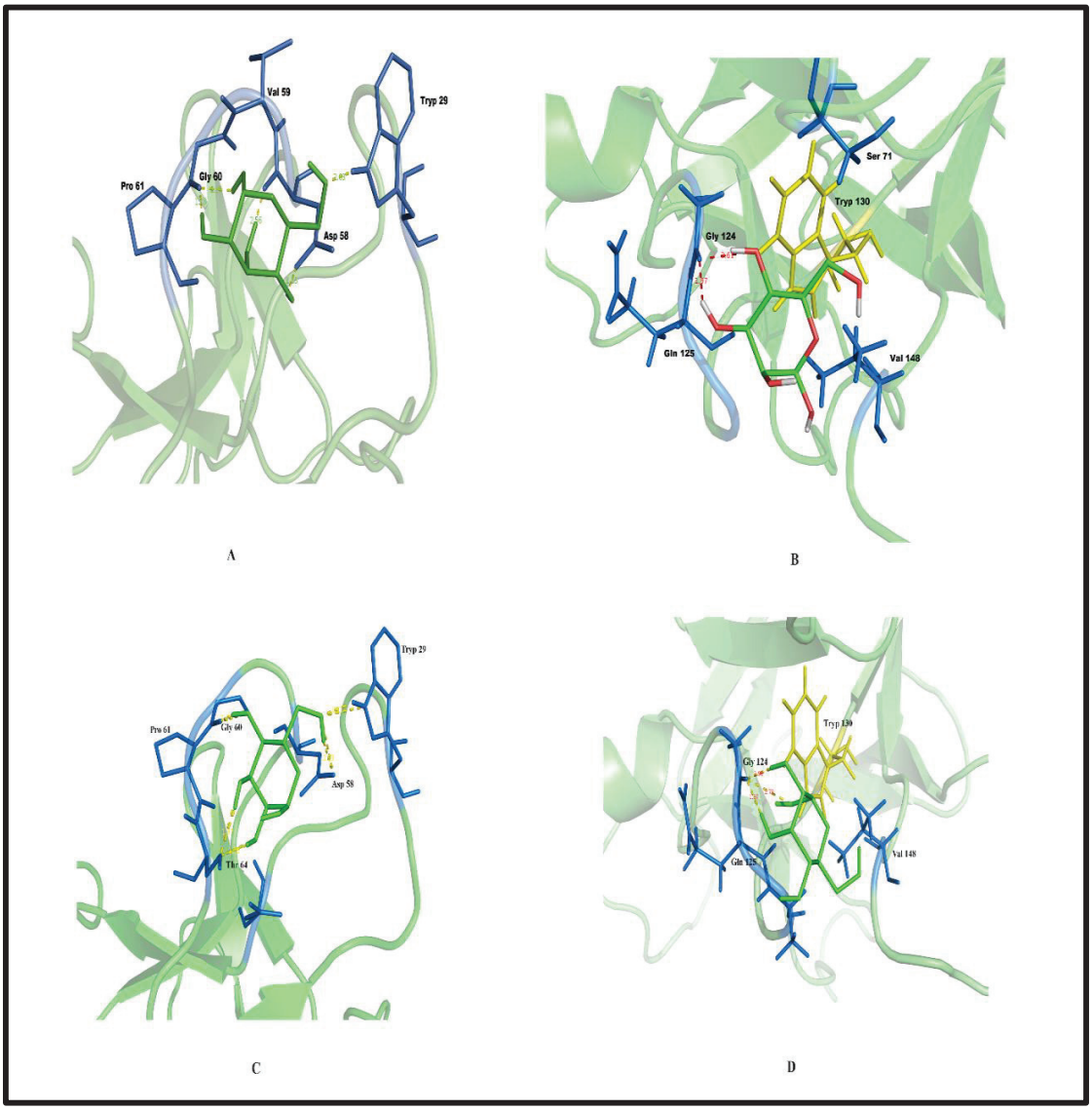


Figure 3.8 Molecular docking studies of PotHg with Mannose & Galactose (A) Docking of mannose with cavity surrounding Trp 29. The amino acids involved in binding are Trp 29, Asp 58, Val 59, Gly 60 and Pro 61. The binding energy was -4.4 Kcal/J. **(B)** Docking of mannose with cavity surrounding Trp 130. Tryptophan did not make any hydrogen bond with mannose. Only Gly 124 was involved in hydrogen formation with mannose. The binding energy was -3.0 Kcal/J. **(C)** Docking of galactose with cavity surrounding Trp 29. The amino acids involved in binding are Trp 29, Val 59, Gly 60 and Pro 61. The binding energy was -4.8 Kcal/J. **(D)** Docking of galactose with cavity surrounding Trp 130. Tryptophan did not make any hydrogen bond with galactose. -3.2 Kcal/J. The distances are shown in Å.

3.7 Discussion

The lectin-like properties of PotHg were evidenced by its hemagglutination activity on rabbit RBC and hemagglutination inhibition assay, which showed that this protein had an affinity for fetuin but did not show any inhibition by simple carbohydrates like mannose and galactose. These observations were further strengthened by thermal shift assay data exhibiting a change in the Tm of protein in the presence of mannose and galactose, though at a very high sugar concentration (>100 mM; data not shown). Fluorescence spectroscopy data further provided the proof of PotHg interaction with galactose and mannose with Kd values of 2.8 x 10^{-3} M and 1.5 x 10^{-3} M, respectively. Thus, the uniqueness of this protein which is a serine protease inhibitor is the presence of lectin-like activity, and this property was further investigated in detail by fluorescence spectroscopy in the present study. This would give insight into the structural basis of sugar interactions of these proteins as in literature there are very few reports of protease inhibitors exhibiting lectin activity. Protease inhibitors with lectin-like activity and specificity towards complex carbohydrates such as asialofetuin, thyroglobulin, ovalbumin, and orosomucoid have been purified from seeds of *Peltophorum dubium and Labramia bojeri* in earlier reports (Macedo et al., 2004; Troncoso et al., 2003; Saitoh et al., 2005).

To decipher the probable sugar binding sites, fluorescence spectroscopy was performed. The tryptophan emission spectra record indicated changes in the tryptophan environment upon addition of carbohydrates. Additional experiments using NBS modified tryptophan and molecular docking confirmed the involvement of tryptophan in carbohydrate interaction. Data obtained from surface analysis of PotHg showed two out of three tryptophan were surface exposed. Charge analysis on the surface of PotHg showed, Try 29 is having a positively charged environment and Trp 130 had negatively charged environment. Molecular docking data also showed surface exposed tryptophan at position 29 to be involved in sugar binding. The data obtained from these two studies support the data of fluorescence quenching studies which suggested that tryptophan has a positively charged environments. These results although preliminary definitely provided information about Trp 29 and its role in carbohydrate recognition and binding.

The structural properties of PotHg were recorded using spectroscopic measurements. Far-UV CD spectra indicated no changes in the secondary structure of protein, after galactose addition. Hence, far-UV CD spectra exhibited high similarities to those of other potato serine protease inhibitors isoforms, other potato cysteine protease inhibitors, and wheat germ agglutinin

(Pouvreau et al., 2005; Rodríguez-Romero et al., 1989). Evidence to support the above observations were provided by bioinformatics data, where PotHg was shown to possess a β-trefoil fold also shared by both, protease inhibitors and lectin family. Thus, it may be concluded that structural similarities exist between PotHg, and protease inhibitors and lectins. Moreover, analysis of near-UV CD spectra of PotHg in absence and presence of galactose showed Λ_{max} at 285nm and 292nm, respectively indicating the presence of tyrosine and tryptophan, in the sequence and also confirmed PotHg structure to be rigid and compact.

Thus, it may be concluded that PotHg is a unique Kunitz-type serine protease inhibitor having lectin-like activity. Further studies are necessary to elucidate the details of serine protease inhibition and regulation by this novel protein, to provide a scientific basis for better pest management strategies.

3.8 Conclusion

In this chapter, we have established carbohydrate binding property of PotHg with fluorescence spectroscopy with simple carbohydrate mannose and galactose. The data obtained from fluorescence spectroscopy combined with data on the effect of NBS on carbohydrate binding (Chapter 2) proved the potential role of tryptophan in carbohydrate binding. To explore in detail and pinpoint the tryptophan, surface charge analysis and molecular docking were used, and results from these studies combined with data obtained from fluorescence quenching with different chemical quencher proved Tryptophan present at position 29 is involved in carbohydrate binding. The circular dichroism data established that PotHg mainly consisting of β-strands and overall structure of PotHg is compact and rigid in nature, which is not affected upon carbohydrate binding. Thermal melt study using CD spectroscopy showed PotHg is stable up to 80°C, which is consistent with the results obtained from temperature stability studies (Chapter 2)