Research Article
Volume 13:04, 2023

Journal of Bioprocessing & Biotechniques

ISSN:2155-9821 Open Access

Mass Spectrometric Characterization of a Novel Galactose-Binding Lectin from *Chlorella sorokiniana* (MW769776)

Savitha J[1]*, Roopashri N Arekal[1], Ashwini Kumar[2], Gunashree BS[3] and Divyashree MS[4]

[1]*Department of Microbiology, Bangalore University, Jnanabharathi Campus, Bengaluru, India*

[2]*Department of Microbiology, Genei Lab Pvt Ltd, Bengaluru, India*

[3]*Department of Microbiology, Mangalore University, Kodagu, India*

[4]*Department of Biotechnology, Manipal Institute of Technology (MIT), Manipal, India*

Abstract

A freshwater green microalgal strain was isolated, and presence of lectin was identified by its strong Hemagglutination Activity (HA). The molecular characterization of algal strain was found to be the *Chlorella sorokiniana* (MW769776). The *Chlorella Sorokiniana* Lectin (CSL) was purified by single-step affinity chromatography method using guar-gum as a resin. The precipitate showed single active peak with titer value of 1024 HU, specific activity of 539 mg/mL, and with purification factor. The purified lectin revealed single band on denaturing electrophoresis with. Liquid Chromatography-Electrospray Ionization-Quadrupole-Time of Flight Mass Spectrometry (LC-ESI-Q-TOF-MS) analysis of tryptic digested purified lectin suggested that it was a monomeric protein. A multiple sequence alignment study revealed that the peptide sequences of CSL exhibited similarity with the H-type lectin domain from Micractinium conductrix. The structure of CSL was studied by FTIR and homology modeling methods, indicating that its secondary structure contained α-helix, β-sheet, and unordered structure, whereas the 3D structure exhibited the similarity with proteins from light-harvesting reaction center complex of photosystem I. The main significance of this study includes the characteristics of CSL are consistent with its identification as a hemagglutinin, a kind of novel lectin, which suggests its candidature for various biological purposes.

Keywords: *Chlorella sorokiniana* • Lectin • Purification • Characterization • Mass spectrometry

Introduction

Lectins are carbohydrate binding proteins of non-immune origin that are capable of agglutinating cells or precipitating carbohydrates without the involvement of enzymatic activity. Lectins are ubiquitous in nature and are found in either soluble or membrane bound forms, with different molecular weights, functions, and structures. They are structurally heterogeneous type of proteins with one or more carbohydrate recognition domains. They bind specifically and reversibly to different types of glycol-conjugates derived oligosaccharides. Because of these benefits, lectins from different sources have been used for the glycan profiling, diagnosis of diseases and cancer biomarkers. Many investigative studies have been carried out to examine their health benefits, and reported that they have important physiological roles, including antibacterial, antifungal, anti-human immunodeficiency virus, antitumor, insecticidal, and analgesic activities. However, it has been reported that lectins occur in algae are a major source to utilize for many biological purposes. Several types of lectins from algae have been detected, studied for their advantages. The lectins existing in the freshwater microalgae have gained great importance and have demonstrated their critical role in suppressing the activity of HIV, insecticidal action, and antimicrobial activity [1].

Green algae, typically included in the genus *Chlorella* are among the most widely distributed and frequently encountered algae in freshwater habitats. They are naturally small (~2 to 10 µm in diameter), unicellular, round in shape, non-motile, and contain a single-cup shaped chloroplast. Some cells have a rigid cell wall, and they are reported to lack a sexual cycle. Like other microalgae, there is an increasing interest in using *Chlorella* in a variety of biotechnological applications, mainly the production of molecules with high economic value. Among these, lectin molecules have significance in many types of biotechnological uses, such as for detection of cancer biomarkers from biological fluids of cancer patients, single molecule interaction study, and other types of pharmacological targets. The presence of lectins in algae was first identified by since then, very few reports have been published on the characterization of lectins from freshwater microalgae. Our research aimed to characterize the *Chlorella* lectin, which would provide a theoretical foundation for the complete utilization of this protein for biotechnological purposes.

In the present study, *Chlorella sorokiniana* MW769776 strain was chosen to isolate and purify lectin by a single-step method using affinity chromatography technique. The properties of the algal lectin were analyzed using different investigational methods. Herein, we report the purification and characterization of a novel galactose-binding lectin from *Chlorella sorokiniana* MW769776 (CSL) using mass spectrometry [2].

***Address for Correspondence:** Savitha J, Department of Microbiology, Bangalore University, Jnanabharathi Campus, Bengaluru, India, Tel: 9886185170; E-mail: drsvtj@yahoo.co.in*

Received: 22 March, 2023, Manuscript No. JBPBT-23-92481; **Editor assigned**: 27 March, 2023, PreQC No. JBPBT-23-92481 (PQ); **Reviewed:** 10 April, 2023, QC No. JBPBT-23-92481; **Revised:** 29 May, 2023, Manuscript No. JBPBT-23-92481 (R); **Published:** 06 June, 2023, DOI: 10.37421/2155-9821.2023.13.573

Materials and Methods

Chemicals and reagents

All chemicals and reagents used in this study were of analytical grade and procured from Himedia (India), Sigma Aldrich (US), and SD Fine (India).

Culture conditions

The green microalgal culture *Chlorella sorokiniana* was isolated from freshwater and maintained on Sueoka's salt medium composed of 50 mL/L of Beijerinck's solution (10 g NH_4Cl, 0.4 g $MgSO_4.7H_2O$, 0.2 g $CaCl_2.2H_2O$, 1 L distilled water), 50 mL/L of phosphate buffer (28.8 g K_2HPO_4, 14.4 g KH_2PO_4, 1 L distilled water), 1 mL/L of Hutner's trace elements, 2 g/L of sodium acetate (pH 7.2) as mentioned in our previous paper. The algal culture was grown at 25°C in 2.0 L Erlenmeyer conical flask containing 1.0 L culture broth under 50-100 µEinsteins/m_2/s irradiance (cool fluorescent white light), which was maintained under a 10 h light-14 h dark photoperiod. Cells were cultured for 10 days, and then harvested at the stationary phase by centrifugation at 8,000 rpm for 15 min. Harvested cells were lyophilized and stored at -20°C until use [3].

Molecular characterization of *Chlorella* species

The molecular characterization of the screened *Chlorella* culture was carried out by target gene sequence homology study. For this, algal culture was subjected to PCR analysis to amplify the 18S rRNA gene fragment, for which universal eukaryotic primers were used to estimate and analyse the chlorophyta variety. In the first step, algal culture was cultivated in Sueoka's salt medium under standard condition of 25°C, 50-100 µEinsteins/m²/s irradiance, and maintained under a 14 h light-10 h dark photoperiod for seven days. Aliquots of culture broth were plated on Sueoka's salt medium to obtain discrete colonies. A single colony was regrown in sueoka's salt medium under the same standard conditions. This was followed by harvesting of algal cells in the course of the exponential growth period, and the extraction of genomic DNA using the conventional method. The targeted region of 18S rRNA gene fragment was amplified using published primers *i.e.,* NS1 forward (GTA GTC ATA TGC TTG TCT C) and NS4 reverse (CTT CCG TCA ATT CCT TTA AG) for algal culture. PCR analysis was performed by preparing 50 µL of reaction mixture containing $MgCl_2$ (2 mM), DNA (50 ng), dNTPs (200 µM), primers (0.6 µM each), and Taq DNA polymerase (1.25 U). PCR analysis was carried out under the following conditions, *i.e.,* 1st cycle at 94°C for 3 min followed by 35 cycles at 94°C for 1 min, 58°C for 1 min, 72°C for 1 min, and a final cycle at 72°C for 7 min. Amplified PCR products were purified using the purification kit (Sigma Aldrich, India), later submitted to the Eurofins Genomics India Pvt. Ltd for gene sequencing. It was carried out using the BDT v3.1 cycle sequencing kit on an ABI 3730xl genetic analyzer. Subsequently, acquired nucleotide sequences were examined using the BLAST (Basic Local Alignment Search Tool) with the NCBI database (National Centre for Biotechnology Information) to evaluate the percentage of homology with closely related strains, which were documented in GenBank. A phylogenetic tree was constructed using Phylogeny.fr [4].

Preparation of crude protein extracts from harvested cells

The *Chlorella* sp. was grown in 250 mL of Sueoka's high salt medium for 10 days. The algal biomass was collected by centrifugation at 7000 rpm, 4°C for 10 min. The resultant pellet (1 gm) was washed thrice with distilled water, and the soluble protein was suspended at a 1:10 (w/v) ratio in 0.025 M Phosphate buffer (pH 7.2), containing 0.15 M NaCl. Biomass was homogenized by sonication at >20 kHz condition for 15 min cycles in an ice bath, followed by storage of the sample mixture at 4°C for overnight. Subsequently, the sample mixture was centrifuged at 7000 rpm for 15 min at 4°C, and the resultant supernatant was collected and designated as the Aqueous Extract (AE) for further investigations.

Purification of lectin from *Chlorella sorokiniana*

Protein precipitation: Aqueous crude protein extract (15 mL) was precipitated by fractionation in five stages, namely, 0-20, 20-40, 40-60, 60-70, and 70-80 (w/v) concentrations of ammonium sulphate. Solid ammonium sulphate was added slowly while the solution was kept stirring to allow for a uniform increase in the concentration and ensure rapid equilibration, and the sample was stored at 4°C for 12 h. Precipitates showing positive hemagglutination reactions (primarily from the 0-80% ammonium sulfate fractions) were pooled, dissolved in 0.025 M PBS (pH 7.2), and dialyzed against the PBS at 4°C for 12 h by refreshing the buffer at every 4 h once. The dialyzed sample was centrifuged to collect the resultant supernatant for further study [5].

Preparation of affinity chromatography column using guar-gum matrix: Affinity chromatography is a technique used to purify proteins that bind non-covalently and reversibly to specific molecules known as ligands. In the present study, guar-gum beads (SRL, India) a galactose polymer was used. As the protein was galactose-specific, the retention of CSL was expected to occur in a single step. Insoluble guar-gum was cross-linked with an epichlorohydrin emulsion (Himedia, India) in 25 mL of 3 N sodium hydroxide. The mixture was stirred vigorously using a spatula until it solidified and was kept at 40°C in a water bath for 24 h with occasional stirring, followed by heating in an oven at 70°C for 4 h. Later, it was homogenized in a mortar pestle to a particle size of approximately 30 mm size in the presence of 0.025 M PBS (pH 7.2, containing 0.15 M NaCl). The homogenized guar-gum beads was then packed into a Borosil column (1.5 cm × 10 cm) and equilibrated with 0.025 MPBS (pH 7.2, containing 0.15 M NaCl) until the absorbance of elutes showed <0.05 reading which was taken at 280 nm using a spectrophotometer (LABMAN UV-visible, India). This was performed to ensure the absence of any other contaminant protein present in the column.

Purification of *Chlorella sorokiniana* lectin: The soluble lectin protein in the precipitated extract was purified using single-step affinity chromatography method. Dialyzed protein precipitates in 0.025 M PBS (pH 7.2, containing 0.15 M NaCl) from ammonium sulphate precipitation (P 60-70) of the crude extract was loaded onto the cross-linked guar-gum matrix column (1.5 cm × 10 cm). The column was then washed with PBS at a flow rate of 0.5 mL min^{-1} until the column effluent showed an absorbance of less than 0.05 at 280 nm. The adsorbed proteins were eluted by the addition of 50 mL of 0.2 M D-galactose (Himedia, India) in 0.025 M PBS (pH 7.2). Fractions of 0.8 mL min^{-1} were collected and checked for the presence of proteins by measuring the absorbance at 280 nm as well as lectin content by HA test using 2% chicken erythrocytes. The active fractions were pooled and dialyzed against distilled water. Finally, the sample was freeze-dried and stored at -20°C until required period [6].

Hemagglutination activity and inhibition assays

Preparation of native erythrocytes sample: Blood samples of animals (chicken and sheep) were obtained by venous puncture of healthy animals from local butcher shops in Bengaluru. One milliliter of each blood sample was centrifuged at 3500 rpm for 15 min at room temperature to collect the erythrocytes, which were then washed thrice with 0.01 M phosphate buffer saline PBS (pH 7.2, containing 0.15 M NaCl). The resultant packed red blood cells were diluted with PBS to obtain a 2% (v/v) blood suspension and stored at 4°C for further use.

Hemagglutination activity assay: The hemagglutination activity assay was performed by a two-fold serial dilution method using a 96-well microtitre V-plate and 2% native erythrocyte suspension (v/v). In this step, 50 µL of purified lectin sample was serially diluted with 50 µL of 0.01 M PBS (pH 7.2), and 50 µL of 2% (v/v) erythrocyte suspension was added to the well. Similarly, a control sample was prepared by adding 100 µL of PBS without adding the protein sample. The plates were then incubated for 1 h at room temperature. Hemagglutination activity was observed visually when more than 50% of RBC in the well was agglutinated, which was depicted as mat formation, and tested as positive. The negative

hemagglutination activity showed button formation at the bottom of the well. The hemagglutination activity was expressed as the Hemagglutination titre (HU) unit. The hemagglutination titre value is defined as the reciprocal of the highest dilution showing positive hemagglutination activity.

Hemagglutination-inhibition assay: The hemagglutination inhibition test was carried out to investigate the carbohydrate binding specificity of the screened lectin using different carbohydrates, such as monosaccharides, oligosaccharides, and glycoproteins. In this study, the following sugars and glycoprotein used were D-glucose, D-galactose, D-maltose, D-mannose, D-fructose, sucrose, D-lactose, D-xylose, L-rhamnose, L-arabinose, D-ribose, D-cellobiose, L-stachyose, and D-raffinose as sugars, and yeast mannan as glycoprotein. The concentration of each sugar was 400 mM, whereas the glycoprotein concentration was maintained at 2 mg/mL. The sugar specificity of the lectin was analyzed in a manner analogous to that of the hemagglutination test. This assay involves a 2-fold serial dilution of sugar solution and then purified lectin solution was added. For this assay, a hemagglutinin solution with a hemagglutination titer unit 8 HU was used. PBS was used as a control to observe the inhibition of lectin hemagglutinating activity by the different carbohydrate groups.

Characterization of purified lectin (CSL)

Homogeneity and molecular mass determination: The degree of protein purity and Molecular mass (Mr) of the lectin was assessed using 16% Tricine SDS-PAGE (Tricine-Sodium Dodecyl Sulphate Polyacrylamide Gel Electrophoresis). Five micrograms of purified protein was subjected to Tricine SDS-PAGE. The protein sample was prepared using a sample buffer containing 12% (w/v) SDS and 6% (v/v) β-mercaptoethanol and heated at 90°C for 5 min. The electrophoretic procedure was performed using a mini electrophoretic system (Bio-Rad, India) under a constant voltage (80 V), and pre-stained markers (2.5 kDa-45 kDa) (SRL, India) were applied to evaluate the molecular weight of the sample. A standard coomassie-brilliant blue R-250 method was used for staining the gel following electrophoresis [7].

Enzymatic in-gel digestion and peptide extraction

The spot signifying the protein of interest was cut out from the Coomassie stained 1-D gel and subjected to in-gel tryptic digestion according to the standard protocol. Concisely, the excised gel piece was washed with 100 µL of destaining solution containing 1:1100 mM Ammonium Bicarbonate (NH_4HCO_3) and 100% Acetonitrile (ACN) until the gel piece appeared colorless and then dehydrated in ACN. Protein reduction was performed in 100 µL of 10 mM DTT in 50 mM NH_4HCO_3 at 56°C for 45 min. Alkylation of protein was prepared in 100 µL of 55 mM iodoacetamide (IAA in 50 mM NH_4HCO_3). Subsequently, the gel piece was destained in 100 µL of a solution containing a 1:1 ratio of 50 mM NH_4HCO_3 and 100% ACN. The final dehydration was performed in 100% ACN for 15 min. To remove the retained moisture in the gel piece, the sample tube was centrifuged, and the resultant supernatant was discarded and completely dried in a speed vac-vacuum concentrator (Thermofisher, India). Trypsin solution (20 ng/µL) (Sigma chemicals-proteomic grade, India) was prepared in 25 mM NH_4HCO_3 and 10 µL of this solution was added to each Eppendorf tube, and kept at 4°C for 30 min for absorption. This was followed by further addition of 10 µL of a 25 mM NH_4HCO_3 solution. The sample was digested at 37°C for 16 h. After digestion, sample was centrifuged, and supernatant containing peptides was collected and transferred to the fresh eppendorf tube rinsed with 100% ACN. Ten microliters of peptide extraction solution containing 1% Triflouroacetic Acid (TFA) and 60% ACN were added to the sample and sonicated for 15 min in an ultrasonic water bath at room temperature. After centrifugation, the supernatant containing peptides was collected, pooled with the respective sample, and dried completely in a speed vac. For MS analysis, the peptides were resuspended in 10 µL of 0.1% TFA containing 30% CAN [8].

LC-ESI-Q-TOF MS analysis: Liquid chromatography-electrospray ionization-quadrupole-time of flight MS analysis of the tryptic digested protein sample was carried out using a high-pressure liquid chromatography (Dionex Ultimate 3000 HPLC) system with an agilent C18 column (4.6 X 150 mm). The mobile phase consisted of water, acetonitrile, and 0.1% formic acid. The gradient was started with 2% of acetonitrile to 98% acetonitrile in 50 min. The digested sample was ran on ESI-Q-TOF impact HD (Bruker Daltonics, Germany) connected to HPLC system in a way that a sample eluted from an HPLC column was directly ionized by Electrospray Ionization (ESI) process and entered into the mass spectrometer. Software for analyzing mass spectra: Flex Analysis 3.1.

The obtained mass spectrum data was submitted to an online database (MASCOT) for protein identification. The database search parameters were set as follows enzyme-trypsin, taxonomy-all entries, fixed modifications-carboxymethyl (C), variable modifications-oxidation (M), mass values-monoisotopic, protein mass-unrestricted, missed cleavage-1. The Peptide Mass Fingerprinting (PMF) data was analyzed after removing the trypsin autolysis, keratin, and matrix peaks using the Flex analysis software version 3.1.

Fourier transform infrared spectroscopy: The infrared spectrum of the *Chlorella* lectin was obtained with a Fourier transform infrared spectrometer (Shimadzu, Kyoto, Japan) within the wave number range of 400 to 4000 cm^{-1}. First, 1 mg of the tested lectin lyophilized powder was ground together with 150 mg of spectroscopic grade KBr powder and pressed into a 1 mm pellet. The experiment was repeated three times. The secondary structure of lectin was analyzed using the FT-IR spectrometer, where the spectra were processed using Peakfit v4.12 software.

Effect of pH, temperature and metal ions on hemagglutinating activity: To determine the thermal stability, freeze-dried CSL (150 µg mL^{-1}) was suspended in PBS, and the sample aliquots were heated in a water bath maintained at different temperatures (30-100°C) with an increment of 10°C for 30 min. At each temperature, an aliquot was drawn and cooled to room temperature to estimate lectin activity by Hemagglutination Assay (HA).

The stability of the purified lectin at different pH values was determined by hemagglutination activity of the lectin samples suspended in different buffers of pH 4 and 5 (25 mM citrate-phosphate buffer), pH 6 and 7 (25 mM sodium phosphate buffer), pH 8 and 9 (25 mM Tris-HCl) and pH 10 and 11 (25 mM glycine-NaOH buffer), and incubated for 18 h at 25°C.

The same procedure was followed to evaluate the effects of Mg^{2+} and Ca^{2+}. In this case, incubation time was 15 min and metal solutions tested were 0.02 M $CaCl^2$ and $MgCl^2$ in 0.025 M PBS (pH 7.2). In all experiments, an aliquot (50 µL) of the mixture was distributed in a 96-well microtitre plate and the HA assay was performed.

Homology based molecular modeling: Swissmodel was used to generate a structural model of the CSL protein. For homology modeling, amino acid sequences were submitted to SWISSMODEL using the automated mode set at default parameters and models were generated using the crystal structure of *Chlorella* lectin. Manual adjustment was carried out on the models using deepview Swiss-Pdb viewer 4.0.1 and rasmol the relative similarity of fit was determined by pairwise measurements of the Root Mean Square Deviation (RNSD) for the Cα atoms of all residues [9].

Total protein estimation

Protein content was determined by the Bradford method and bovine serum albumin was used as standard.

Statistical analysis

The data were expressed as the mean or the percent mean ± Standard Deviation (SD). All the experiments were tested and analyzed in triplicates. Microsoft Excel 2010 (USA) was used for the regression analysis.

Results and Discussion

In this study, the lectin present in *Chlorella* sp. was identified, and its

amino acid sequences was determined by Peptide Mass Fingerprinting (PMF) and tandem Mass Spectrometric (MS/MS) analysis of tryptic peptides. Previous information on already submitted freshwater green microalgal lectins was very useful to investigate the presence of isoformic variants.

Molecular characterization-pcr detection for genus specificity

In the present study, a native green microalgal strain was isolated from a freshwater habitat in Bengaluru, India and subjected to microscopic observation to confirm its purity. Based on the cell morphology, the screened microalga was identified as *Chlorella*. As confirmatory evidence, PCR analysis was also carried out with universal eukaryotic forward and reverse primers to identify its genus specificity. A single distinct PCR amplicon band of ~1000 bp was detected when resolved on 2% Agarose gel, as shown in Figure 1. The resultant phylogenetic tree exhibited the evolutionary relationship of the selected strain with closely related microalgal isolates deposited in the NCBI GenBank. The algal strain showed 100% homology with the *Chlorella sorokiniana* isolates with accession number AM423162.1, EU402596.1, LK021940.1, KF444207.1, etc. After molecular characterization and phylogenetic analysis, the screened isolate was identified as *Chlorella sorokiniana*, belonging to the division Chlorophyta as shown in the generated tree [10].

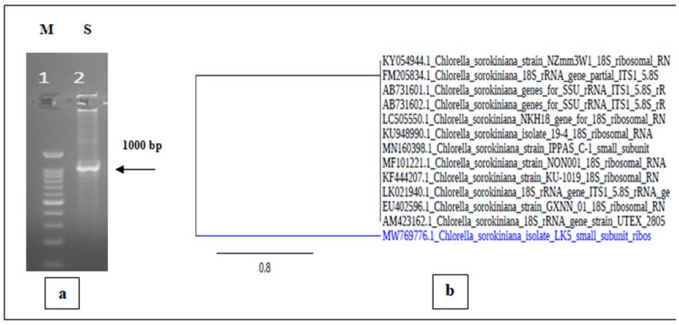

Figure 1. (A) Agarose gel electrophoretic pattern of PCR amplicon with 18S rRNA gene fragment in a native isolate of *Chlorella sorokiniana*. B) Phylogenetic relationship of isolated *Chlorella sorokiniana* (MW769776.1) based on small subunit ribosomal RNA gene.

Purification of *Chlorella Sorokiniana* Lectin (CSL) and analysis of molecular weight by Tricine SDS-PAGE

The crude extract of *Chlorella sorokiniana* MW769776 showed 32 fold hemagglutination activities towards chicken erythrocytes. However, no activity was found with human erythrocytes of A, B and O groups. Other *Chlorella* species such as *C. ellipsoidal* and *C. vulgaris* exhibited high hemagglutination activity with rabbit and human O erythrocytes. The crude extract of *Chlorella sorokiniana* MW769776 was pretreated by precipitation with 60% ammonium sulfate. The lectin purification was achieved in a single step by affinity chromatography on a guar-gum column. The elution profile of the purified lectin on the affinity column exhibited a single peak as shown in Figure 2. Fractions corresponding to the single peak were collected and dialyzed against 0.025 M PBS to remove contaminants and later subjected for hemagglutination test and showed positive activity. The yields and specific hemagglutination activities of the chromatographic fractions are listed in Table 1. The purity of active fractions was determined by Trichina SDS-PAGE in a 16% gel. The electrophoresis-based results suggested that the lectin existed as a monomer with a molecular weight of 16 kDa, which was estimated from the calibration curve (Figure 3). The recovery yield constituted approximately 27%, which was in agreement with the previously reported results from other *Chlorella* species. The results demonstrated strong inhibition of the hemagglutinating activity of CSL by glycoprotein yeast mannan and the simple sugars tested, including galactose, sucrose, and lactose indicating complex sugar specificity (Table 1). Similar carbohydrate-binding properties have been reported for other types of freshwater microalgal lectins, such as *Chlorella vulgaris, Tetradesmus obliquus, Chlorella sp.*, and *Scenedesmus obliquus* [11].

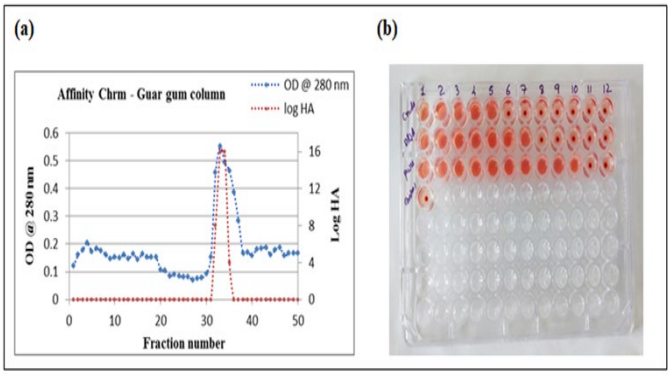

Figure 2. (A) Purification of *Chlorella Sorokiniana* Lectin (CSL) by affinity chromatography using cross-linked guar gum column (1.5 x 10 cm), equilibrated with 0.025M PBS (pH 7.2). Ammonium sulphate protein precipitate (P50-60) was applied to the column, flow rate adjusted to 0.5 mL/min. The adsorbed lectin was eluted with 0.2 M D-galactose in 0.025M PBS (pH 7.2, containing 0.15 M NaCl). Collected protein fractions were assayed for hemagglutinating activity using native 2% chicken erythrocytes. Absorbance at 280 nm; Hem agglutinating activity. CL=Lectin from *Chlorella* sp. HU=hem agglutinating units.

Sample	Volume (mL)a	HA (HU/mL)b	Total activity (axb)	Total protein (mg/mL) c	Specific activity (b/c)	Recovery Yield (%)d	Purification (fold)
Aqueous Extract (AE)	15	32	480	3	11	100	1
P 60-70	3	128	384	2.2	58	80	5
Affinity Column (AC)	0.1	1024	102	1.9	539	27	9

Table 1. Purification of lectin from biomass of *Chlorella sorokiniana* MW769776.

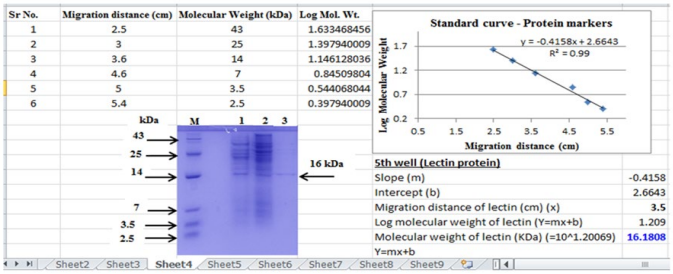

Figure 3. Calibration curve for the estimation of molecular weight of CSL; migration distance plotted against log molecular weight of standard marker proteins along with the photography of tricine SDS-PAGE of the purified lectin (CSL); electrophoresis was carried out on a 16% polyacrylamide gel in the presence of β-mercaptoethanol. The gel was stained with Coomassie brilliant blue R-250.

Identification of protein by LC-ESI-Q-TOF-MS analyses

The single protein band was excised from the 16% tricine SDS-PAGE, digested using trypsin and analyzed by LC-ESI-Q-TOF MS in the reflector mode. The peptide masses thus obtained for purified lectin were searched against the Mass Spectrometry Protein Sequence Database (MSDB) and the National Center for Biotechnology Information (NCBI) database using the MASCOT search engine. The results showed identity of peptide masses with previously submitted sequences of uncharacterized protein of *Chlamydomonas reinhardtii*. The database search using the taxonomy-all entries provided three protein sequences matched with the accession numbers as follows A0A2K3E1A2, A0A2K3DAQ2, and A0A2K3CQJ2, which showed a very less percentage of protein sequence coverage. The protein view of purified lectin Q1021 sequences exhibited

2% of protein sequence coverage with an uncharacterized protein of *Chlamydomonas reinhardtii* (A0A2K3CQJ2), with a Monoisotopic mass (Mr) of 15.48 kDa, and calculated pI of 5.77. The NCBI blast search showed the following features: The source was from *Chlamydomonas reinhardtii,* the product was an uncharacterized protein, and the calculated molecular weight of the protein was 15.30 kDa. The MS/MS fragmentation of tryptic peptides is shown in Figure 4, these amino acid sequences matched with the NCBI submitted sequence XP_042914775.1, which is an uncharacterized protein of Chlamydomonas reinhardtii. The tryptic peptides of the purified protein contained 38 amino acid residues, including WLPEPPATPATALPAGLGTVSAALPNSLRGSALDQALR. The fragmentation pattern of the tryptic peptides at observed m/z: 360.17, 397.21, 487.21, 495.21, 516.27, 643.29, 672.88, 907.44, 1071.53, 1619.73 was analyzed. The mass of 'y' and 'b' ions equal the mass of its amino acid residues. The 'y' ion represents the C-terminal fragment, whereas 'b' ion represents the N-terminal fragment. The peak at m/z 643.29 was the stronger peak of the y-cleavage peptide SLRGSALDQALR was observed, whereas the peak at m/z 487.21 was the strongest peak of the y-cleavage peptide QALR. The peak at m/z 397.21 was the weaker peak of the b-cleavage peptide WLP. The peaks at m/z 360.17, 516.27, 672.88, and 1071.53 corresponded to the b cleavage peptides WLPEPP, WLPEPPATPA, WLPEPPATPATAP, and WLPEPPATPATAPAGLGTVSAA, respectively. A BLAST search in UniProtKB revealed 100% similarity of the resultant peptide sequences with an uncharacterized protein of *Chlamydomonas reinhardtii* (A0A2K3CQJ2) and 61.5% identity with the Carbohydrate Binding Module1 (CBM1) domain-containing protein of *Verticillium longisporum* (A0A0G4N7X6) [12].

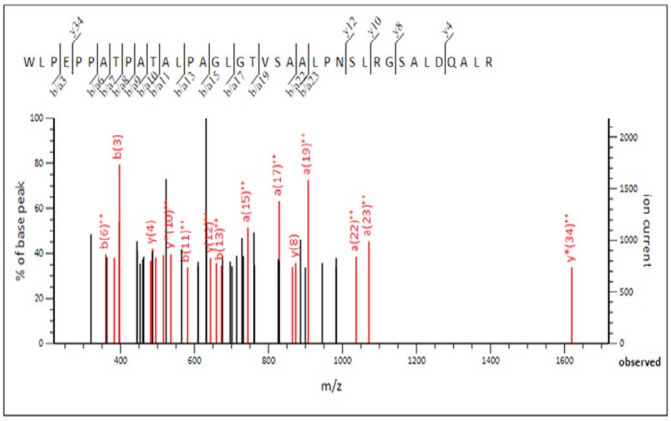

Figure 4. LC-ESI-Q-TOF-MS/MS spectrum of the tryptic peptide. Peptide corresponding to the sequences WLPEPPATPATALPAGLGTVSAALPNSLRGSALDQALR, Monoisotopic mass (Mr) 15.48 kDa. These sequences matched with NCBI submitted sequences XP_042914775.1 uncharacterized protein of *Chlamydomonas reinhardtii*.

Further confirmation evidence was accumulated by carrying out a multiple sequence alignment study for the obtained peptide sequences (A0A2K3E1A2, A0A2K3DAQ2, and A0A2K3CQJ2) with those of different lectins from other *Chlorella* species using CLUSTAL analysis. The peptide sequences of diverse lectins from algal species were selected randomly from the NCBI database under the taxonomy viridiplantae as shown in Table 2. The resultant peptide sequences obtained from A0A2K3E1A2 and A0A2K3DAQ2 exhibited no similarities with the aligned amino acid sequences of the different lectins. Whereas in the case of Q1021 peptide sequences acquired from A0A2K3CQJ2 showed slight similarities with the H-type lectin domain of Micractinium conductrix, which belongs to the Chlorellaceae family, as shown in Figure 5. Based on visual inspection of the peptide sequence alignment map suggests a minor level of similarity with the H-type lectin domain. However, amino acid substitutions were observed in peptide Q1021 (2LPEPPATPATA12, 21SAALPNSLRGSALDQ35) when compared with peptides A0A2P6UZH1, A0A2P6VI72, and A0A2P6VLT8 (34VADPLCQTCKT44, 53DSTLPDSETGKCLLR67). This was also confirmed by aligning the peptide sequences of *Chlorella sorokiniana* MW769776 purified protein with the amino acid sequences of the SUEL-type lectin domain containing protein of *Chlorella variabilis* (E1Z896) which showed 1.129% identity by hits of 9 identical positions as shown. A similar alignment study was performed with the peptide sequences of hypothetical proteins of *Auxenochlorella protothecoides* or *Chlorella protothecoides* (A0A087SU92), which showed 1.918% identity by hits of 16 identical positions as. Mass spectra and amino acid sequence analysis of tryptic peptides established the existence of amino acid substitutions among various isoforms of lectins from different sources [13].

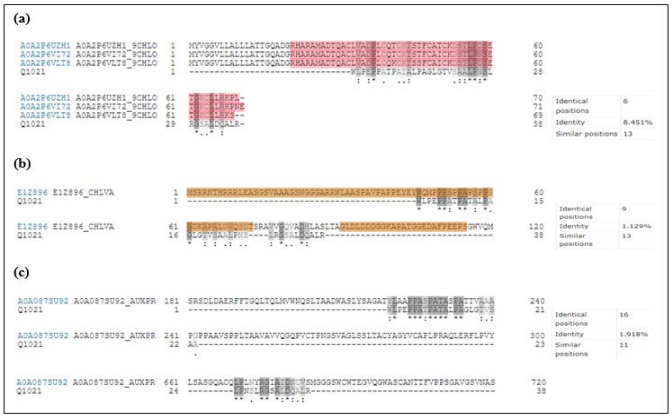

Figure 5. Multiple sequence alignment study using CLUSTAL. Identical positions in amino acids showed by asterisk sign, similar positions in amino acids showed by colon and dot. A) A0A2P6UZH1, A0A2P6VI72, A0A2P6VLT8=H-type lectin domain, *Micractinium conductrix*; Q1021=*Chlorella sorokiniana* MW769776; B) E1Z896=SUELtype lectin domain, *Chlorella variabilis*; C) A0A087SU92=Hypothetical protein, *Chlorella protothecoides*. Chain of amino acid residues is depicted in pink color, similarity is depicted in grey color, and region is depicted in brown color.

SL. No.	Accession Number	Name of the lectin	Source	Identity (%)
1	PRW33211.1	Rhamnose-binding lectin (SAL)	*Chlorella sorokiniana*	1.43
2	PRW20453.1	Concanavalin A-like lectin glucanase	*Chlorella sorokiniana*	1
3	PRW33614.1	H-type lectin domain	*Chlorella sorokiniana*	1
4	E1Z896	SUEL-type lectin domain-containing protein	*Chlorella variabilis*	1.13
5	A0A087SU92	Hypothetical protein F751_1672	*Auxenochlorella protothecoides*	1.92
6	PRW60571.1	Rhamnose-binding lectin	*Chlorella sorokiniana*	1
7	A0A2P6VDD7	Endoplasmic reticulum lectin1-like isoform A	*Micractinium conductrix*	4.07
8	A0A2P6VGB6	H-type lectin domain	*Micractinium conductrix*	9.46
9	A0A2P6TH56	H-type lectin domain	*Chlorella sorokiniana*	5.19
10	A0A2P6UZH1	H-type lectin domain	*Micractinium conductrix*	8.57
11	A0A2P6VI72	H-type lectin domain	*Micractinium conductrix*	9.86

12	A0A2P6VLT8	H-type lectin domain	*Micractinium conductrix*	8.57
13	A0A2P6V220	G-type lectin S-receptor-like serine threonine-kinase	*Micractinium conductrix*	5.26

Table 2. Multiple sequence alignment data.

FT-IR analysis

Fourier-Transform Infrared Spectroscope (FTIR) was used to study the secondary structure of the *Chlorella* lectin. It is a useful method to determine the type of functional group and the secondary structure of proteins. The results of the FTIR spectrum showed prominent peaks at 3625.61, 3417.95, 2091.31, 1650.96, 1644.19, and 1633.69 cm^{-1} (Figure 6). The strong and wide absorption band at 3625.61 cm^{-1} was the characteristic absorption band of the protein, attributed to the superposition between O-H and N-H stretching vibrations, which suggested the existence of intra-molecular hydrogen bonds, intermolecular hydrogen bonds, and amidogens in CSL [14]. The three absorption bands at 3700-2100 cm^{-1} indicated the asymmetric stretching vibration of methyl ($-CH_3$), asymmetric stretching vibration of methylene ($=CH_2$), and symmetrical stretching vibration of methyl ($-CH_3$), respectively. The absorption band at 1650.96 cm^{-1} (amide I band) was an important characteristic absorption band of protein, which indicated C=O stretching vibration. The band at frequency regions from 1633.69 cm^{-1} was assigned to β-sheet, accounting for 39.55%. The peak of 1644.19 cm^{-1} could be assigned to unordered structure. Unordered structure accounted for 39.38% of CSL. The α-helix absorption accounting for 39.23% appeared at 1650.96 cm^{-1} [15].

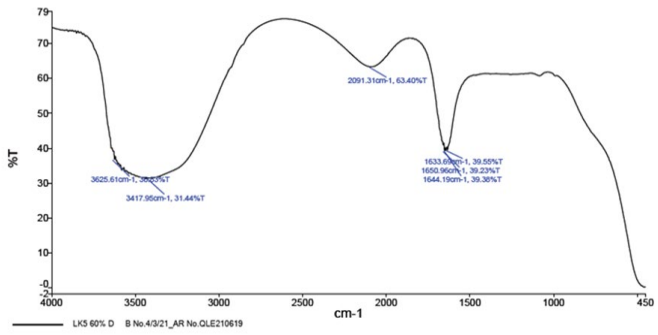

Figure 6. FTIR spectra of *Chlorella Sorokiniana* MW769776 lectin (CSL).

Temperature and pH stability

Purified CSL was stable at neutral pH up to 70°C for 30 min. It also exhibited tolerance under wide pH ranging from 6 to 11. Changes in pH may influence the ionization of amino acid residues in the lectin molecule, which can interfere with its binding to the carbohydrates. The CSL remained active over a wide pH range from 6.0 to 11.0 (Figure 7). Similar results were reported for *Chlorella pyrenoidosa* (4.0-11.0), marine red alga *Meristiella echinocarpa* (5.0-10.0), and *Bryothamnion triquetrum* (4.0-11.0). Algal lectins generally have high molecular stability owing to the presence of disulfide bonds and small molecular masses, hence able to maintain the hemagglutinating activity over a wide pH range. The CSL hemagglutinating activity resisted heating at 70°C for up to 30 min, but 50% of the activity was reduced by heating beyond 70°C for 30 min. Whereas, hemagglutinating activity was completely arrested at 100°C after 30 min. Similarly, lectins from other freshwater microalgae such as Tetradesmus obliquus, and marine red algae, such as Gracilaria ornate and *Pterocladiella capillacea* were denatured at 70°C. Therefore, it can be inferred that the CSL structure is as stable as that of other algal lectins, was previously reported [16].

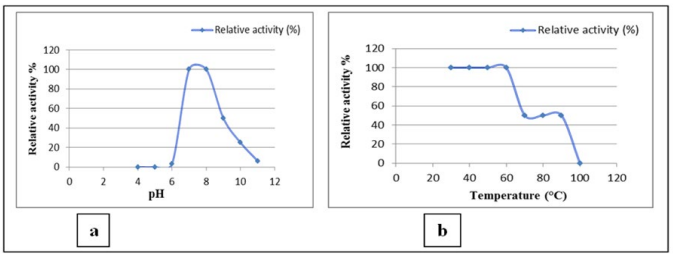

Figure 7. Effect of pH A) Temperature B) Hemagglutination activity of *Chlorella Sorokiniana* MW769776 lectin (CSL).

The CSL hemagglutinating activity was not inhibited by divalent metal ions, such as Ca^{2+}, but was slightly inhibited by Mg^{2+} when compared to the initial hemagglutination activity (4 HU), showing its sensitivity to this ion. These data are in contrast with a previous report on lectin from Tetradesmus obliquus, which showed inhibition with Fe^{2+}, Ca^{2+}, Mg^{2+}, Zn^{2+}, and Cu^{2+}. Similarly, marine red algae such as *Gracilaria ornate* and *Pterocladiella capillacea* depend on metal ions for their hemagglutination activity [17].

Molecular modeling

Molecular modeling of related proteins is important for understanding the role of amino acid residues, which are associated with their structural and functional properties. In the present study, the 3D structure of CSL was obtained by homology modeling using SWISSMODEL. The structure of CSL appeared to superimpose with the photosynthetic light harvest-reaction center (LH1-RC) super complex of *Thiorhodovibrio* strain 970. The structure of CSL was 22.22% identical to the LH1-RC protein complex and overall Root Mean Square Deviation (RMSD) value was 2.8 Å as shown in Figure 8. Generally, the acceptable range of RMSD when the target protein overlaps with the template is less than 2 Å, which means it is considered to be more similar to each other. In this case, the resultant RMSD value suggested that the similarity between CSL and the LH1-RC super complex of *Thiorhodovibrio* strain 970 was quite low. In addition, the CSL amino acid residues did not show any similarity with other types of lectin proteins that have been submitted in the database. Previous studies have reported that in photosynthetic organisms, lectin family proteins play important roles in capturing and releasing photosynthetic *via* an endogenous lectin cycle. Notably, a stromal thylakoid membrane lectin from unicellular green alga *Dunaliella salina* was shown to be associated with the light-harvesting complex of photosystem I. A computational analysis of known amino acid sequences revealed that the light-harvesting complex of photosystem I of *Dunaliella salina* contains a fragment that was homologous to the carbohydrate-binding site of the galactose-specific lectins. Based on these data, we can strongly claim that the functional property of CSL might be associated with the LH1-RC of photosystem I. Since the purified protein did not show much similarity with other types of lectins, hence it can be considered as a unique galactose-binding lectin protein of *Chlorella sorokiniana* MW769776 [18].

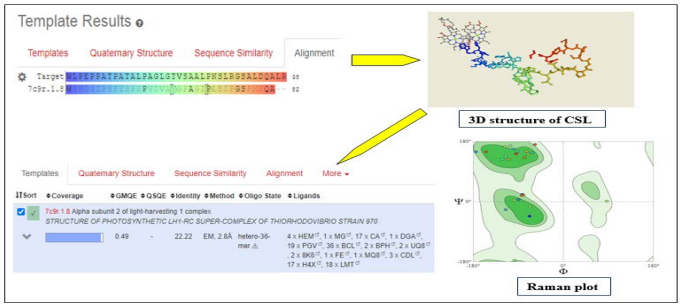

Figure 8. The 3D model of the galactose-binding lectin of *Chlorella sorokiniana* MW769776 (CSL) generated by SWISSMODEL in automated mode using LH1-RC super complex of *Thiorhodovibrio* strain 970 as a template.

Conclusion

In the present study, a novel galactose-binding lectin with a molecular weight of 16 kDa from *Chlorella sorokiniana* MW769776 (CSL) was successfully isolated and purified. The CSL hemagglutination activity was inhibited by D-galactose, and its activity was not influenced by metal ions such as $Mg^{2+}$and Ca^{2+}. It was stable in alkaline buffer solution. The amino acid sequences of tryptic peptides contained abundant proline, threonine, glycine, leucine, and serine. The secondary structure of CSL mainly contained a turn structure, whereas the 3D structure revealed that the functional property of CSL was related to the LH1-RC of photosystem I. Therefore, we conclude that the lectin identified in the present study from *Chlorella sorokiniana* MW769776, may be a novel protein that has not yet been reported. This protein can be used for further single-molecule interaction studies because of its low molecular weight and bioactivity mechanism that would be worth to conduct in the future. These findings have been submitted to the claim of an Indian patent.

Acknowledgement

The authors wish to gratefully acknowledge University Grants Commission (UGC)-Dr. D.S. Kothari Postdoctoral fellowship scheme (No.F.4-2/2006(BSR)/BL/18-19/0159), India, for providing financial support to carry out this research work and Bangalore university for allowing us to utilize the laboratory facilities at the department of microbiology, biotechnology and food technology, Bengaluru, India. They wish to acknowledge the Indian Institute of Science (IISC), Bengaluru, India, for having carried out the LC-ESI-Q-TOF-MS/MS analysis.

Data Availability Statement

Data available on request from the authors.

Conflicts of Interest

"The authors declare no potential conflicts of interest."

References

1. Goldstein, Irwin J, R Colin Hughes, Michel Monsigny, and Toshiaki Osawa, et al. "What should be called a lectin?." *Nature* 285 (1980): 66-66.
2. Dodd, Roger B, and Kurt Drickamer. "Lectin-like proteins in model organisms: implications for evolution of carbohydrate-binding activity." *Glycobiology* 11 (2001): 71-79.
3. Silva, Mariana CC, Lucimeire A Santana, Reinhard Mentele, and Rodrigo S Ferreira, et al. "Purification, primary structure and potential functions of a novel lectin from *Bauhinia forficata* seeds." *Process Biochem* 47 (2012): 1049-1059.
4. Lagarda-Diaz, Irlanda, Ana M Guzman-Partida, Gabriel Urbano-Hernandez, and M Magdalena Ortega-Nieblas, et al. "Insecticidal action of PF2 lectin from *Olneya tesota* (Palo Fierro) against *Zabrotes subfasciatus* larvae and midgut glycoconjugate binding." *J Agric Food Chem* 57 (2009): 689-694.
5. Seriramalu, Ramarao, Wei Wei Pang, Jaime Jacqueline Jayapalan, and Emida Mohamed, et al. "Application of champedak mannose-binding lectin in the glycoproteomic profiling of serum samples unmasks reduced expression of alpha-2 macroglobulin and complement factor B in patients with nasopharyngeal carcinoma." *Electrophor* 31 (2010): 2388-2395.
6. Macedo, Maria Ligia Rodrigues, Maria das Gracas Machado Freire, Maria Barbosa Reis da Silva, and Luana Cassandra Breitenbach Barroso Coelho, et al. "Insecticidal action of *Bauhinia monandra* leaf lectin (BmoLL) against *Anagasta kuehniella* (Lepidoptera: Pyralidae), *Zabrotes subfasciatus* and *Callosobruchus maculatus* (Coleoptera: Bruchidae)." *Comp Biochem Physiol a Mol Integr Physiol* 146 (2007): 486-498.
7. Lagarda-Diaz, Irlanda, Ana Maria Guzman-Partida, and Luz Vazquez-Moreno. "Legume lectins: proteins with diverse applications." *Int J Mol Sci* 18 (2017): 1241-1242.
8. de O Mello, Erica, Izabela S dos Santos, Andre de O Carvalho, and Luisa S de Souza, et al. "Functional expression and activity of the recombinant antifungal defensin PvD1r from *Phaseolus vulgaris* L.(common bean) seeds." *BMC Biochem* 15 (2014): 1-13.
9. Lusvarghi, Sabrina, and Carole A. Bewley. "Griffithsin: an antiviral lectin with outstanding therapeutic potential." *Viruses* 8 (2016): 295-296.
10. Shen, Li, Zhiguo Luo, Junbo Wu, and Li Qiu, et al. "Enhanced expression of α2, 3-linked sialic acids promotes gastric cancer cell metastasis and correlates with poor prognosis." *Int J Oncol* 50 (2017): 1201-1210.
11. Blomme, Bram, Christophe van Steenkiste, Nico Callewaert, and Hans van Vlierberghe, et al. "Alteration of protein glycosylation in liver diseases." *J Hepatol* 50 (2009): 592-603.
12. Fekete, Ferenc, Balazs Fadgyas, Eva Papp, and Agnes Szilagyi, et al. "The role of mannose binding lectin on fever episodes in pediatric oncology patients." *Pathol Oncol Res* 22 (2016): 139-143.
13. Nunes, Breno S, Nathalie S Rensonnet, Daniela Dal-Secco, and Silvio M Vieira, et al. "Lectin extracted from Canavalia grandiflora seeds presents potential anti-inflammatory and analgesic effects." *Naunyn Schmiedebergs Arch Pharmacol* 379 (2009): 609-616.
14. Liao, WR, and R Huang. "Agglutination of human and animal erythrocytes in marine unicellular algae." *J Ind Microbiol Biotechnol* 24 (2000): 262-266.
15. Singh, Ram Sarup, and Amandeep Kaur Walia. "Lectins from red algae and their biomedical potential." *J Appl Phycol* 30 (2018): 1833-1858.
16. Silva, AJ, VLR Cavalcanti, ALF Porto, and WAGama, et al. "The green microalgae *Tetradesmus obliquus* (*Scenedesmus acutus*) as lectin source in the recognition of ABO blood type: purification and characterization." *J Appl Phycol* 32 (2020): 103-110.
17. Cavalcanti, Vivianne Lays Ribeiro, Romero Marcos Pedrosa Brandao-Costa, Emmanuel Viana Pontual, and Alexsandra Frazao de Andrade, et al. "*Chlorella vulgaris* lectin kills Aedes aegypti larvae." *Algal Res* 56 (2021): 102290.
18. Naeem, Aabgeena, M Saleemuddin, and Rizwan Hasan Khan. "Glycoprotein targeting and other applications of lectins in biotechnology." *Curr Protein Pept Sci* 8 (2007): 261-271. .

How to cite this article: J, Savitha, Roopashri N Arekal, Ashwini Kumar, Gunashree BS and Divyashree MS. "Mass Spectrometric Characterization of a Novel Galactose-Binding Lectin from Chlorella sorokiniana (MW769776)." *J Bioprocess Biotech* 13 (2023): 573.