Characterization of novel lectins from Burkholderia pseudomallei and Chromobacterium violaceum with seven-bladed β-propeller fold

Petra Sýkorová, Jitka Novotná, Gabriel Demo, Guillaume Pompidor, Eva Dubská, Jan Komarek, Eva Fujdiarová, Josef Houser, Lucia Hároníková, Annabelle Varrot, et al.

Characterization of novel lectins from *Burkholderia pseudomallei* and *Chromobacterium violaceum* with seven-bladed β-propeller fold

Petra Sýkorová[a,b], Jitka Novotná[b,c], Gabriel Demo[b,c,1], Guillaume Pompidor[d], Eva Dubská[b], Jan Komárek[b,c], Eva Fujdiarová[b,c], Josef Houser[b,c], Lucia Hároníková[a], Annabelle Varrot[e], Nadezhda Shilova[f], Anne Imberty[e], Nicolai Bovin[f], Martina Pokorná[b,c], Michaela Wimmerová[a,b,c,*]

michaw@chemi.muni.cz

[a]Department of Biochemistry, Faculty of Science, Masaryk University, Brno, Czech Republic

[b]Central European Institute of Technology, Masaryk University, Brno, Czech Republic

[c]National Centre for Biomolecular Research, Faculty of Science, Masaryk University, Brno, Czech Republic

[d]EMBL Hamburg c/o DESY, Hamburg, Germany

[e]CERMAV, CNRS Université de Grenoble Alpes, Grenoble, France

[f]Shemyakin–Ovchinnikov Institute of Bioorganic Chemistry, Russian Academy of Sciences, Moscow, Russian Federation

[1]current address: RNA Therapeutics Institute, Department of Biochemistry and Molecular Pharmacology, University of Massachusetts Medical School, Worcester, Massachusetts, USA

[*]Corresponding author.

ABSTRACT

Burkholderia pseudomallei and *Chromobacterium violaceum* are bacteria of tropical and subtropical soil and water that occasionally cause fatal infections in humans and animals. Microbial lectins mediate the adhesion of organisms to host cells, which is the first phase in the development of infection. Here we report the discovery of two novel lectins from the above-mentioned bacteria - BP39L and CV39L. The crystal structures revealed that the lectins possess a seven-bladed β-propeller fold. Functional studies conducted on a series of oligo- and polysaccharides confirmed the preference of BP39L for mannosylated saccharides and CV39L for rather more complex polysaccharides with a monosaccharide preference for β-L-fucose. The presented data indicate that the proteins belong to a currently unknown family of lectins.

Keywords

Lectin; *Burkholderia pseudomallei*; *Chromobacterium* v*iolaceum*; Protein structure; Seven-bladed β-propeller fold

1. Introduction

The bacterium *Burkholderia pseudomallei* is a rod-shaped bacterium found in contaminated water and soil, which belongs to the *Burkholderia cepacia* complex (Bcc). It is a group of Gram-negative bacteria which includes at least 21 related species [1, 2, 3] from which many are resistant to antibiotics [4]. The genome of *B. pseudomallei* is composed of 2 chromosomes with 4.07 and 3.17 megabase pairs [5]. This human pathogen causes a disease called melioidosis or Whitmore's disease. Melioidosis is endemic in East Asia and north Australia. It is a febrile illness with a mortality rate of 40% connected with the production of abscesses in the lungs and also in

the liver, spleen, skeletal muscles and prostate [6]. Due to its biofilm formation and resistance to antibiotics, the treatment is very problematic and often unsuccessful, with currently no existing vaccine [7, 8]. The bacterium can survive in a hostile environment for years [9]. *B. pseudomallei* is also closely related to *Burkholderia mallei* [8]. *B. mallei* is an obligate mammalian pathogen causing the disease glanders, which occurs mainly in horses [10].

Chromobacterium v*iolaceum* is a Gram-negative bacterium living in tropical and subtropical water and soil. The bacterium is saprophytic and an opportunistic pathogen of humans causing sepsis, liver and lung abscesses with high mortality [11]. This β-proteobacterium was described at the end of the 19th century, and the complete genome of *C.* v*iolaceum* strain ATCC 12472 was sequenced in 2003. The genome consists of a single circular chromosome with 4.75 megabase pairs [12]. The *C.* v*iolaceum* genome is most similar to the plant pathogen *Ralstonia solanacearum* [13], and almost 10% of *C. violaceum* genes are closely related to genes of the opportunistic human pathogen *Pseudomonas aeruginosa* [14].

Lectins are proteins which are capable of specific and reversible binding to carbohydrate moieties. They usually consist more than one carbohydrate binding site per protein molecule and such multivalency favors the strong avidity to glycoconjugates available on cell surfaces. The carbohydrate-mediated recognition plays an important role in the ability of pathogenic bacteria to adhere to the surface of the host cell in the first step of their invasion and infectivity [15]. The first characterized lectin from *C. violaceum* was named CV-IIL [16, 17]. CV-IIL is a homotetrameric fucose-binding lectin and is homologous to the fucose-specific PA-IIL (LecB) from *P. aeruginosa* [18] and the mannose-specific RS-IIL from *R. solanacearum* [16, 19].

While many lectins reach multivalency through oligomerization, for so-called β-propeller lectins, multivalency is created by repeats of conserved carbohydrate binding domains [20]. The

repeats with multiple binding sites are typically located on one face of the β-propeller donut shape. Therefore, topologically the carbohydrate binding domains are perfectly positioned to bind carbohydrate epitopes on cell surfaces to maximize the lectin adherent function.

Several six- or seven-bladed β-propeller lectins were found in fungi, such as *Psathyrella velutina* (PVL) [21] or *Agrocybe aegerita* (AAL2) [22], belonging to PVL-like lectin family, and in *Aleuria aurantia* (AAL) [23], being the member of AAL-like lectin familly. They are likely to be involved in defense against worms feeding on mushrooms. Other β-propeller lectin AFL (AAL-like family) has been identified in pathogenic fungus *Aspergillus fumigatus* [24]. This fucose-specific lectin plays a role in eliciting host immune response [25, 26]. Considering Prokaryota, there were described six-bladed β-propeller lectins in *Ralstonia solanacearum* (RSL) [27] and human pathogen *Burkholderia ambifaria* (BambL) [28], both belonging to AAL-like family. A separate family of seven-bladed β-propeller bangle lectins have been characterized by lectins from two species of *Photorhabdus* bacteria (PHL and PLL) with evidence for dual specificity for fucose and galactose and probable role in interaction between bacteria and their hosts [29, 30].

Since some of the β-propeller lectins are involved in host-pathogen recognition, they represent promising targets for glycomimetics with anti-infectious properties to compete with host recognition. The multivalent potential of these lectins in pathogenic microorganisms has been a key to design high-affinity inhibitors [31, 32, 33].

In this paper we structurally and biochemically characterize two novel lectin members: BP39L and CV39L from *B. pseudomallei* and *C.* v*iolaceum*, respectively. The crystal structures reveal that the proteins possess a seven-bladed β-propeller fold. The saccharide specificity of both lectins has been determined and the putative saccharide binding sites were characterized.

The knowledge about their structure together with sequence and structure predictions indicate that the proteins belong to an undescribed family of lectins.

2. Materials and methods

2.1. Preparation of Constructs

The genes coding for BP39L and CV39L flanked with the NdeI and HindIII restriction sites were synthesized by GeneArt (ThermoFisher Scientific). The N-terminus of proteins was designed to carry both the His_6-tag and the specific site (ENLYFQS) recognized by the TEV protease. After digestion with NdeI and HindIII, the DNA sequences were cloned into the pET-25(b+) vector (Novagen, Madison, United States). The constructs *pET25bp39l* and *pET25cv39l* were transformed into *E. coli* Tuner (DE3) cells (Novagen, Madison, United States).

2.2. Protein Expression and Purification

E. coli Tuner(DE3) cells containing the *pET25bbp39l* or *pET25bcv39l* plasmid were grown in LB medium with 100 $\mu g \cdot ml^{-1}$ ampicilin until OD_{600} reached 0.5. Gene expression was initiated with 1 mM IPTG (Isopropyl β-D-1-thiogalactopyranoside). After a 16-hour gene expression at 18°C, the cultures were harvested, centrifuged at 12 000 g at 4°C for 10 min and resuspended in the buffer containing 20 mM Tris/HCl, 100 mM NaCl, pH 7.5. Cells were disintegrated by ultrasonic vibration and centrifuged at 21,000 g at 4°C for 60 min. The standard yield was 5 mg and 90 mg per litre of cell culture for BP39L and CV39L, respectively.

The recombinant proteins were purified by chelating chromatography (IMAC) using a Ni-NTA column (Qiagen, Hilden, Germany). Cell extract was loaded onto a column in 20 mM Tris/HCl, 200 mM NaCl, pH 7.5, for the protein BP39L. Unbound proteins were washed off with

the running buffer. After washing, BP39L was eluted in the same buffer with 60 mM imidazole. For the purification of CV39L, buffer containing 20 mM Tris/HCl, 300 mM NaCl, pH 7.5 was used and the protein was eluted with 210 mM imidazole. Both proteins were also purified in a mannose-agarose column. In this case, the cell extract was loaded onto a column in 20 mM Tris/HCl, 100 mM NaCl, pH 7.5. The protein was eluted from the column with loading buffer containing 100 mM D-mannose. Protein purity was assessed by 12% SDS-PAGE and mass spectrometry. The proteins were dialyzed against 20 mM Tris/HCl, 20 mM NaCl, pH 7.5 for 2 days (IMAC purification) and 5 days (mannose-agarose purification). Proteins were lyophilized and stored at -80°C or used immediately and concentrated with Macrosep Centrifugal Devices (3 kDa molecular weight cut-off; Pall, Port Washington, United States).

2.3. Microcalorimetry Experiments

ITC experiments were performed with MicroCal ITC_{200} and MicroCal Auto-ITC_{200} (Malvern). Protein concentrations used for ITC experiments varied from 50 to 500 μM and ligand solution from 5 to 45 mM, respectively. Injections of 0.5-2.0 μl aliquots of a carbohydrate solution dissolved in the same buffer as the protein (20 mM Tris/HCl, 20 mM NaCl, pH 7.5 for BP39L and 20 mM Tris/HCl, 150 mM NaCl, pH 7.5 for CV39L) were added automatically at 3-min intervals to the protein solution present in the calorimeter cell. The temperature of the cell was maintained at 25 ± 0.1°C. Injections of the ligand into the buffer solution were used as control experiments. The ITC experiments were performed in triplicates.

Data were fitted with the MicroCal Origin 7 software using a one binding site model. The association constants (K_a) and enthalpies (ΔH) of binding were obtained from the data fitted to a theoretical titration curve.

2.4. Glycan Microarray

The affinity of the lectins towards mammalian glycans and bacterial polysaccharides was determined using a glycan microarray constructed by Semiotik LLC (Moscow, Russia), similar to an array described earlier [34]. Pure BP39L and CV39L were labelled with the fluorescent dye DyLight 488 NHS Ester (Thermo Scientific, Waltham, United States) according to the product manual. The array was composed of more than 600 tested glycans, each in 6 replicates. Aliquots of BP39L and CV39L solution (50 μg/ml) in PBS (0.01 M Na_2HPO_4, 0.01 M NaH_2PO_4, 0.138 M NaCl, and 0.0027 M KCl, pH 7.4) containing 1 % BSA and 1% Tween-20 were applied to slides (pre-treated with PBS containing 0.1 % Tween-20 for 15 min) and incubated in a humidified chamber at 37°C for 1 h. After washing with PBS containing 0.05 % Tween-20 and deionized water, slides were scanned with an InnoScan1100 AL (Innopsys, Carbonne, France) with a 488 nm laser at 100 PMT and high laser power mode. The obtained data were processed using the software Mapix 7.3.1 and Mapix 8.2.2 (Innopsys, France) and the fixed 100 μm-diameter ring method. Data are reported as the relative fluorescence units (RFU) median of six spot replicates. The signals with fluorescence intensity exceeding the background value by a factor of five were considered to be significant.

2.5. Crystallization and Data Collection

Crystallization trials were performed with Qiagen crystallization screens Classic, Classic II, Classic Lite, PEGs I, PEGs II, PACT, AmSO4 and Compass using the sitting drop technique. Optimized crystals were obtained by the hanging-drop vapor diffusion method using 2 - 4 μL drops containing a 50:50 and/or 34:66 (v/v) mix of a reservoir solution combined with BP39L or

CV39L solution at 17°C. The solution of BP39L had a protein concentration of 12 mg/ml in 20 mM Tris/HCl, 20 mM NaCl, pH 7.5. The crystals of BP39L were obtained after two months with the following initial conditions: 0.1 M MES, 15% PEG 5000 MME, 0.2 M ammonium sulphate, pH 6.5. Optimization of the initial crystallization conditions yielded crystals with the precipitant solution 0.1 M MES, 10 - 15% PEG 6000, 0.2 M ammonium sulphate, pH 6.5. PEG 6000 (28%) for BP39L was added to the mother liquor as a cryoprotectant prior to crystal freezing in a mounted loop at 100 K. A europium derivative was prepared by soaking a crystal in a 50 mM solution of Eu-DTPA-BMA (NetX-ray, Inc.) [35] for 24 hours.

The solution of CV39L protein had a concentration of 5 mg/ml in the same buffer as for BP39L. The crystals of CV39L were obtained after two weeks under the following crystallization conditions: 0.1 M sodium acetate, 30% PEG 2000 MME or 25% PEG 4000, 0.2M ammonium sulphate, pH 4.6. Optimization of the initial crystallization conditions yielded crystals with a precipitant solution of 0.1 M sodium acetate, 30% PEG 2000 MME, 0.2M ammonium sulphate, pH 4.6. For CV39L PEG 400 (40%) was used as a cryoprotectant prior to crystal freezing in a mounted loop at 100 K. A europium derivative was prepared by soaking a crystal in a 50 mM solution of Eu-DTPA-BMA (NetX-ray, Inc.) [35] for 10 minutes.

Native and derivative diffraction data were collected on the P13 EMBL beam line of PETRA III (Hamburg, Germany) with the use of the PILATUS detector (Dectris Ltd., Switzerland). Image processing was performed using XDS [36] before being scaled and converted to structural factors using SCALA in the CCP4 program suite [37]. The initial structure of BP39L was solved using SAD data (x-ray wavelength of 1.776 Å) with the help of the program suite hkl2map (structure determination, phasing and solvent flattening) [38] together with the automatic model building module of the ARP/wARP program package [39].

For SAD data (x-ray wavelength of 1.776 Å) of the CV39L lectin, the automatic structure solution pipeline Auto-Rickshaw [40] was used. All further computing was performed using the CCP4 [37] suite unless otherwise stated.

2.6. Structure Determination

A molecular replacement technique was used to solve the two native structures with MOLREP [41] applying the monomeric coordinates of the initial BP39L and CV39L structure from the SAD data collection. Crystallographic refinement was carried out with the program REFMAC (Version 5.8.0049) [42] and manual model building was done with Coot (Version 0.7.2) [43]. Water molecules were introduced automatically using Coot and inspected manually. Alternative conformations were constructed where necessary (with occupancies estimated from the refined relative B-factors of the conformations). The stereochemical quality of the models was assessed with the programs Procheck [44] and Molprobity [45]. Molecular drawings were prepared using PyMOL (The PyMOL Molecular Graphics System, Version 1.5, Schrödinger, LLC). The final models of the BP39L and CV39L lectins were deposited in the PDB database under ID codes 6GXR and 6GXS.

2.7. Analytical ultracentrifugation

Analytical ultracentrifugation experiments were performed using a ProteomeLab XL-A analytical ultracentrifuge (Beckman Coulter) equipped with An-60 Ti rotor. Before analysis, purified proteins BP39L and CV39L were brought into the experimental buffer (20 mM Tris/HCl, 150 mM NaCl, pH 7.4) by dialysis and the dialysate was used as an optical reference.

Sedimentation velocity experiments were conducted in a standard double-sector centerpiece cell loaded with 420 μl of protein sample (0.06-0.38 $mg \cdot ml^{-1}$ for BP39L and 0.05-0.32 $mg \cdot ml^{-1}$ for CV39L) and 430 μl of reference solution. Data were collected using absorbance optics at 20°C at a rotor speed of 42,000 rpm. Scans were performed at 280 nm at 10 min intervals and 0.003 cm spatial resolution in continuous scan mode. The partial specific volume of proteins together with solvent density and viscosity were calculated from the amino acid sequence and buffer composition, respectively, using the software Sednterp 1.09. The sedimentation profiles were analyzed with the program Sedfit 14.3 [46]. A continuous size-distribution model for non-interacting discrete species providing a distribution of apparent sedimentation coefficients was used.

Sedimentation equilibrium experiments were performed at 20° C in a six-channel centerpiece cell loaded with 110 μl of protein samples (0.05-0.18 $mg \cdot ml^{-1}$ for BP39L and 0.07 and 0.14 $mg \cdot ml^{-1}$ for CV39L) and 120 μl of reference solution. Samples were gradually spun at three different rotor speeds in order to generate concentration gradients with shapes that include both shallow and steeper gradients (the rotor speeds of 14,000 rpm, 16,700 rpm and 29,000 rpm for BP39L and 9,000 rpm, 11,000 rpm, and 18,500 rpm for CV39L). After the equilibrium was achieved, data were collected at 280 nm by averaging 20 replicates with 0.001 cm spatial resolution in step mode. Data from the multi-speed experiment were globally analyzed with SEDPHAT 10.58 [47] using the model for one non-interacting discrete species.

2.8. Circular Dichroism Spectroscopy and Thermal Stability Experiments

CD measurements were performed using a Jasco J-815 spectropolarimeter (Jasco, Hachioji, Japan) at 20°C in 0.1 cm path-length quartz cuvettes with 4.5 μM BP39L or CV39L in

buffer containing 20 mM Tris/HCl, 20 mM NaCl, pH 7.5. The CD spectra were recorded from 200 to 260 nm. Each experiment was repeated five times and spectra were averaged. All spectra were carried out at the interval of 0.1 nm with a scanning speed 1 nm/min, 1 s response time and 1 nm bandwidth.

To evaluate the effect of ligands and bivalent ions on lectin stability, thermal dependent CD measurements were performed. 4.5 μM BP39L was mixed with 31.5 μM D-mannose and 4.5 μM CV39L with 31.5 μM L-fucose or proteins with/without a ligand were mixed with 45 μM $CaCl_2$, $MgCl_2$ or $MnCl_2$. The mixtures were incubated for 16 hours at 4°C. The thermal stability of proteins was measured in 0.1cm path-length cuvettes. The protein solutions were heated from 20°C to 80°C at a rate of 1°C/min with a scanning speed of 20 nm/min. The changes in ellipticity were monitored at 210 nm (BP39L) or 230 nm (CV39L). The thermal experiments were repeated three times and spectra were averaged. The temperature of the cell holder was regulated using a JASCO PTC-423L temperature controller.

2.9. Molecular docking

For docking experiments, we used ICM-Pro modelling software [48] based on the based on the Internal Coordinate Mechanics (ICM) approach (Molsoft, L.L.C., La Jolla, USA). Both X-ray crystal structures (BP39L; PDB 6GXR and CV39L; PDB 6GXS] were converted into an ICM object (i.e. prepared for use in ICM environment) using the standard ICM procedure [49]. The protein atoms were assigned to the correct atom types and charge based on a modified ECEPP/3 force field[50], the ligand atom types and charges were assigned based on the Merck molecular force field (MMFF94)[51].Missing hydrogen atoms were added. Side chains with added atoms and polar hydrogen atoms, or side chains with multiple tautomeric or crystallographically

ambiguous rotational conformations such as asparagine, glutamine and histidine, were sampled and optimized. The grid box was centred to the binding cavity for each of the lectin. Standard ICMdock grid potential maps within 3D box that encompass the ligand-binding pocket residues were calculated with a 0.5 grid spacing [52]. These maps represent electrostatics, hydrophobicity, hydrogen bonding, and the soft van der Waals potentials for hydrogens and for heavy atoms. Semi-flexible docking, which keeps the protein itself rigid while the ligand is flexible, was performed. ICM uses a Monte Carlo global optimization procedure[53] to predict binding poses for the ligand in the binding pocket. Optimised 3D structures of saccharides were downloaded from [54] or were built in the carbohydrate builder of the GLYCAM web interface [55] and automatically assigned charges by ICM where necessary. The compounds were then docked into the predicted lectin binding using the batch docking method of ICM. The docking results were evaluated by the ICM scoring function, which is weighted according to the parameters like, internal force-field energy, entropy loss between bound and unbound states of the ligand, ligand receptor hydrogen bond interactions, polar and non-polar solvation energy differences between bound and unbound states, electrostatic energy, hydrophobic energy, and hydrogen bond donor or acceptor desolvation.

3. Results

3.1. CV39L and Related Hypothetical Lectins

The CV39L lectin was discovered during the affinity purification of a *C. violaceum* CCM 2076 culture (obtained from the Czech Collection of Microorganisms) on a mannose-agarose column. The protein was identified by LC-MS/MS as the hypothetical protein CV_1052 from the

C. violaceum ATCC 12472 translated genome (accession number WP_011134607). It consisted of 363 amino acids (including the initial methionine) and was named CV39L. Using a BLAST search, a 357-amino-acid hypothetical lectin from *B. pseudomallei* was identified with 26% sequence identity (55% similarity) and was named BP39L (accession number WP_004536448). Hypothetical proteins from the related species *B. mallei* and *B. ubonensis* have strong sequence similarity with BP39L (sequence identity of 99% and 93%, respectively), while hypothetical proteins from *C. vaccinii* and *Pseudogulbenkiania ferrooxidans* are very similar to CV39L (sequence identity of 98% and 96%, respectively) (Fig. S1). Based on multiple sequence alignment analysis of multiple known β-propeler lectins and BP39L and CV39L (Fig. S2), these proteins are distantly similar to recently described lectins PLL and PHL from *Photorhabdus* family, however they can be considered as an individual lectin family.

To study the role of the BP39L and CV39L lectins, the genes were commercially synthesized. The BP39L and CV39L lectins were produced in a recombinant form in *Escherichia coli*. The proteins were purified by immobilized metal ion affinity chromatography or by mannose-agarose column and the protein purity and identity was verified by SDS-PAGE and mass spectrometry analysis (Fig. S3).

3.2. Oligosaccharide Specificity

To obtain a broad view of potential ligands, a glycan microarray consisting of over 600 different glycans and bacterial polysaccharides (Semiotik LLC., Moscow, Russia) was used. The results (Fig.1 and Tables S1-S3) revealed the strongest binding of BP39L to the Manα1-6(Manα1-3)Manα1-6(Manα1-3)Manβ pentasaccharide. Additionally, BP39L had a strong binding response to the Lewis b epitope and the Manα1-6(Manα1-3)Manβ trisaccharide and a weaker response to

the Manα1-6Manβ disaccharide. For CV39L, the strongest binding was observed to Manβ1-4GlcNAc and L-β-fucose. The next significant binding was observed to Glcα1-4Glcβ epitope (D-maltose). As for bacterial exopolysaccharides (Tables S1-S3), BP39L only exhibited a weak affinity for *E. coli* O73, whereas CV39L bound strongly to several exopolysaccharides from Gram-negative bacteria (*Shigella flexneri*, *Escherichia coli*, *Cronobacter sakazakii*, *Aeromonas hydrophila*). Most of these CV39L-recognized exopolysaccharides contain α-L-rhamnose in their core saccharide sequence.

To characterize the interactions between lectins and saccharides, isothermal titration microcalorimetry (ITC) was performed. A series of mono-, di-, trisaccharides and some bacterial exopolysaccharides were tested (Fig. 2 and Table 1). BP39L binding to branched mannotriose (Manα1-6(Manα1-3)Man) was confirmed (Fig. 2). Titration curves indicate a medium affinity with a dissociation constant of 190 μM. CV39L exhibited interactions with the monosaccharides L-galactose and L-fucose within the same affinity range (K_D = 200-300 μM). The analysis of thermodynamic parameters (Table 1) indicates that the interactions are enthalpy-driven (with an enthalpy value ΔH ~ 29 kJ/mol for BP39L and ΔH ~ 36 kJ/mol for CV39L) with unfavourable entropy contributions, which are typical for protein-saccharide interactions [56]. Moreover, BP39L and CV39L bind D-mannose (for CV39L also D-maltose) with an affinity weaker than 1 mM (data not shown).

3.3. Structural Description

BP39L crystallizes in the space group *I*222 and CV39L in the space group $P2_12_12_1$ with 1 and 4 protein molecules in the asymmetric unit, respectively (data collection and refinement statistics are in Table 2). Both BP39L and CV39L exhibit a single-domain structure with the seven-bladed

β-propeller fold organized around a sevenfold pseudoaxis of symmetry (Fig. 3A and 3C). Each repeat for BP39L and CV39L (also called the W-motif; see Fig. 3A and 3C) contains a twisted four-stranded antiparallel β-sheet (A-B-C-D; Fig. 3A and 3C) connected by relatively short loops. In the structure of CV39L, a short α-helix is present in each repeat connecting β-strands A and B (Fig. 3C). The observed circular dichroism (CD) spectra of BP39L and CV39L lectins are in qualitative agreement with the theoretical curve for the β-turns of type II and/or I (Fig. S4), suggesting a dominance of β-turns in BP39L and CV39L structures in solution.

The overall shape for both lectins is a small torus, with a diameter of ~45 Å and a height of ~30 Å for BP39L, and ~50 Å and ~30 Å for CV39L. In the β-propeller fold, the first β-strand in each blade (A) defines the tunnel in the centre of the protein and the last one (D) is the most exposed to solvent at the protein surface (Fig. 3). The β-strands in both lectins are mostly hydrophobic except for β-strand A, which exhibits an amphiphilic character. The inner tunnel broadens from one side of monomer to the other one from 13 to 20 Å for BP39L and from 17 to 25 Å for CV39L. The BP39L N-terminus is located in the central cavity, stabilized mainly by hydrophobic and polar interactions (hydrogen bonds) mediated by water molecules (Fig. 3A). The N-terminus of CV39L interacts with the C-terminus of the lectin and is also mainly stabilized by hydrophobic and polar interactions mediated by β-strand D of the W7 repeat (Fig. 3C). Superposition of the seven β-strand blade repeats gives overall RMSD (C_α-backbone) values of 0.36 Å and 0.43 Å for BP39L and CV39L, respectively, indicating high structural similarity of the repeats for each lectin.

DaliLite [57] was used to search for proteins with structural similarity. BP39L and CV39L exhibited the highest structural identity to PHL, a 7-bladed β-propeller lectin from *Photorhabdus asymbiotica* (with RMSD 2.1 to 2.2 Å, PDB 5MXE) [29] and the related PLL lectin form *P.*

luminescens (with RMSD 2.0 to 2.3 Å, PDB 5C9O) [30] with Z-scores in a range from 32.8 to 36.7 (Table S5). These proteins have sequence identities with BP39L and CV39L ranging from 18.9% to 21.4%. 7-bladed lectins from PVL family (PVL from *Psathyrella velutina* [58] and AAL2 from *Agrocybe aegerita* [59]) did not show strong structural similarity to CV39L and BP39L with Dali Z-scores in a range from 12.6 to 19.5 corresponding to relatively low sequence identities in a range from 11.1% to 20%.

3.4. Oligomeric State

In the crystalline state, BP39L is a monomer with no extensive contact with any neighbour, while CV39L is a dimer with an interface of 1420 $Å^2$, as analysed by the software PISA (CCP4 package) [60]. The two monomers of CV39L (Fig. S5) are related by a pseudo-2-fold axis of symmetry and superposition of their C_α-backbones gives an RMSD value of 0.23 Å. The polar interactions mediated by loops play a key role in dimer association. The main hydrogen bonds in CV39L dimerization are established between the side chains of Thr39 (W1/B-C) - Glu91 (W2/B-C) and Arg140 (W3/B-C), Ser328 (W7/A-B) – Ser328 (W7/A-B), and Gln350 (W7/B-C) - Thr276 (W6/A-B) and Ile295 (W6/B-C). This dimerization mode creates a “back to back” association of two CV39L molecules (Fig. S5).

The oligomerization state of both lectins in solution was investigated by analytical ultracentrifugation (AUC) (Fig. 4). Analytical ultracentrifugation experiments revealed that in solution both BP39L and CV39L behave as single homogeneous species. BP39L sedimented at 3.09 S ($s_{20,w}$ = 3.20 after extrapolation to standard conditions in Sednterp), suggesting the presence of a monomeric species with a frictional ratio of 1.26 (Fig. 4A). The sedimentation coefficient of CV39L was 5.21 S ($s_{20,w}$ = 5.48). This value is clearly much higher than the

predicted maximum value for a spherical monomer (3.86 S as calculated in Sednterp), suggesting that a dimer with a moderately elongated shape (frictional ratio of 1.27) is formed (Fig. 4B). Global analysis of the sedimentation equilibrium gave a molecular weight of 37.0 kDa for BP39L and 75.8 kDa for CV39L (Fig. 4). These values correspond well with the theoretical molecular weight of BP39L for a monomer (37.7 kDa) and CV39L for a dimer (79.3 kDa). The AUC analysis confirmed that the protein oligomeric states observed in the crystal structure for both lectins are maintained in solution.

3.5. Putative Binding Sites

Both co-crystallization and soaking experiments with BP39L (D-mannose, D-galactose, Man-α1,3(Manα1,6)Man) and CV39L (D-mannose, D-galactose, L-fucose) did not result in formation of complex with saccharides. The analysis of the surface revealed seven possible binding sites in the form of shallow pockets in between blades (Fig. 3B, 3D, 5A and 5B), similar to observations of other seven-bladed β-propeller lectins such as PLL from *P. luminescens* [30] and PHL from *P. asymbiotica* [29]. Comparison of the two lectins' putative binding sites indicated the presence of conserved tryptophan residues (possibly enabling the hydrophobic interaction) and polar amino acids (responsible for the stabilization of the saccharide via hydrogen bonds) (Fig. 5), as classically observed in the propeller carbohydrate binding sites ([23, 24, 30, 29]). For both lectins, the binding sites are created by amino acids of blade Wi and blade Wi+1 (Fig. 5). The potential seven binding sites in BP39L lectin are highly conserved (Fig. 5A and 5D). In CV39L lectin, the seven binding sites are highly similar to each other and their overall architecture resembles the sites of BP39L by position of stacking Trp, and most polar and hydrophobic residues. (Fig. 5B and 5D). Most of the amino acids that can participate in

saccharide binding are located in the four β-strands, except for the Asn and Tyr residues of CV39L (Fig. 5B) that contribute to the formation of the short α-helix.

The related *B. cenocepacia* (species of *Bcc* as *B. pseudomallei*) and *C. violaceum* contain C-type lectins that require a calcium ion [61, 62, 17]. Two calcium ions help to further strengthen the binding between the protein and a carbohydrate through hydrogen bonds. The potential ion-dependent carbohydrate-binding of BP39L and CV39L lectins was tested using thermal denaturation CD spectroscopy. The presence of D-mannose (BP39L) and L-fucose (CV39L) alone in solution increased the thermal stability of the lectins by ~ 4°C, whereas the presence of cations (Ca^{2+}, Mg^{2+}, Mn^{2+}) had no effect on the stability of both lectins (Fig. S6).

As the crystallization with saccharides failed to provide protein-saccharide complex structures, a molecular modelling approach was performed to dock monosaccharides into the putative binding site of both lectins (Fig. 6 and Table S4). The O1-methylated saccharides were used in docking studies (methyl α/β-D-mannoside and Manα1-3(Manα1-6)Man mannotriose for BP39L, methyl α/β-L-fucoside, methyl α/β-D-glucoside, methyl β-D-mannoside, methyl β-L-galactoside and methyl α/β-D-galactoside for CV39L). For simplicity and due to their very similar behaviour, only the data considering the binding site between repeats W1 and W2 for both lectins are described (Fig. 6 and Table S4).

Docking of the two anomeric forms of D-mannoside resulted in very different orientation of the saccharide ring in the BP39L binding site (Fig. 6A, 6B and Table S4). Analysis of the binding mode of docked α-D-mannoside (Fig. 6A) shows stabilization via polar interactions (at O2, O3, O5 and O6 hydroxyl positions) by Q22, D101 and W103 residues of the BP39L binding pocket. Moreover, the orientation of the saccharide revealed that the W44 residue should be responsible for the stacking interaction necessary for saccharide stabilization (Fig. 6A). The O1-methyl

group is exposed to the solvent side of the binding pocket, enabling the possible binding of mannosylated oligosaccharides as observed in the glycan microarray. Docking of Manα1-3(Manα1-6)Man mannotriose revealed several possible orientations with one of the terminal mannoses being coordinated in binding site keeping the same binding mode as observed for α-D-mannoside (Fig 6C, 6D). Additional mannose units extend towards the loop between β-sheets C and D, where they are further stabilized by polar interactions to W44, A50 and N52 residues. In either of the binding mode, the reducing O1 end of oligosaccharide is exposed to the solvent enabling possible extension by more complex oligosaccharide. In contrast, β-D-mannoside was not properly accommodated by the binding site (Fig. 6B). The most likely orientation from docking correspond to distorted saccharide ring with O1-methyl group pointing into the binding pocket. This would be not compatible with oligosaccharide binding what corresponds to results from glycan array.

For CV39L, both anomeric forms of L-fucoside were accommodated by the binding site, however clear differences in the binding mode were observed (Fig. 6E and 6F). For β-L-fucoside (Fig. 6F), the saccharide ring orientation allows the O1-methyl group of β-L-fucoside to be exposed to solvent and hence terminal β-L-fucoside is well accessible. Considering the stacking and polar interactions which might occur in the stabilization of β-L-fucoside, the N26, Q31, Q101, G105 and W107 residues were found to possibly facilitate the polar interactions and the W47 residue to provide the stacking interaction (Fig. 6F). The most probable orientation of α-L-fucoside is flipped bottom-up compared to β-L-fucoside (Fig. 6E). It would also enable extension at O1 as the O1-methyl group is exposed to solvent, however the predicted net of hydrogen bonds is less defined and the O4 oxygen is oriented towards W47 side chain resulting in repulsion. This is in good agreement with glycan array data, where β-L-fucoside was among

best binders, while no significant recognition of α-L-fucoside was observed. Docking of O1-methyl glucosides show formation of more complex polar contacts for α anomer (Fig. 6G, 6H). O1-methyl group is facing the solvent in both cases, however saccharide molecules are rotated by 180 degrees with respect to each other. Preference for α-D-glucoside observed in glycan array analysis is therefore likely to be caused by additional contacts rather than by position of glucose moiety itself. Docking of b-D-mannoside resulted in formation of hydrogen bonds between O3, O4 and O6 sugar atoms and protein binding site, however O2 hydroxyl points towards apolar part of W47 side chain resulting in possible repulsion. We also tried docking of Manβ1-4GlcNAc (most preferred ligand in glycan array), however the results did not converge (data not shown) and therefore the binding preference cannot be clearly explained in this case. Docking of both α and β D-galactoside resulted in similar saccharide orientation with O1-methyl group pointing towards solvent (Fig. 6K, 6L). However, in both cases the orientation of O4 hydroxyl group might cause sterical clash with W47 side chain. Additionally, since L-fucose and L-galactose are similar, with respect to chemical composition and geometrical conformation, the molecular docking of β-L-galactoside revealed that they share a binding mode in the CV39L binding site (Fig. 6F and 6J) with the O1 hydroxyl oriented towards the solvent side of the binding site of CV39L.

4. Discussion

4.1. Comparison of putative binding sites between BP39L and CV39L

B. pseudomallei and *C. violaceum* are pathogens known to cause infections in humans, the mechanism of which is not well understood. However, bacterial lectins produced by these pathogens are believed to play a key role in their pathogenicity, host recognition and adhesion.

BP39L and CV39L lectins were identified in *B. pseudomallei* and *C. violaceum*, sharing 26% identity in their amino acid sequence. The structures of BP39L and CV39L lectins revealed a 7-bladed β-propeller fold with up to seven putative binding sites, indicating the multivalent potential of both lectins. The multivalent interactions of lectins are intensively studied, since they can be targeted by glycomimetics to compete with host recognition [63].

The structural superimposition of the BP39L and CV39L lectins and their putative binding sites revealed $RMSD_{(C\alpha)}$ values of 3.6 Å and 2.3 Å respectively, indicating a close structural relationship. A closer comparison of their putative binding sites showed that the hydrophobic region ($F32_{BP39L}/L33_{CV39L}$, $V42_{BP39L}/L44_{CV39L}$, $W44_{BP39L}/W47_{CV39L}$) is well preserved and the set of two Trp residues is highly conserved ($W44_{BP39L}/W47_{CV39L}$, $W103_{BP39L}/W107_{CV39L}$) (Fig. 7). The W44 residue in the BP39L putative site is in a flipped position compared to W47 in the CV39L site (Fig. 5C). The inner core of the binding cavities is composed of polar residues ($D30_{BP39L}/Q31_{CV39L}$ and $Q22_{BP39L}/N26_{CV39L}$) which may contribute to the saccharide stabilization via hydrogen bonds. The largest difference was found in the outer part of the putative binding sites, where a change was observed from hydrophilic to hydrophobic character or *vice versa* (loop region – $Q28_{BP39L}/Y29_{CV39L}$ and $V97_{BP39L}/Q101_{CV39L}$).

The indicated change in the chemical character of BP39L and CV39L binding sites (outer part) relates to the contrasting saccharide specificity of the lectins obtained by glycan microarray and ITC. The experiments conducted on a series of oligo- and polysaccharides confirmed the preference of BP39L for mannosylated saccharides with Manα1-6Man being the preferred motif over Manα1-3Man and Manα1-2Man isomers (Table S1). CV39L recognize mainly rather more complex bacterial exopolysaccharides with an interesting monosaccharide preference for β-L-fucose. This could be related to the versatility and high environmental adaptability of *C.*

violaceum [12], as β-L-fucose cannot be found in mammals, while it was detected in amphibians [64]. In case of both lectins, they do not exhibit strict specificity towards one type of saccharide. As the binding sites are highly conserved for each of the studied lectins, this can be explained by binding site promiscuity, which was already reported for several other lectins, such as CV-IIL from *Chromobacterium violaceum* [17]. We also cannot exclude the existence of second type of binding site with alternated specificity, as was reported for lectin PHL [29], however there are no clear structural evidences for this.

The ability of BP39L and CV39L lectins to bind oligo- and polysaccharides correlates with the wide shallow pocket shape of the putative binding sites. This fact coincides very well with the stronger binding constant of Manα1-6(Manα1-3)Man trisaccharide (K_D = ~ 0.2 mM) to BP39L compared to D-mannose. Additionally, the docking simulations of D-mannose (BP39L) and L-fucose (CV39L) revealed expected CH-π interaction between the apolar part of the saccharide with Trp residue and coordination of OH groups by a net of hydrogen bonds. For all studied saccharides (with exception of O1-methyl β- D-mannoside in BP39L binding site), we observed possible O1 hydroxyl orientation of the saccharides, which could be exposed to the solvent side of the putative binding sites to allow the binding of the oligo- and polysaccharides.

4.2. Comparison of seven-bladed β-propeller lectins

Currently, only four lectins other than BP39L and CV39L are known to carry a 7-bladed β-propeller fold. The 7-bladed β-propellers of PVL from *Psathyrella velutina* [21] and AAL2 from *Agrocybe aegerita* [22] are fungal lectins belonging to the same lectin family specifically recognizing GlcNAc. The PVL and AAL2 lectins did not show strong structural similarity to CV39L and BP39L (Table S5). Dali Z-scores are in a range from 12.6 to 19.5 corresponding to

relatively low sequence identities in a range from 11.1% to 20%. These lectins have a different formation of the blades in the seven-bladed β-propeller fold compared to CV39L and BP39L. Additionally, the localization of their binding sites between the respective blades is on the opposite edge of the propeller (Fig. 7A and 7D).

The other two lectins, PHL from *P. asymbiotica* [29] and PLL from *P. luminescens* [30], exhibited the highest structural similarity to CV39L and BP39L with Z-scores in a range from 32.8 to 36.7 (Table S5). These proteins have sequence identities with BP39L and CV39L ranging from 18.9% to 21.4%. PHL is the first known bangle lectin to possess dual specificity with well-defined binding sites for two unrelated saccharides: L-fucose and D-galactose. The fucose-binding PLL lectin is sequentially and structurally related to PHL. Both lectins share the characteristic hydrophobic binding mode of two Trp residues to stabilize L-fucose. However, the fucose-binding sites of PHL and PLL are located on the opposite site of the blades to BP39L and CV39L (Fig. 7B and 7E). The diverse employment of two Trp residues in the putative BP39L and CV39L binding sites compared to the fucose-binding sites of PHL and PLL lectins [29, 30] shows that they are unrelated.

Interestingly, the architecture of the D-galactose-binding site of PHL lectin resembles putative binding sites of BP39L and CV39L lectins (Fig. 7C and 7F), mainly considering the highly conserved set of two Trp residues. The main signature of these two Trp residues in the D-galactose-binding site of PHL is that the first Trp residue stabilizes the hydrophobic part of the saccharide and the second Trp residue creates a network of polar interactions with other amino acids in the binding pocket to stabilize the saccharide. Molecular docking simulations in putative binding sites of the BP39L and CV39L lectins revealed a potentially very similar binding mode to the galactose-binding site of PHL that commonly also occurs in six-bladed β-propeller lectins

such as AAL [23] or AFL [24]. Although BP39L and CV39L lectins possess a very similar structural fold and probable binding scheme compared to PHL and its D-galactose-binding site, they are sequentially very distinct, making the sequence alignment impossible, even at the level of single blades and the respective binding sites. The conserved sequence motifs of the PHL and PLL lectins show the presence of two conserved sets of two Trp residues which are located in the L-fucose-binding sites (PHL and PLL) and in the D-galactose-binding sites (PHL). In contrast, the conserved sequence motifs of BP39L and CV39L contain only one set of two highly conserved Trp residues located in putative binding sites.

Despite having the same three-dimensional fold, the BP39L and CV39L lectins belong to a different family than the PHL and PLL lectins, mainly due to the absence of significant sequence similarity and a different amount of conserved sets of Trp residues, which define the binding sites. A similar situation can be found in the field of glycosyltransferases, where to date only four 3-D folds are commonly known, but 97 distinct families have been identified [65].

5. Conclusions

This work indicates that BP39L and CV39L should belong to an undescribed new family of lectins. Furthermore, the ability of BP39L to bind mannosylated polysaccharides and CV39L to bind bacterial exo-polysaccharides may help to design additional functional and structural studies to synthesize potential glyco-clusters which could be used in the development of anti-infective drugs.

Acknowledgments

This work was supported by the Czech Science Foundation (18-18964S) and by the Czech Ministry of Education, Youth and Sports within programme INTER COST (LTC17076) and rom European Regional Development Fund-Project „CIISB4HEALTH“ (No. CZ.02.1.01/0.0/0.0/16_013/0001776). Work of N.B. and N.S. was supported by the Russian Science Foundation (project No. 14-50-00131). We wish to thank Synchrotron PETRA III – P13 (Hamburg, Germany) for access to their synchrotron data-collection facilities and the allocation of synchrotron radiation beamtime. A.V. and A.I. acknowledge the support of French ANR through Glyco@Alps (ANR-15-IDEX-02) and Labex ARCANE (ANR-11-LABX-003). We wish to acknowledge the Biomolecular Interactions and Crystallization and the Nanobiotechnology Core Facilities of CEITEC supported by the CIISB research infrastructure (LM2015043 funded by MEYS CR) for their support with obtaining scientific data presented in this paper.

Conflict of interest

The authors declare that they have no conflicts of interest with the contents of this article.

Author Contributions

P.S., J.N., G.D., M.P. and M.W. conceived and designed the experiments. P.S., J.N., G.D., E.D., J.K., E.F., L.H., N.S., and M.P. performed the experiments. P.S., J.N., G.D., J.H., E.D., J.K., E.F., A.V., A.I., N.S., N.B., M.P. and M.W. analyzed the data. G.D. solved the crystal structures. P.S., G.D., J.H. and M.W. wrote the manuscript.

References

[1] M.J. Medina-Pascual, S. Valdezate, P. Villalón, N. Garrido, V. Rubio, J.A. Saéz-Nieto, Identification, molecular characterisation and antimicrobial susceptibility of genomovars of the *Burkholderia cepacia* complex in Spain, Eur. J. Clin. Microbiol. Infect. Dis. 31 (2012) 3385–3396.

[2] C. Peeters, J.E.A. Zlosnik, T. Spilker, T.J. Hird, J.J. LiPuma, P. Vandamme, *Burkholderia pseudomultivorans* sp. nov., a novel *Burkholderia cepacia* complex species from human respiratory samples and the rhizosphere, Syst. Appl. Microbiol. 36 (2013) 483–489.

[3] T. Spilker, J.L. Ginther, B.J. Currie, J.E.A. Zlosnik, B. De Smet, J.J. LiPuma, P. Keim, T.J. Hird, D.M. Wagner, J.A. Jacobs, M. Kaestli, C. Peeters, T.J. Kidd, S.C. Bell, M. Mayo, P. Vandamme, *Burkholderia stagnalis* sp. nov. and *Burkholderia territorii* sp. nov., two novel *Burkholderia cepacia* complex species from environmental and human sources, Int. J. Syst. Evol. Microbiol. 65 (2015) 2265–2271.

[4] S.A. Sousa, C.G. Ramos, J.H. Leitão, *Burkholderia cepacia* Complex: Emerging Multihost Pathogens Equipped with a Wide Range of Virulence Factors and Determinants, Int. J. Microbiol. 2011 (2011) 1–9.

[5] W.J. Wiersinga, B.J. Currie, S.J. Peacock, Melioidosis, N. Engl. J. Med. 367 (2012) 1035–1044.

[6] N.J. White, Melioidosis, Lancet Lond. Engl. 361 (2003) 1715–1722.

[7] L.-C. Choh, G.-H. Ong, K.M. Vellasamy, K. Kalaiselvam, W.-T. Kang, A.R. Al-Maleki, V. Mariappan, J. Vadivelu, *Burkholderia* vaccines: are we moving forward?, Front. Cell. Infect. Microbiol. 3 (2013)..

[8] E.E. Galyov, P.J. Brett, D. DeShazer, Molecular Insights into *Burkholderia pseudomallei* and *Burkholderia mallei* Pathogenesis, Annu. Rev. Microbiol. 64 (2010) 495–517.

[9] J. Gilad, I. Harary, T. Dushnitsky, D. Schwartz, Y. Amsalem, *Burkholderia mallei* and *Burkholderia pseudomallei* as bioterrorism agents: national aspects of emergency preparedness, Isr. Med. Assoc. J. IMAJ. 9 (2007) 499–503.

[10] G.C. Whitlock, D. Mark Estes, A.G. Torres, Glanders: off to the races with *Burkholderia mallei*, FEMS Microbiol. Lett. 277 (2007) 115–122..

[11] C.-H. Yang, Y.-H. Li, *Chromobacterium violaceum* infection: A clinical review of an important but neglected infection, J. Chin. Med. Assoc. 74 (2011) 435–441.

[12] Brazilian National Genome Project Consortium., The complete genome sequence of *Chromobacterium violaceum* reveals remarkable and exploitable bacterial adaptability, Proc. Natl. Acad. Sci. 100 (2003) 11660–11665.

[13] M. Salanoubat, S. Genin, F. Artiguenave, J. Gouzy, S. Mangenot, M. Arlat, A. Billault, P. Brottier, J.C. Camus, L. Cattolico, M. Chandler, N. Choisne, C. Claudel-Renard, S. Cunnac, N. Demange, C. Gaspin, M. Lavie, A. Moisan, C. Robert, W. Saurin, T. Schiex, P. Siguier, P. Thébault, M. Whalen, P. Wincker, M. Levy, J. Weissenbach, C.A. Boucher, Genome sequence of the plant pathogen *Ralstonia solanacearum*, Nature. 415 (2002) 497–502.

[14] C. Stephens, Microbial genomics: tropical treasure?, Curr. Biol. CB. 14 (2004) R65-66.

[15] H. Lis, N. Sharon, Lectins: Carbohydrate-Specific Proteins That Mediate Cellular Recognition, Chem. Rev. 98 (1998) 637–674.

[16] K. Zinger-Yosovich, Production and properties of the native *Chromobacterium violaceum* fucose-binding lectin (CV-IIL) compared to homologous lectins of *Pseudomonas*

aeruginosa (PA-IIL) and *Ralstonia solanacearum* (RS-IIL), Microbiology. 152 (2006) 457–463.

[17] M. Pokorná, G. Cioci, S. Perret, E. Rebuffet, N. Kostlánová, J. Adam, N. Gilboa-Garber, E.P. Mitchell, A. Imberty, M. Wimmerová, Unusual Entropy-Driven Affinity of *Chromobacterium violaceum* Lectin CV-IIL toward Fucose and Mannose † , ‡, Biochemistry. 45 (2006) 7501–7510.

[18] A. Imberty, M. Wimmerová, E.P. Mitchell, N. Gilboa-Garber, Structures of the lectins from *Pseudomonas aeruginosa*: insight into the molecular basis for host glycan recognition, Microbes Infect. 6 (2004) 221–228.

[19] D. Sudakevitz, N. Kostlánová, G. Blatman-Jan, E.P. Mitchell, B. Lerrer, M. Wimmerová, D.J. Katcoff, A. Imberty, N. Gilboa-Garber, A new *Ralstonia solanacearum* high-affinity mannose-binding lectin RS-IIL structurally resembling the *Pseudomonas aeruginosa* fucose-specific lectin PA-IIL: *Ralstonia solanacearum* lectin RS-IIL, Mol. Microbiol. 52 (2004) 691–700.

[20] F. Bonnardel, A. Kumar, M. Wimmerova, M. Lahmann, S. Perez, A. Varrot, F. Lisacek, A. Imberty, Architecture and Evolution of Blade Assembly in β-propeller Lectins, Structure. 27 (2019) 764-775.e3.

[21] G. Cioci, E.P. Mitchell, V. Chazalet, H. Debray, S. Oscarson, M. Lahmann, C. Gautier, C. Breton, S. Perez, A. Imberty, β-Propeller Crystal Structure of *Psathyrella velutina* Lectin: An Integrin-like Fungal Protein Interacting with Monosaccharides and Calcium, J. Mol. Biol. 357 (2006) 1575–1591.

[22] X.-M. Ren, D.-F. Li, S. Jiang, X.-Q. Lan, Y. Hu, H. Sun, D.-C. Wang, Structural Basis of Specific Recognition of Non-Reducing Terminal N-Acetylglucosamine by an *Agrocybe aegerita* Lectin, PLOS ONE. 10 (2015) e0129608.

[23] M. Wimmerova, E. Mitchell, J.-F. Sanchez, C. Gautier, A. Imberty, Crystal structure of fungal lectin: six-bladed beta-propeller fold and novel fucose recognition mode for *Aleuria aurantia* lectin, J. Biol. Chem. 278 (2003) 27059–27067.

[24] J. Houser, J. Komarek, N. Kostlanova, G. Cioci, A. Varrot, S.C. Kerr, M. Lahmann, V. Balloy, J.V. Fahy, M. Chignard, et al., A soluble fucose-specific lectin from *Aspergillus fumigatus* conidia--structure, specificity and possible role in fungal pathogenicity, PloS One. 8 (2013) e83077.

[25] S.C. Kerr, G.J. Fischer, M. Sinha, O. McCabe, J.M. Palmer, T. Choera, F. Yun Lim, M. Wimmerova, S.D. Carrington, S. Yuan, C.A. Lowell, S. Oscarson, N.P. Keller, J.V. Fahy, FleA Expression in Aspergillus fumigatus Is Recognized by Fucosylated Structures on Mucins and Macrophages to Prevent Lung Infection, PLOS Pathog. 12 (2016) e1005555.

[26] N. Richard, L. Marti, A. Varrot, L. Guillot, J. Guitard, C. Hennequin, A. Imberty, H. Corvol, M. Chignard, V. Balloy, Human Bronchial Epithelial Cells Inhibit *Aspergillus fumigatus* Germination of Extracellular Conidia via FleA Recognition, Sci. Rep. 8 (2018).

[27] N. Kostlánová, E.P. Mitchell, H. Lortat-Jacob, S. Oscarson, M. Lahmann, N. Gilboa-Garber, G. Chambat, M. Wimmerová, A. Imberty, The Fucose-binding Lectin from *Ralstonia solanacearum*: A New Type of β-propeller Architecture Formed by Oligomerization and Interacting with fucoside, fucosyllactose, and plant xyloglucan, J. Biol. Chem. 280 (2005) 27839–27849.

[28] A. Audfray, J. Claudinon, S. Abounit, N. Ruvoën-Clouet, G. Larson, D.F. Smith, M. Wimmerová, J. Le Pendu, W. Römer, A. Varrot, A. Imberty, Fucose-binding Lectin from Opportunistic Pathogen *Burkholderia ambifaria* Binds to Both Plant and Human Oligosaccharidic Epitopes, J. Biol. Chem. 287 (2012) 4335–4347.

[29] G. Jančaříková, J. Houser, P. Dobeš, G. Demo, P. Hyršl, M. Wimmerová, Characterization of novel bangle lectin from *Photorhabdus asymbiotica* with dual sugar-binding specificity and its effect on host immunity, PLOS Pathog. 13 (2017) e1006564.

[30] A. Kumar, P. Sýkorová, G. Demo, P. Dobeš, P. Hyršl, M. Wimmerová, A Novel Fucose-binding Lectin from *Photorhabdus luminescens* (PLL) with an Unusual Heptabladed β-Propeller Tetrameric Structure, J. Biol. Chem. 291 (2016) 25032–25049.

[31] D. Goyard, V. Baldoneschi, A. Varrot, M. Fiore, A. Imberty, B. Richichi, O. Renaudet, C. Nativi, Multivalent Glycomimetics with Affinity and Selectivity toward Fucose-Binding Receptors from Emerging Pathogens, Bioconjug. Chem. 29 (2018) 83–88..

[32] G. Jančaříková, M. Herczeg, E. Fujdiarová, J. Houser, K.E. Kövér, A. Borbás, M. Wimmerová, M. Csávás, Synthesis of α- L -Fucopyranoside-Presenting Glycoclusters and Investigation of Their Interaction with *Photorhabdus asymbiotica* Lectin (PHL), Chem. - Eur. J. 24 (2018) 4055–4068.

[33] T. Machida, A. Novoa, É. Gillon, S. Zheng, J. Claudinon, T. Eierhoff, A. Imberty, W. Römer, N. Winssinger, Dynamic Cooperative Glycan Assembly Blocks the Binding of Bacterial Lectins to Epithelial Cells, Angew. Chem. Int. Ed. 56 (2017) 6762–6766.

[34] Y.A. Knirel, H.-J. Gabius, O. Blixt, E.M. Rapoport, N.R. Khasbiullina, N.V. Shilova, N.V. Bovin, Human tandem-repeat-type galectins bind bacterial non-βGal polysaccharides, Glycoconj. J. 31 (2014) 7–12..

[35] E. Girard, M. Stelter, J. Vicat, R. Kahn, A new class of lanthanide complexes to obtain high-phasing-power heavy-atom derivatives for macromolecular crystallography, Acta Crystallogr. D Biol. Crystallogr. 59 (2003) 1914–1922.

[36] W. Kabsch, *XDS*, Acta Crystallogr. D Biol. Crystallogr. 66 (2010) 125–132.

[37] M.D. Winn, C.C. Ballard, K.D. Cowtan, E.J. Dodson, P. Emsley, P.R. Evans, R.M. Keegan, E.B. Krissinel, A.G.W. Leslie, A. McCoy, S.J. McNicholas, G.N. Murshudov, N.S. Pannu, E.A. Potterton, H.R. Powell, R.J. Read, A. Vagin, K.S. Wilson, Overview of the *CCP* 4 suite and current developments, Acta Crystallogr. D Biol. Crystallogr. 67 (2011) 235–242.

[38] T. Pape, T.R. Schneider, *HKL2MAP* : a graphical user interface for macromolecular phasing with *SHELX* programs, J. Appl. Crystallogr. 37 (2004) 843–844.

[39] G. Langer, S.X. Cohen, V.S. Lamzin, A. Perrakis, Automated macromolecular model building for X-ray crystallography using ARP/wARP version 7, Nat. Protoc. 3 (2008) 1171–1179.

[40] S. Panjikar, V. Parthasarathy, V.S. Lamzin, M.S. Weiss, P.A. Tucker, On the combination of molecular replacement and single-wavelength anomalous diffraction phasing for automated structure determination, Acta Crystallogr. D Biol. Crystallogr. 65 (2009) 1089–1097.

[41] A. Vagin, A. Teplyakov, Molecular replacement with *MOLREP*, Acta Crystallogr. D Biol. Crystallogr. 66 (2010) 22–25.

[42] A.A. Vagin, R.A. Steiner, A.A. Lebedev, L. Potterton, S. McNicholas, F. Long, G.N. Murshudov, *REFMAC* 5 dictionary: organization of prior chemical knowledge and guidelines for its use, Acta Crystallogr. D Biol. Crystallogr. 60 (2004) 2184–2195.

[43] P. Emsley, B. Lohkamp, W.G. Scott, K. Cowtan, Features and development of *Coot*, Acta Crystallogr. D Biol. Crystallogr. 66 (2010) 486–501.

[44] R.A. Laskowski, M.W. MacArthur, D.S. Moss, J.M. Thornton, PROCHECK: a program to check the stereochemical quality of protein structures, J. Appl. Crystallogr. 26 (1993) 283–291.

[45] V.B. Chen, W.B. Arendall, J.J. Headd, D.A. Keedy, R.M. Immormino, G.J. Kapral, L.W. Murray, J.S. Richardson, D.C. Richardson, *MolProbity* : all-atom structure validation for macromolecular crystallography, Acta Crystallogr. D Biol. Crystallogr. 66 (2010) 12–21.

[46] P. Schuck, Size-Distribution Analysis of Macromolecules by Sedimentation Velocity Ultracentrifugation and Lamm Equation Modeling, Biophys. J. 78 (2000) 1606–1619.

[47] P. Schuck, On the analysis of protein self-association by sedimentation velocity analytical ultracentrifugation, Anal. Biochem. 320 (2003) 104–124.

[48] R. Abagyan, M. Totrov, D. Kuznetsov, ICM-A new method for protein modeling and design: applications to. J Comp Chem 15 (1994) 488–506]

[49] R. Abagyan, ICM user manual. https://www.molsoft.com/, 2017 (accessed 1 July 2019)

[50] G. Némethy, K.D. Gibson, K.A. Palmer, C.N. Yoon, G. Paterlini, A. Zagari, S. Rumsey, H.A. Scheraga, Energy parameters in polypeptides. X. Improved geometrical parameters and nonbonded interactions for use in the ECEPP/3 algorithm, with application to proline-containing peptides J Phys Chem 96 (1992) 6472-6484

[51] T.A. Halgren, Merck molecular force field. I. Basis, form, scope, parameterization, and performance of MMFF94. J Comp Chem 17 (1996) 490–519

[52] M. Totrov, R. Abagyan, Flexible protein-ligand docking by global energy optimization in internal coordinates. Proteins 29(Suppl 1) (1997) 215-220,

[53] R. Abagyan, M. Totrov, Biased probability Monte Carlo conformational searches and electrostatic calculations for peptides and proteins. J Mol Biol, 235 (1994) 983–1002

[54] S. Pérez, A. Sarkar, A. Rivet, C. Breton & A. Imberty, Glyco3D : A portal for structural glycosciences. Meth. Mol. Biol. 1273 (2015) 241-258.

[55] Woods Group. (2005-2014) GLYCAM Web. Complex Carbohydrate Research Center, University of Georgia, Athens, GA. http://www.glycam.com (accessed 1 July 2019).

[56] T.K. Dam, C.F. Brewer, Probing Lectin–Mucin Interactions by Isothermal Titration Microcalorimetry, in: S.R. Stowell, R.D. Cummings (Eds.), Galectins, Springer New York, New York, NY, 2015, pp. 75–90.

[57] L. Holm, P. Rosenström, Dali server: conservation mapping in 3D, Nucleic Acids Res. 38 (2010) W545-549.

[58] G. Cioci, E.P. Mitchell, V. Chazalet, H. Debray, S. Oscarson, M. Lahmann, C. Gautier, C. Breton, S. Perez, A. Imberty, β-Propeller Crystal Structure of *Psathyrella velutina* Lectin: An Integrin-like Fungal Protein Interacting with Monosaccharides and Calcium, J. Mol. Biol. 357 (2006) 1575–1591.

[59] G. Yang, A.J. Dowling, U. Gerike, R.H. ffrench-Constant, N.R. Waterfield, Photorhabdus virulence cassettes confer injectable insecticidal activity against the wax moth, J. Bacteriol. 188 (2006) 2254–2261.

[60] E. Krissinel, K. Henrick, Inference of macromolecular assemblies from crystalline state, J. Mol. Biol. 372 (2007) 774–797.

[61] O. Šulák, G. Cioci, E. Lameignère, V. Balloy, A. Round, I. Gutsche, L. Malinovská, M. Chignard, P. Kosma, D.F. Aubert, C.L. Marolda, M.A. Valvano, M. Wimmerová, A.

Imberty, *Burkholderia cenocepacia* BC2L-C Is a Super Lectin with Dual Specificity and Proinflammatory Activity, PLoS Pathog. 7 (2011) e1002238.

[62] E. Lameignere, L. Malinovská, M. Sláviková, E. Duchaud, E.P. Mitchell, A. Varrot, O. Šedo, A. Imberty, M. Wimmerová, Structural basis for mannose recognition by a lectin from opportunistic bacteria *Burkholderia cenocepacia*, Biochem. J. 411 (2008) 307–318..

[63] S. Cecioni, A. Imberty, S. Vidal, Glycomimetics versus Multivalent Glycoconjugates for the Design of High Affinity Lectin Ligands, Chem. Rev. 115 (2015) 525–561.

[64] T.N. Druzhinina, N.S. Utkina, K. Chan, G. Strecker, V.N. Shibaev, Activity of enzymes catalyzing formation of beta-L-fucosyl phosphate and GDP-beta-L-fucose in amphibian tissues and their application in chemo-enzymic synthesis of GDP-beta-L-fucose, Biochem. Biokhimiia. 64 (1999) 783–787.

[65] A. Varki (Ed.), Essentials of glycobiology, third ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, 2017.

Captions to illustrations

Fig. 1. Box and whisker plot for glycan array screening for BP39L and CV39L labelled with DyLight 488 NHS Ester. (A) BP39L and (B) CV39L - the top 20 saccharides were selected for display. The 10 saccharides that gave the highest average signals are depicted. Linker formula -sp2: -O-$(CH_2)_2NH_2$; sp3: -O-$(CH_2)_3NH_2$; sp4: -$NHCOCH_2NH_2$. The bottom and top of the box are the first and third quartiles; the band inside the box is the second quartile (the median); the ends of the whiskers represent the minimum and maximum values of the data and the small squares inside the boxes represent the mean relative fluorescence units (RFU). The raw glycan array results are given in Tables S2 and S3.

Fig. 2. ITC titration curves of BP39L and CV39L with different saccharides. (A) BP39L (0.1 mM) titrated with Manα1-6(Manα1-3)Man (10 mM). (B) CV39L (0.1 mM) titrated with L-galactose (12.1 mM) and (C) L-fucose (12.1 mM). The upper graphs show raw data after the baseline integration. The lower plots (black squares) show the total heat released as a function of total ligand concentration for the titration shown in the upper panel. The solid line represents the best least-squares fit. The grey squares represent the baseline.

Fig. 3. Crystal structure of BP39L and CV39L. (A) Crystal structure of BP39L shown in secondary structure representation (β-strands in yellow and loops in green) (B) Surface representation (yellow) of BP39L colour-coded according to putative binding sites (blue). (C) Crystal structure of CV39L shown in secondary structure representation (β-strands in violet, α-helices in cyan and loops in pink). (D) Surface representation (violet) of CV39L color-coded according to putative binding sites (blue), α-helices shown in cyan.

Fig. 4. Analytical ultracentrifugation experiments of BP39L and CV39L. (A) Equilibrium distributions of BP39L (0.18 mg·ml^{-1}) were obtained at rotor speeds of 14,000 (purple curve), 16,700 (dark blue) and 29,000 rpm (light blue). The residual plot (lower panel) shows the goodness of the fit. For the inserted sedimentation coefficient distributions c(s) of BP39L, the continuous size distribution of sedimenting species resulted in a single peak. (B) Equilibrium distributions of CV39L (0.14 mg·ml^{-1}) were obtained at rotor speeds of 9,000 (purple curve), 11,000 (dark blue) and 18,500 rpm (light blue). For the inserted sedimentation coefficient distributions c(s) of CV39L, the continuous size distribution of sedimenting species resulted in a single peak. Global analysis of multi-speed experiments (for multiple loading concentrations of proteins) was done using SEDPHAT. The figures were created in the program GUSSI 1.0.8 written by Chad Brautigam (University of Texas, Southwestern Medical Centre).

Fig. 5. Characterization of BP39L and CV39L putative binding sites. (A) Stick representation with shown chemistry of putative binding site between blades W1 and W2 of BP39L and (B) CV39L of monomer A. (C) Structural alignment of BP39L (green) and CV39L (pink) putative binding sites (between repeats W1-W2) shown in stick representation with respective color-coded labelling. (D) Sequence alignment of BP39L and CV39L according to putative binding sites of 7 repeats (W1-W7). Amino acids forming binding sites are highlighted with green (BP39L) and pink (CV39L) background. Seven repetitions (W1-W7) are indicated with the main secondary structure elements (color-coded for each β-propeller). In both alignments, black spheres indicate an example of one binding pocket created by the amino acids of blade Wi (W1) and blade Wi+1 (W2).

Fig. 6. Orientations of docked monosaccharides in BP39L and CV39L binding pockets. Individual ligands are coloured and labelled with corresponding colour highlight. O1 atoms of monosaccharides are labelled. Binding site of BP39L (green) with docked methyl α/β-D-mannoside (yellow sticks; A/B), Manα1-3(Manα1-6)Man mannotriose in two different binding

modes (orange or brown sticks, respectively; C/D). Arrows point towards possible oligosaccharide extension. Binding site of CV39L (pink) with docked methyl α/β-L-fucoside (blue sticks; E/F), methyl α/β-D-glucoside (green sticks; G/H), methyl β-D-mannoside (yellow sticks; I), methyl β-L-galactoside (teal sticks; J) and methyl α/β-D-galactoside (cyan sticks; K/L).

Fig. 7. Structural comparison of BP39L and CV39L with other seven-bladed β-propeller lectins. Structural alignment of BP39L (green-yellow) putative binding site (shown site I - green sticks) with (A) PVL (orange) (PDB 2BWR) GlcNAc-binding site (orange sticks) and with (B) PHL (violet) (PDB 5MXE) galactose-binding site (violet sticks). (C) Close-up view of BP39L aligned with site VI of PHL. Structural alignment of CV39L (magenta-pink) putative binding site (shown site I - pink sticks) with (D) PVL (orange) (PDB 2BWR) GlcNAc-binding site (orange sticks) and (E) PHL (violet) (PDB 5MXE) galactose-binding sites (violet sticks). (F) Close-up view of CV39L aligned with site VI of PHL.
The labelling of the binding sites is color-coded to match the colour of the stick representation.

Table 1. Microcalorimetry titration data for binding of saccharides to BP39L and CV39L. Thermodynamic parameters with standard deviations (averaged from three independent measurements) for the binding of saccharides to BP39L and CV39L. (n = number of binding sites, K_a = association constant, K_d = dissociation constant, ΔG = change in Gibb's free energy, ΔH = change in enthalpy, -TΔS = change in entropy)

Protein	Ligand	n	K_a (10^3 M^{-1})	K_d (mM)	−G (kJ/mol)	−H (kJ/mol)	−TΔS (kJ/mol)
BP39L							
	α1-3,α1-6-Mannotriose	8.56 ± 0.04	5.15 ± 0.09	0.19 ± 0.003	-21.2 ± 0.05	-28.8 ± 0.22	7.6 ± 0.26
	D-Man	7*	0.78 ± 0.045	1.28 ± 0.074	-16.5 ± 0.95	-37.9 ± 1.42	21.4 ± 1.24
	Man α1-2- Man	7*	0.23 ± 0.060	4.38 ± 1.140	-13.5 ± 3.51	-20.0 ± 5.67	6.54 ± 1.70
n. b.+	L-Fuc, methyl-β-L-fucoside, Lewis b, exopolysaccharides (*B. cenocepacia*), N-acetylglucosamine						
CV39L							
	L-Gal	8.56 ± 0.06	4.67 ± 0.12	0.21 ± 0.005	-21.0 ± 0.06	-35.8 ± 0.40	14.9 ± 0.47

L-Fuc	7.87 ± 0.12	3.21 ± 0.14	0.31 ± 0.010	-20.0 ± 0.11	-36.7 ± 1.08	16.7 ± 1.19
D-maltose	7*	0.26 ± 0.002	3.74 ± 0.025	-13.9 ± 0.00	-26.8 ± 0.57	12.9 ± 0.61
D-Man	7*	0.14 ± 0.026	7.21 ± 1.320	-12.3 ± 0.45	-25.6 ± 3.65	13.3 ± 4.13

n. b.[+] D-Gal, L-rhamnnose, N-acetylgalactosamine, BGA, BGB, sialic acid, exopolysaccharides (*B. cenocepacia*), Fuc1-2Gal, Galβ1-3GalNAc, GalNAcα (1-3Fuc)1-2Gal, Fuc1-6NAcGlc, Man1-3(Man1-6)Man, lactose

*The stoichiometry value was fixed to expected number of binding sites due to low affinity interaction.

[+]No observed binding

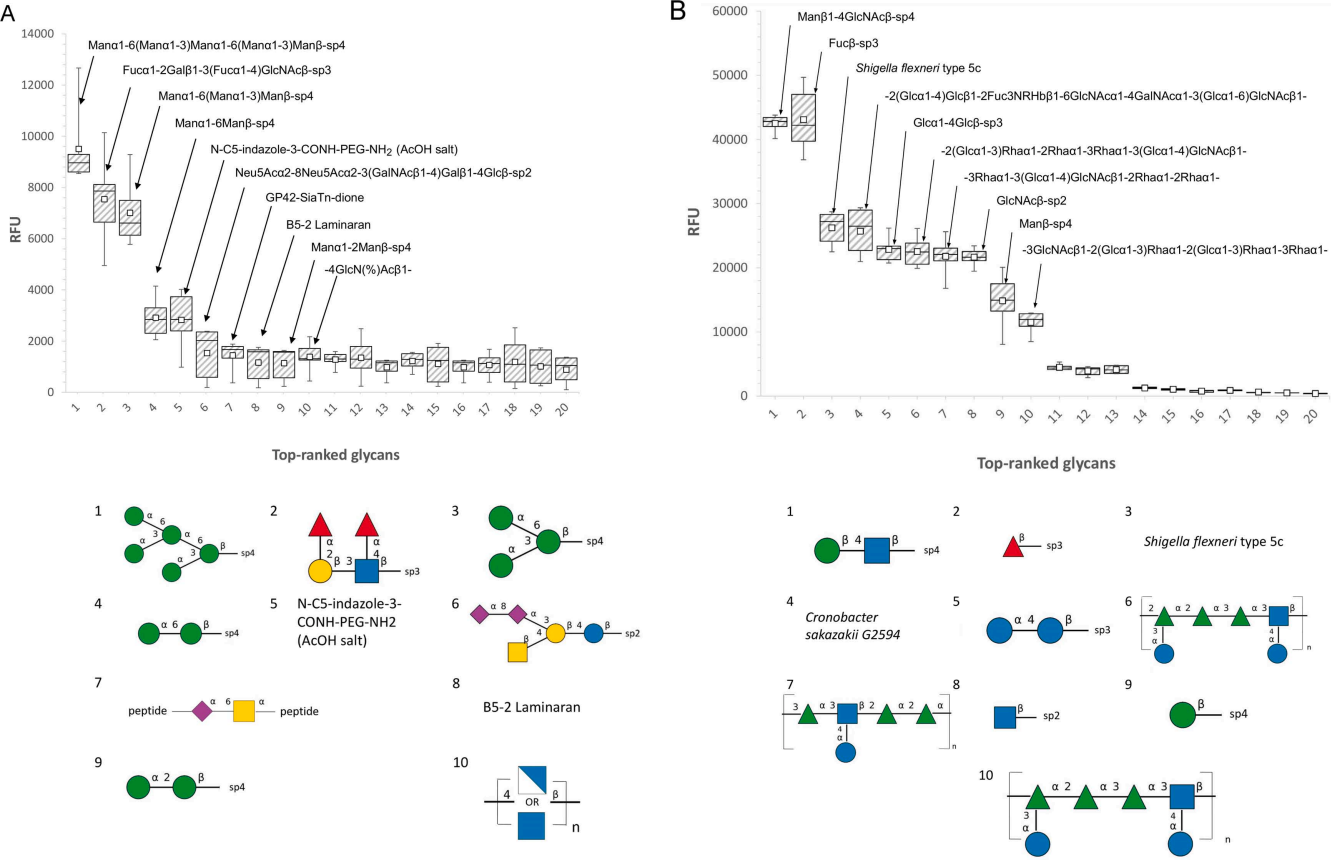
Table 2. Data collection and refinement statistics. Values in parentheses are for the highest resolution shell.

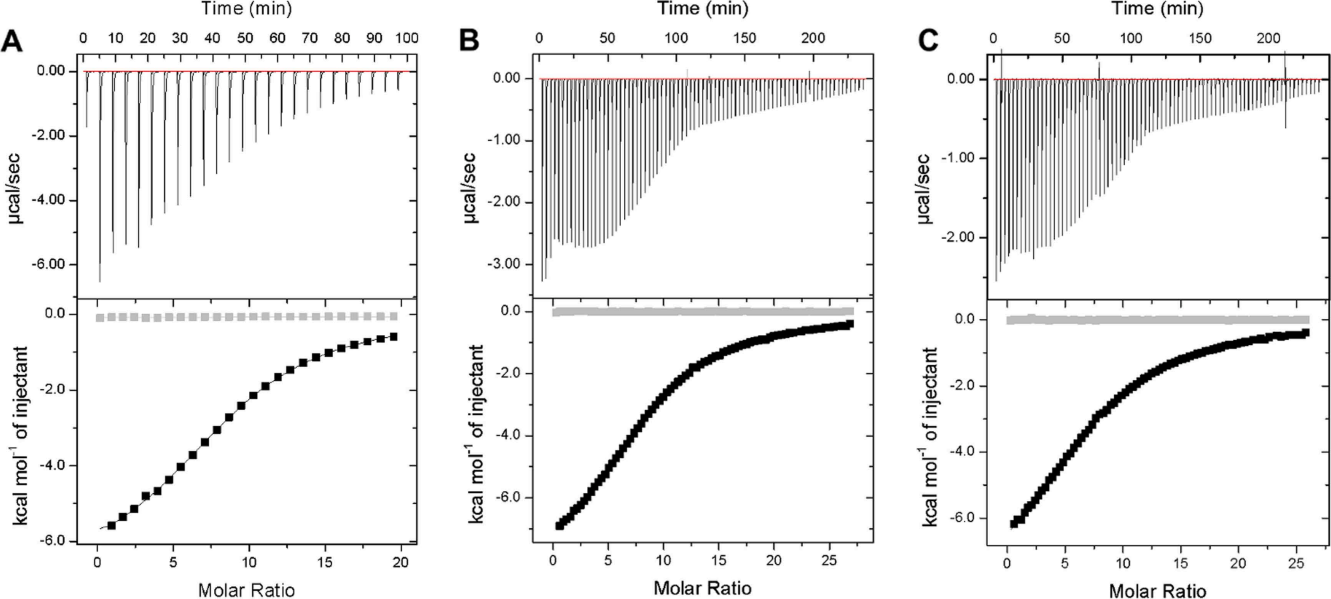
	BP39L		**CV39L**	
	Eu derivative	Native	Eu derivative	Native
PDB code	-	6GXR	-	6GXS
Space group	I222	I222	$P2_12_12_1$	$P2_12_12_1$
Wavelength (Å)	1.776	0.979	1.776	0.967
Unit-cell parameters				
a (Å)	46.07	46.08	65.69	65.71
b (Å)	87.52	87.55	124.10	123.78
c (Å)	159.29	159.10	181.00	180.93
Resolution range (Å)	76.70-1.84 (1.94-1.84)	79.55–1.70 (1.79 – 1.70)	124.11-2.55 (2.63 – 2.55)	80.93-1.80 (1.90 – 1.80)
Reflections measured	695843 (77031)	476560 (70482)	1273263 (118313)	1026231 (148690)
Unique reflections	27798 (3723)	33933 (4942)	48771 (4457)	137285 (19822)
Completeness (%)	97.8 (91.8)	94.7 (95.9)	99.3 (99.6)	100.0 (100.0)
$CC_{1/2}$ (%)	99.8 (98.3)	100.0 (92.4)	99.7 (80.3)	99.9 (95.3)
R_{merge}[†]	0.099 (0.317)	0.071 (0.430)	0.279 (1.783)	0.102 (0.363)
Multiplicity	25.0 (20.7)	14.0 (14.3)	26.1 (26.5)	7.5 (7.5)
$<I/\sigma (I)>$	27.8 (11.9)	25.1 (6.2)	15.0 (2.4)	14.7 (4.9)
Refinement resolution (Å)	-	79.55–1.70	-	73.04-1.80
σ cutoff	-	0	-	0
Reflections used	-	32232	-	130295
Reflections used for Rfree	-	1615	-	6547
R factor[‡] (%)	-	15.74	-	15.28
Rfree[‡] (%)	-	19.61	-	18.85
R.m.s.d. bond lengths (Å)	-	0.023	-	0.023
R.m.s.d. bond angles (°)	-	2.099	-	2.014
No. of water molecules	-	317	-	1275

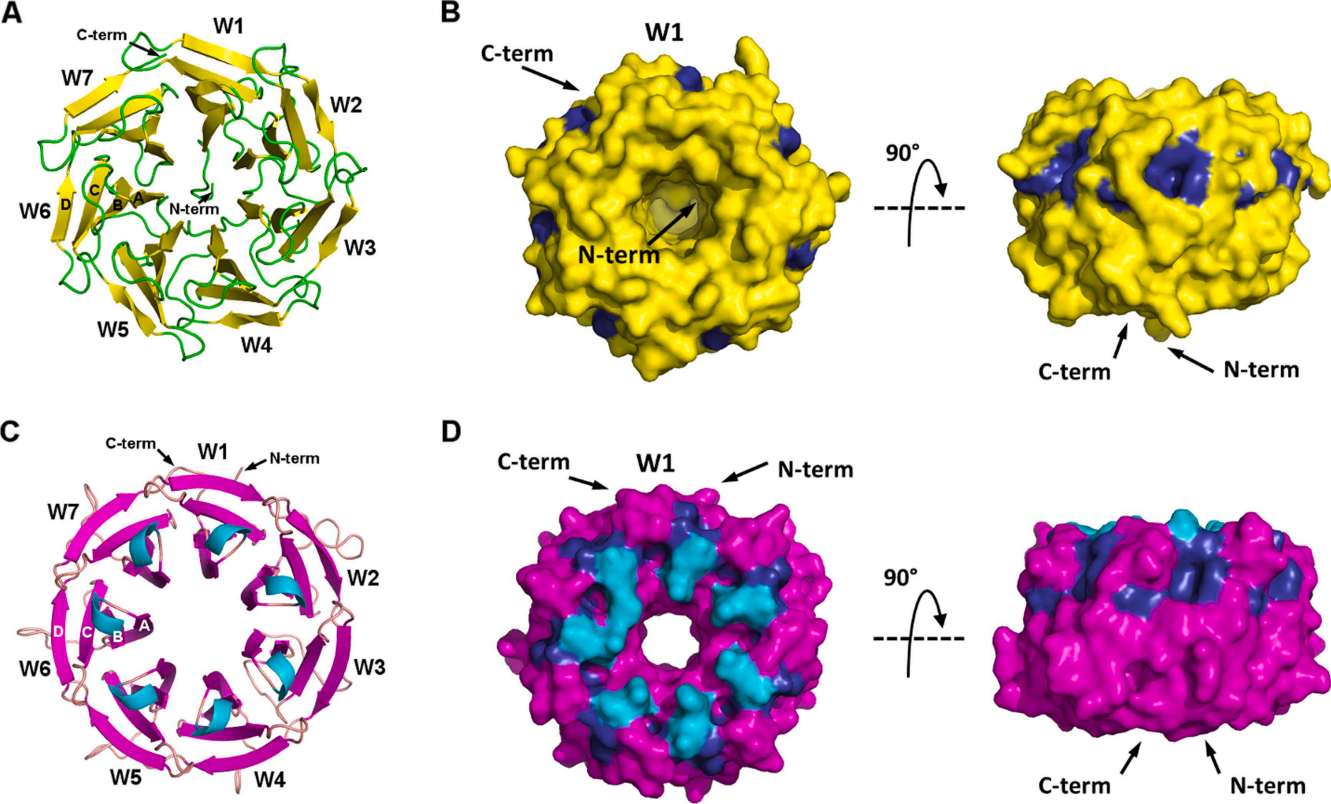
No. of non-H atoms (total)	-	2928	-	12196
Ramachandran plot				
Residues in most favourable regions (%)	-	98.0	-	98.0
Residues in additional allowed regions (%)	-	2.0	-	2.0

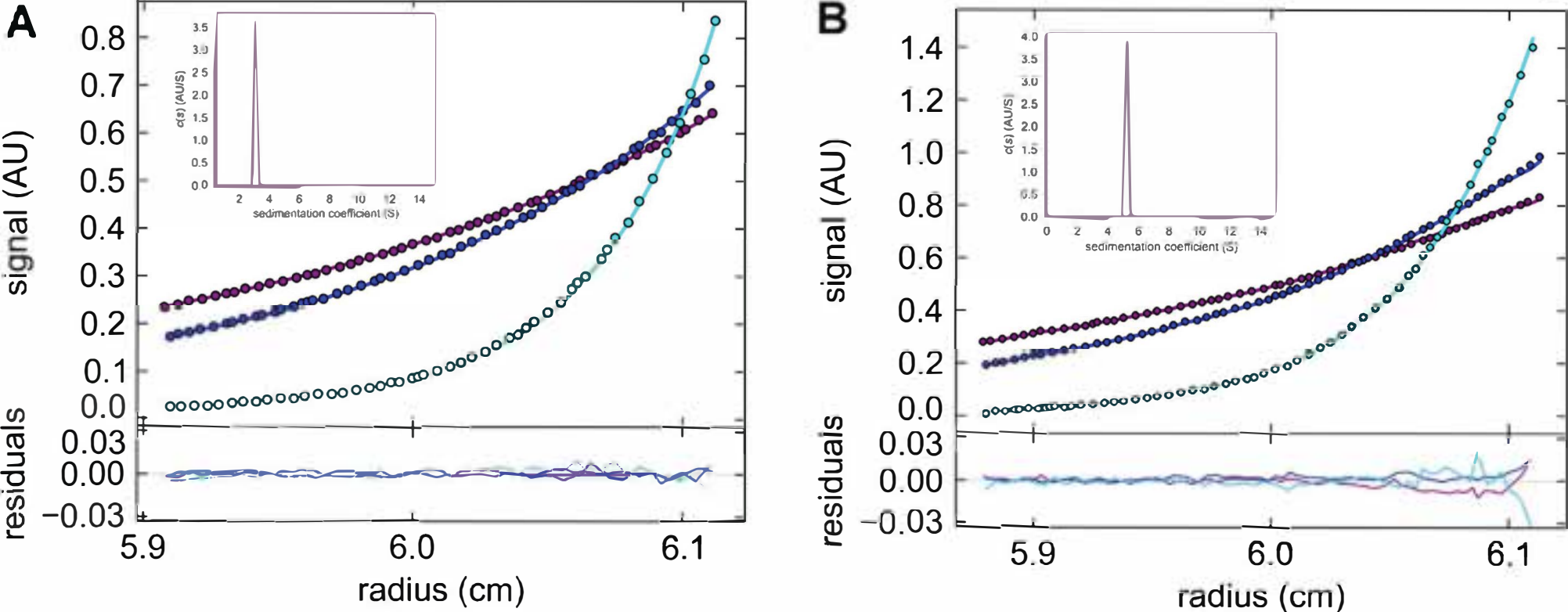
†Rmerge = $\sum I_i - \langle I \rangle | / \sum I_i$, where I_i is the intensity of observation and <I> is the mean value for that reflection.

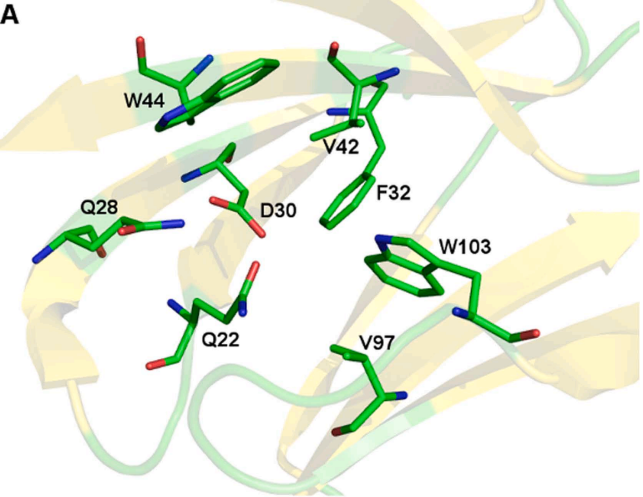
‡R factor = $\sum ||F_o(h)| - |F_c(h)|| / \sum_h F_o(h)$, where F_o and F_c are the observed and calculated structure-factor amplitudes, respectively.

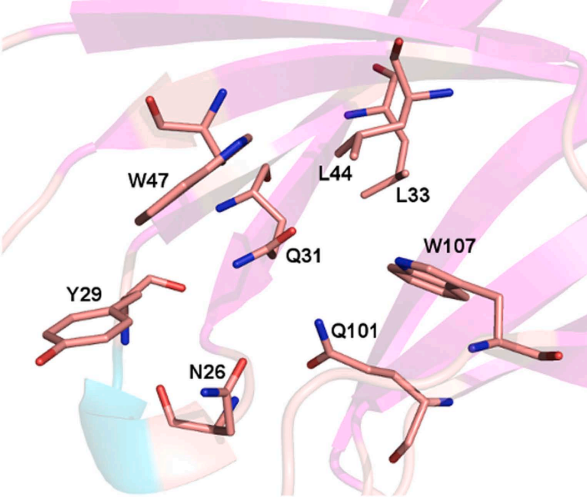

A
RFU
Manα1-6(Manα1-3)Manα1-6(Manα1-3)Manβ-sp4
Fucα1-2Galβ1-3(Fucα1-4)GlcNAcβ-sp3
Manα1-6(Manα1-3)Manβ-sp4
Manα1-6Manβ-sp4
N-C5-indazole-3-CONH-PEG-NH2 (AcOH salt)
Neu5Acα2-8Neu5Acα2-3(GalNAcβ1-4)Galβ1-4Glcβ-sp2
GP42-SiaTn-dione
B5-2 Laminaran
Manα1-2Manβ-sp4
-4GlcN(%)Acβ1-
Top-ranked glycans
B
RFU
Manβ1-4GlcNAcβ-sp4
Fucβ-sp3
Shigella flexneri type 5c
-2(Glcα1-4)Glcβ1-2Fuc3NRHbβ1-6GlcNAcα1-4GalNAcα1-3(Glcα1-6)GlcNAcβ1-
Glcα1-4Glcβ-sp3
-2(Glcα1-3)Rhaα1-2Rhaα1-3Rhaα1-3(Glcα1-4)GlcNAcβ1-
-3Rhaα1-3(Glcα1-4)GlcNAcβ1-2Rhaα1-2Rhaα1-
GlcNAcβ-sp2
Manβ-sp4
-3GlcNAcβ1-2(Glcα1-3)Rhaα1-2(Glcα1-3)Rhaα1-3Rhaα1-
Top-ranked glycans
peptide
peptide
B5-2 Laminaran
N-C5-indazole-3-CONH-PEG-NH2 (AcOH salt)
Shigella flexneri type 5c
Cronobacter sakazakii G2594
OR

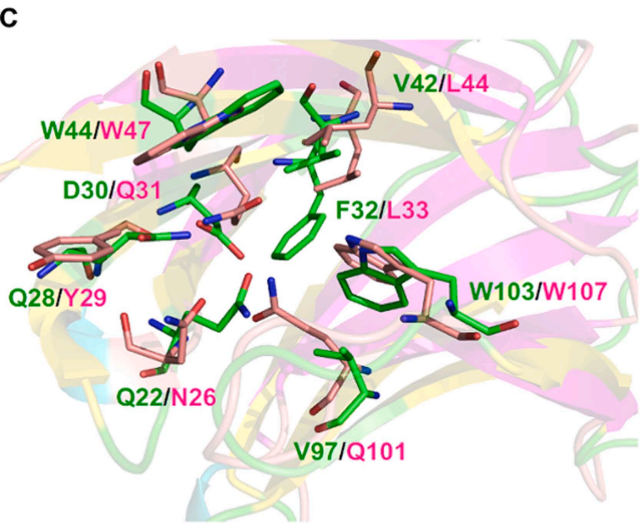

A
B
C
Time (min)
µcal/sec
kcal mol^{-1} of injectant
Molar Ratio


A
C-term
W1
W7
W2
W6
D
C
B
A
N-term
W3
W5
W4
B
W1
C-term
N-term
90°
C-term
N-term
C
C-term
W1
N-term
W7
W2
D
C
B
A
W6
W3
W5
W4
D
C-term
W1
N-term
90°
C-term
N-term

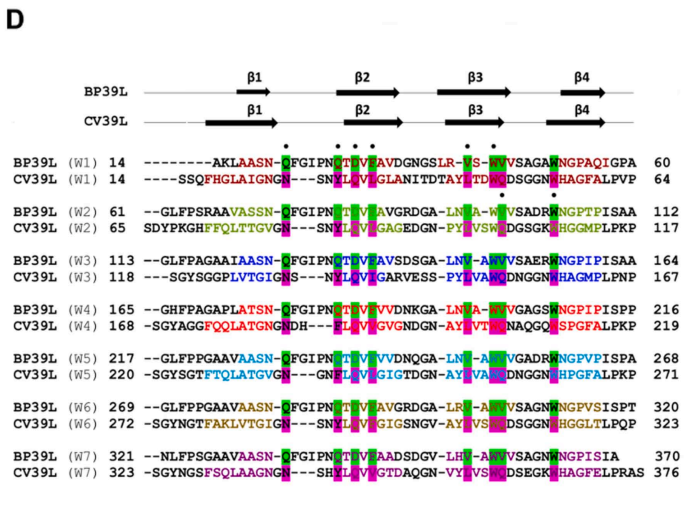

A
B
signal (AU)
residuals
radius (cm)
c(s) (AU/S)
sedimentation coefficient (S)

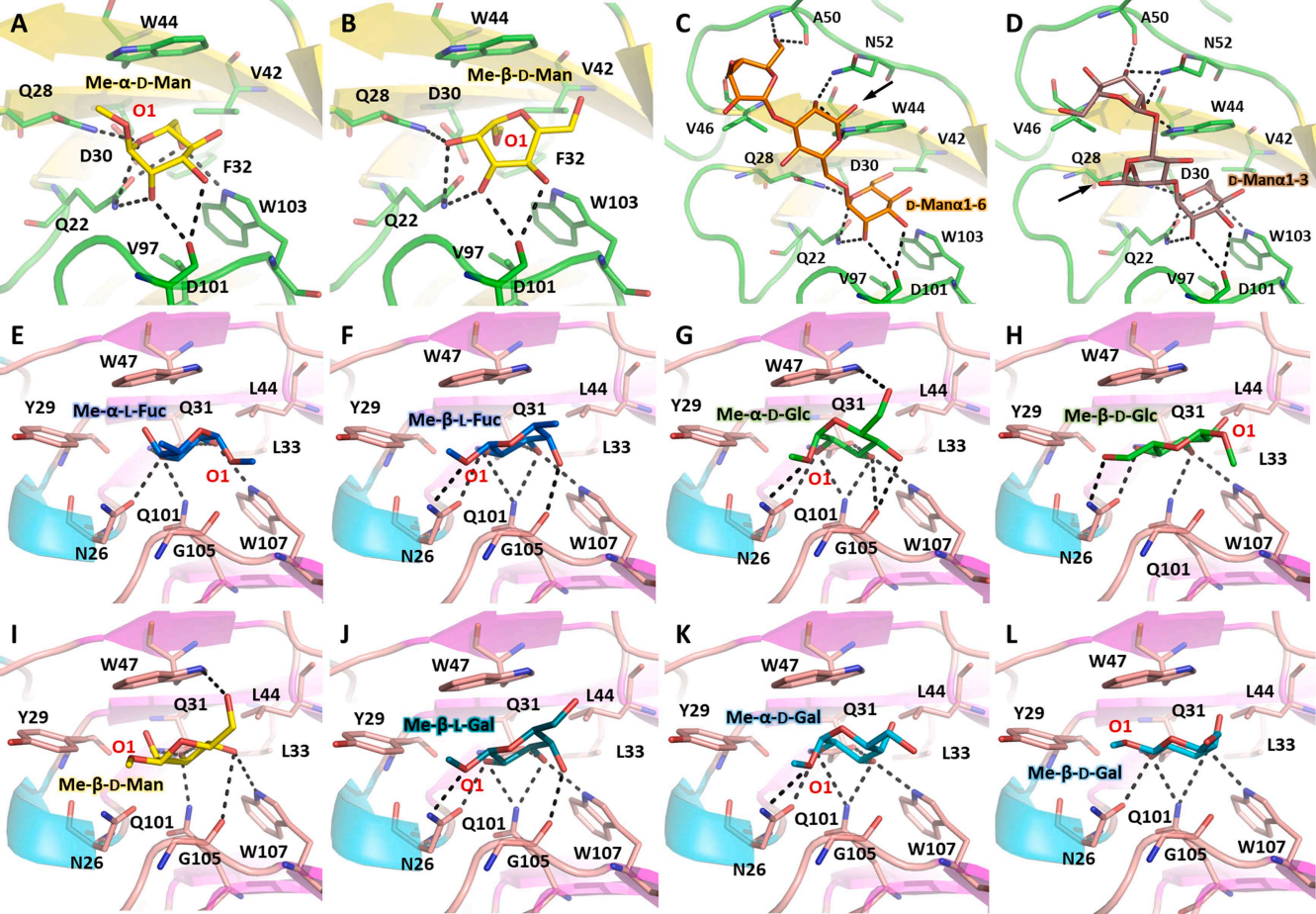

A
W44
V42
Q28
D30
F32
W103
Q22
V97

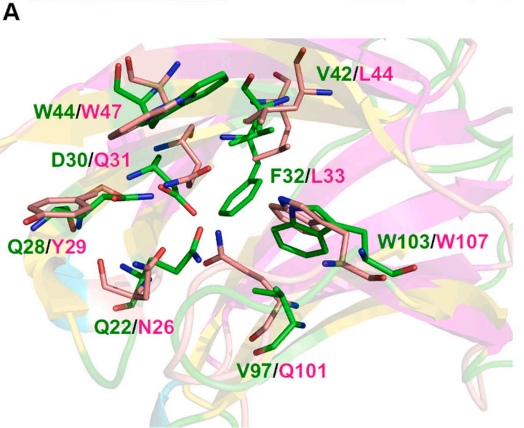

B
W47
L44
L33
Q31
W107
Y29
Q101
N26


C
V42/L44
W44/W47
D30/Q31
F32/L33
Q28/Y29
W103/W107
Q22/N26
V97/Q101

D

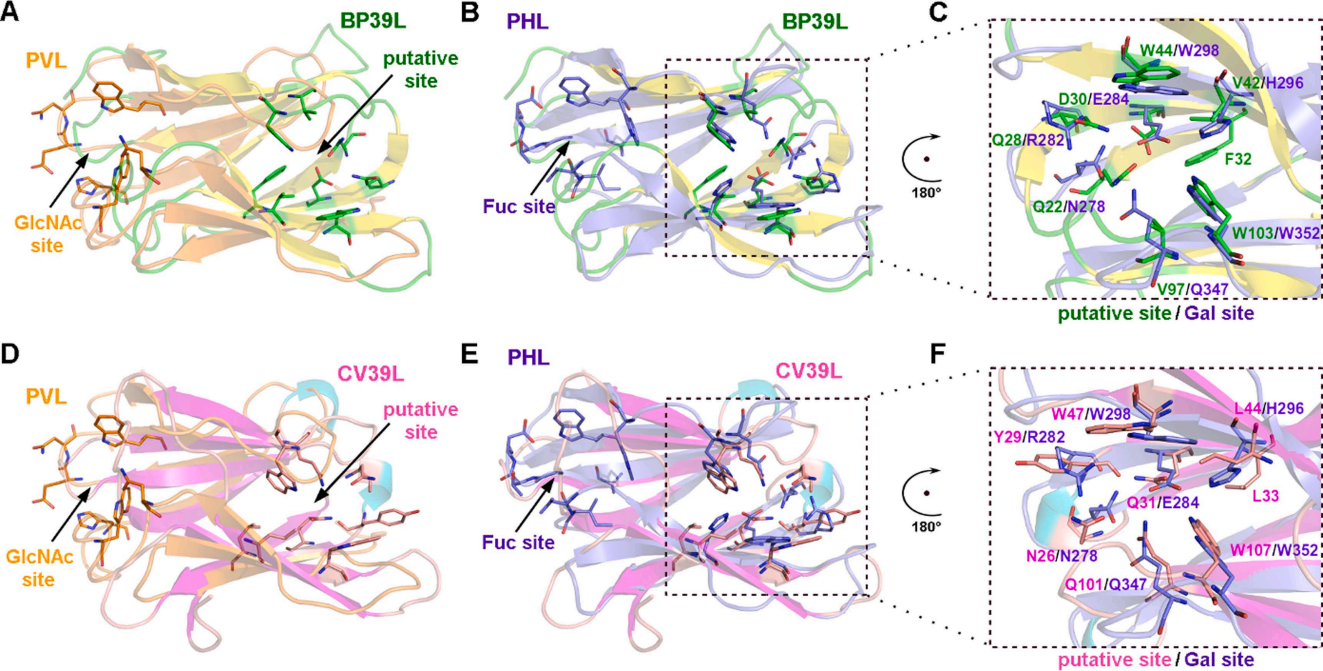

BP39L β1 β2 β3 β4
CV39L β1 β2 β3 β4
BP39L (W1) 14 --------AKLAASN-QFGIPNLTTVTAVDGNGSLR-LS-AVVSAGAWNGPAQIGPA 60
CV39L (W1) 14 ----SSQFHGLAIGNGN---SNYLQVVGLANITDTAYLTDWQDSGGNMHAGFALPVP 64
BP39L (W2) 61 --GLFPSRAAVASSN-QFGIPNLTTVTAVGRDGA-LNLA-AVVSADRWNGPTPISAA 112
CV39L (W2) 65 SDYPKGHFFQLTTGVGN---SNYLQVVGAGEDGN-PYLVSWLDGSGKMHGGMPLPKP 117
BP39L (W3) 113 --GLFPAGAAIAASN-QFGIPNLTTVTAVSDSGA-LNV-AAVVSAERWNGPIPISAA 164
CV39L (W3) 118 ---SGYSGGPLVTGIGNS---NYLQVVGARVESS-PYLVAWQDNGGNMHAGMPLPNP 167
BP39L (W4) 165 --GHFPAGAPLATSN-QFGIPNLTTVTVVDNKGA-LNLA-AVVGAGSWNGPIPISPP 216
CV39L (W4) 168 -SGYAGGFQQLATGNGNDH---YLQVVGVGNDGN-AYLVTGCNAQGQQSPGFALPKP 219
BP39L (W5) 217 --GLFPGAAVAASN-QFGIPNLTTVTVVDNQGA-LNL-AAVVGADRWNGPVPISPA 268
CV39L (W5) 220 -SGYSGTFTQLATGVGN---GNYLQVVGIGTDGN-AYLVAGCDNGGNMHPGFALPKP 271
BP39L (W6) 269 --GLFPPGAAVAASN-QFGIPNLTTVTAVGRDGA-LRL-AAVVSAGNWNGPVSISPT 320
CV39L (W6) 272 -SGYNGTFAKLVTGIGN---SNYLQVVGIGSNGV-AYLVSGCDSGGNMHGGLTLPQP 323
BP39L (W7) 321 --NLFPSGAAVAASN-QFGIPNLTTVTAADSDGV-LHL-AAVVSAGNWNGPISIA 370
CV39L (W7) 323 -SGYNGSFSQLAAGNGN---SHYLQVVGTDAQGN-VYLVSGCDSEGKMHAGFELPRAS 376


A
W44
Me-α-D-Man
V42
Q28
O1
D30
F32
W103
Q22
V97
D101
B
W44
Me-β-D-Man
V42
Q28
D30
O1
F32
W103
Q22
V97
D101
C
A50
N52
W44
V46
V42
Q28
D30
D-Manα1-6
W103
Q22
V97
D101
D
A50
N52
W44
V46
V42
Q28
D30
D-Manα1-3
W103
Q22
V97
D101
E
W47
L44
Y29
Me-α-L-Fuc
Q31
L33
O1
Q101
N26
G105
W107
F
W47
L44
Y29
Me-β-L-Fuc
Q31
L33
O1
Q101
N26
G105
W107
G
W47
L44
Y29
Me-α-D-Glc
Q31
L33
O1
Q101
N26
G105
W107
H
W47
L44
Y29
Me-β-D-Glc
Q31
O1
L33
Q101
N26
W107
I
W47
Q31
L44
Y29
O1
L33
Me-β-D-Man
Q101
N26
G105
W107
J
W47
L44
Y29
Me-β-L-Gal
Q31
L33
O1
Q101
N26
G105
W107
K
W47
L44
Y29
Me-α-D-Gal
Q31
L33
O1
Q101
N26
G105
W107
L
W47
L44
Y29
O1
Q31
L33
Me-β-D-Gal
Q101
N26
G105
W107



B

```
BP39L (W1):  AKLAASN-----QFGIPNQTDVFAVDGNGSLR-VS-WVVSAGAWNGPAQIGPAG     61
CV39L (W1):  SSQFHGLAIGNGN---SNYLQVLGLANITDTAYLTDWQDSGGNWHAGFALPVPS     65

BP39L (W2):  LFPSRAAVASSN---QFGIPNQTDVFAVGRDGALNVA-WVVSADRWNGPTPISAAG  113
CV39L (W2):  DYPKGHFFQLTTGVGN---SNYLQVLGAGEDGNPYLVSWQDGSGKWHGGMPLPKPS  118

BP39L (W3):  LFPAGAAIAASNQFGIPNQTDVFAVSDSGALNV-AWVVSAERWNGPIPISAAG  165
CV39L (W3):  GYSGGPLVTGIGNS---NYLQVIGARVESSPYLVAWQDNGGNWHAGMPLPNPS  168

BP39L (W4):  HFPAGAP-LATSN-QFGIPNQTDVFVVDNKGALNVA-WVVGAGSWNGPIPISPPG  217
CV39L (W4):  GYAGGFQQLATGNGNDH---FLQVVGVGNDGNAYLVTWQNAQGQWSPGFALPKPS  220

BP39L (W5):  LFPPGAAVAASN--QFGIPNQTDVFVVDNQGALNV-AWVVGADRWNGPVPISPAG  269
CV39L (W5):  GYSGTFTQLATGVGN---GNFLQVLGIGTDGNAYLVAWQDNGGNWHPGFALPKPS  272

BP39L (W6):  LFPPGAAVAASN--QFGIPNQTDVFAVGRDGALRV-AWVVSAGNWNGPVSISPTN  321
CV39L (W6):  GYNGTFAKLVTGIGN---SNYLQVFGIGSNGVAYLVSWQDSGGNWHGGLTLPQPS  324

BP39L (W7):  LFPSGAAVAASN--QFGIPNQTDVFAADSDGVLHV-AWVVSAGNWNGPISIA    370
CV39L (W7):  GYNGSFSQLAAGNGN---SHYLQVVGTDAQGNVYLVSWQDSEGKWHAGFELPRAS  376
```


A
PVL
BP39L
putative site
GlcNAc site
B
PHL
BP39L
Fuc site
180°
C
W44/W298
V42/H296
D30/E284
Q28/R282
F32
Q22/N278
W103/W352
V97/Q347
putative site / Gal site
D
PVL
CV39L
putative site
GlcNAc site
E
PHL
CV39L
Fuc site
180°
F
W47/W298
L44/H296
Y29/R282
L33
Q31/E284
N26/N278
W107/W352
Q101/Q347
putative site / Gal site