STRUCTURAL INSIGHTS INTO THE ANTI-HIV ACTIVITY OF THE OSCILLATORIA AGARDHII AGGLUTININ HOMOLOG LECTIN FAMILY

Leonardus M.I. Koharudin¹, Sireesha Kollipara², Christopher Aiken², and Angela M. Gronenborn¹

From the Departments of Structural Biology¹, University of Pittsburgh School of Medicine, Pittsburgh, PA 15260, USA and ²Department of Pathology, Microbiology and Immunology², Vanderbilt University School of Medicine, Nashville, TN 37232, USA

*Running title: Structural, carbohydrate binding and functional studies of OAAHs

Address correspondence to: Angela M. Gronenborn, Department of Structural Biology, University of Pittsburgh School of Medicine, 3501 Fifth Avenue, BST3/Rm1050, Pittsburgh, PA 15260, Tel: (412) 648 9959, Fax: (412) 648 9008, Email: amg100@pitt.edu.

Keywords: Carbohydrate binding protein; protein complex; HIV; NMR; X-ray crystallography

Background: The Oscillatoria Agardhii agglutinin homolog (OAAH) proteins constitute a novel lectin family.
Results: 3D structures, carbohydrate binding specificities, and antiviral activity data for several members were determined.
Conclusion: All members display potent anti-HIV activity.
Significance: Our results uncovered the structural basis of protein-carbohydrate recognition in this novel lectin family and provide insights into the molecular basis of their HIV inactivation properties.

SUMMARY
Oscillatoria Agardhii agglutinin homolog (OAAH) proteins belong to a recently discovered lectin family. All members contain a sequence repeat of ~66 amino acids, with the number of repeats varying among different family members. Apart from data for the founding member OAA, neither three-dimensional structures, nor information about carbohydrate binding specificities, nor antiviral activity data is available up to now for any other members of the OAAH family. In order to elucidate the structural basis for the antiviral mechanism of OAAHs, we determined the crystal structures of Pseudomonas fluorescens and Myxococcus Xanthus lectins. Both proteins exhibit the same fold, resembling the founding family member, OAA, with minor differences in loop conformations. Carbohydrate binding studies by NMR and X-ray structures of glycan-lectin complexes reveal that the number of sugar binding sites corresponds to the number of sequence repeats in each protein. Like for OAA, tight and specific binding to α3,α6-mannopentaose was observed. All the OAAH proteins described here exhibit potent anti-HIV activity at comparable levels. Altogether, our results provide structural details of the protein-carbohydrate interaction for this novel lectin family and insights into the molecular basis of their HIV inactivation properties.

The development of carbohydrate binding proteins (lectins) as microbicidal drugs represents a novel therapeutic approach in the fight against HIV transmission (1-3). This strategy is particularly useful for women in some regions of the world where social and psychological barriers are substantial and not easily overcome. For example, due to economic and societal pressures, diagnosis and treatment for HIV infections may not be readily available or are stigmatized. Therefore, using microbicides designed to significantly reduce transmission of sexually transmitted viral pathogens when applied topically to genital mucosal surfaces is potentially a powerful strategy, given that application in cream form is discreet and can be completely controlled by women.

One of several lectins that were selected as a candidate for lectin-based microbicides is Cyanovirin-N (CV-N) (4-6). It exhibits potent
antiviral activity by blocking virus entry into the host target cells through specific and tight binding to Manα(1-2)Man linked mannose substructures (7-12), in particular the D1 or D3 arms of the remarkably enriched high-mannose N-linked glycans, present on the HIV envelope glycoprotein 120 (gp120), thereby compromising the required conformational changes in gp120/gp41 for fusion with the target cell membrane. Following CV-N’s discovery, there have been a growing number of lectins that are also known to bind to the high mannose glycans on gp120, such as DC-SIGN (13), a dendritic cell surface receptor that captures HIV-1 and facilitates infection of HIV-1 permissive cells in trans by endocytosis (14), as well as a variety of cyanobacterial- or actinomycetal-derived proteins like Scytovirin (15), Griffithsin (16), MVL (17), and Actinohivin (18). Notably, the target epitopes on Man-8/9 and binding modes for these different lectins are quite distinct (7,10,19-22). In an effort to characterize as many possible mannose-targeting lectins structurally, as well as with respect to their recognition epitopes, we recently investigated a novel antiviral lectin from the cyanobacterium Oscillatoria Agardhii (named Oscillatoria Agardhii Agglutinin; OAA) (23,24). Surprisingly, the compact, β-barrel-like architecture of OAA is very different from previously characterized lectin structures and unique among all available protein structures in the protein data bank. Most importantly, so far OAA’s carbohydrate recognition of Man-9 is also unique, compared to all other anti-viral lectins. While most of the known HIV-inactivating lectins recognize the reducing or non-reducing end mannoses of Man-8/9, OAA recognizes the branched core unit of Man-8/9.

Genes coding for OAA-homologous proteins have recently been discovered in a number of other prokaryotic microorganisms, including cyanobacteria, proteobacteria, and chlorobacteria, as well as in an eukaryotic marine red alga (25). Similar to OAA, these proteins, termed OAAHs, contain a sequence repeat of ~66 amino acids, with the number of repeats varying for different family members. For example, the 133 residue containing Pseudomonas fluorescens and Herpetosiphon aurantiacus homologs contain two sequence repeats, like OAA, while Lyngbya sp., Burkholderia oklahomensis EO147, Stigmatella aurantiaca DW4/3-1, Myxococcus Xanthus, and Eucheuma serra homologs contain four sequence repeats over lengths of 246 to 268 residues (25).

Apart from data for the founding member OAA, neither three-dimensional structures nor information about carbohydrate binding specificities nor antiviral activity is available up to now for any other member of the OAAH family. In order to further characterize this important lectin family, we carried out structural analyses of two additional members of the OAAH family. We determined the structures of the Pseudomonas fluorescens and the Myxococcus Xanthus OAAHs that contain two and four sequence repeats, respectively. We delineated the conserved sequence and structure features in these proteins and determined their sugar binding specificities by NMR titrations and X-ray crystallography. In addition, HIV assays were carried out and we observed that the antiviral potency for all OAAH family members investigated here is comparable and is not significantly enhanced when more binding sites are present. Altogether, our results provide the basis for expanding our knowledge with respect to protein-carbohydrate interaction in general and recognition of α3,α6-mannopentaose by this class of lectins in particular.

**EXPERIMENTAL PROCEDURES**

**Protein expression, purification, and crystallization** - PFA and MBHA were expressed and purified as described previously for OAA (23). Briefly, synthetic PFA and MBHA genes encoding residues 1−133 and 1−267 (25) were cloned into the pET-26b(+) expression vector (Novagen), using NdeI and XhoI restriction sites at the 5′ and 3′ ends, respectively. *E. coli* Rosetta2 (DE3) cells (Novagen) were transformed with the pET-26b(+) vector for protein expression. Cells were initially grown at 37°C, induced with 1 mM IPTG, and grown for ~18 h at 16°C for protein expression.

The hybrid OAAH protein (OPA) was created by combining the OAA and PFA genes via a six nucleotide linker encoding N-G residues, thus generating a 267 residue chimera. Individual DNA
fragments encoding residues M1 to T133 of OAA (PCR1) and residues S2 to E133 of OAA (PCR2) were amplified using the parent vectors as templates. PCR1 (amplification of OAA utilized oligonucleotides 5'-CGGAAATTAATACGACTCACTATAGGGG-3' (T7 promoter) as forward and 5'-CCGCATATTTGCTACCGTTCGTCA-GGGTACCTTTGAAGCCGATC-3' (R1) as reverse primers. Note that the first 13 (italic), the ACCGTT (bolded), and the last 25 (underlined) nucleotides in the reverse primer (R1) are nucleotides that code for the first few N-terminal residues of PFA, the N-G linker residues, and the last few C-terminal residues of OAA. PCR2 (amplification of PFA) was performed with the oligonucleotides 5'-GGTACCGTGACGAGTACAATATGCGGTGGCCAACCAGTG-3' as forward (F2) and 5'-CAAAAAACCCC-TCAAGACCCGTTAGAG-3' (T7 terminator) as reverse primers. For PCR2, the first 12 nucleotides (underlined), the ACCGTT (bolded), and the last 26 (italic) nucleotides in the forward primer (F2) are nucleotides that encode for the last few C-terminal residues of OAA, the N-G linker residues, and the first few N-terminal residues of PFA. PCR3, as the final step was performed after mixing equal amounts of these two PCR products and amplifying with T7 promoter and T7 terminator primers, respectively. Once the complete DNA fragment encoding the full length OPA was generated, the fragment was cloned into the pET-26b(+) expression vector (Novagen), using NdeI and XhoI restriction sites at the 5' and 3' ends, respectively. E. coli Rosetta2 (DE3) cells (Novagen) were transformed with the pET-26b(+)–OPA vector for protein expression. As for PFA and MBHA, cells were initially grown at 37 °C, induced with 1 mM IPTG, and grown for ~18 h at 16°C for protein expression. Protein was prepared from the soluble fraction after rupturing the cells by sonication. The cell lysate was centrifuged to remove cell debris and, after centrifugation, the supernatant was dialyzed over night against 20 mM Tris-HCl buffer (pH 8.5) for OPA and against 20 mM NaAcetate buffer (pH 5.0) for PFA and MBHA. The first purification step involved anion-exchange chromatography on a Q(HP) column (GE Healthcare) for OPA and a SP(HP) column (GE Healthcare) for PFA and MBHA. Fractions containing pure protein were collected and concentrated up to 40 mg/ml using centriprep devices (Millipore). For crystallization and NMR, the buffer was exchanged to 20 mM Tris-HCl buffer, 100 mM NaCl, 3 mM NaN₃, (pH 8.0) and to 20 mM NaAcetate, 3 mM NaN₃, 90/10% H₂O/D₂O (pH 5.0), respectively.

Crystallization trials of apo PFA and apo MBHA were carried out by the sitting drop vapor diffusion at room temperature using drops consisting of 2 µl protein and 2 µl of reservoir solutions at a protein concentration of 40 mg/ml. Well-diffracting crystals were obtained in 1.0 M sodium citrate and 0.1 M imidazole (pH 8.0) for PFA and in 0.2 M ammonium sulphate, 0.1 M sodium cacodylate trihydrate (pH 6.5) and 30 % polylene glycol 8000 for MBHA after ~1 week. The crystals for the carbohydrate-bound MBHA were obtained by co-crystallization. Drops consisted of 2 µl reservoir solution and 2 µl of premixed protein and α₃,α₆-mannopentaose at a molar ratio of 1:6, with the protein concentration at 30 mg/ml. Well-diffracting crystals of α₃,α₆-mannopentaose-bound MBHA were obtained using 0.05 M monobasic potassium phosphate and 20 % w/v polylene glycol 3,350.

**Diffraction data collection and structure determination of apo PFA, apo MBHA, and glycan-bound MBHA**

X-ray diffraction data for apo PFA, apo MBHA, and glycan-bound MBHA crystals were collected up to 1.70, 1.60, and 1.76 Å resolution, respectively, on flash-cooled crystals (-180°C) using a Rigaku FR-E generator with a Saturn 944 CCD detector or a R-AXIS IV image plate detector at a wavelength corresponding to the copper edge (1.54 Å). All diffraction data were processed, integrated, and scaled using d*TREK software (26), and eventually converted to mtz format using the CCP4 package (27). The unit-cell dimensions of the C2 apo PFA crystal were...
\(a = 58.95\ \text{Å}, \ b = 70.31\ \text{Å}, \ \text{and} \ c = 62.29\ \text{Å} \ (\beta = 111.73^\circ)\) with an estimated solvent content of 42.52 % (\(V_m = 2.14\ \text{Å}^3/\text{Da}\)), based on Matthews Probability Calculator (http://www.ruppweb.org/Mattprob/). The crystals contained two PFA molecules per asymmetric unit. The unit-cell dimensions of the P3\(_1\) apo and the P2\(_{1}2_12_1\) \(\alpha_3\alpha_6\)-mannopentaose-bound MBHA crystals were \(a = 76.10\ \text{Å}, \ b = 76.10\ \text{Å}, \ \text{and} \ c = 37.67\ \text{Å} \ (\gamma = 120^\circ)\) and \(a = 45.59\ \text{Å}, \ b = 57.32\ \text{Å}, \ \text{and} \ c = 105.97\ \text{Å} \ (\alpha, \beta, \text{and} \ \gamma = 90^\circ)\) with an estimated solvent content of 45.48 % (\(V_m = 2.26\ \text{Å}^3/\text{Da}\)) and 50.37 % (\(V_m = 2.48\ \text{Å}^3/\text{Da}\)), respectively. Both crystals contained one polypeptide molecule per asymmetric unit.

Phases for both apo PFA and MBHA crystals were determined by molecular replacement using the previously determined structure of OAA (PDB:3S5V) (23,24) as structural probe in PHASER (28). After generation of the initial model, the chain was re-built using the program Coot (29). Iterative refinement was carried out by alternating between manual re-building in Coot (29) and automated refinement in REFMAC (30). Similar procedures were adopted for the diffraction data collected from the complex MBHA crystal. In the later case, the structure obtained from the apo MBHA protein was used as the structural probe for molecular replacement.

All final models exhibit clear electron density for all residues (2-133 and 2-267 for PFA and MBHA, respectively). Note that the first residue (M1) was completely removed by the E.coli N-terminal methionine aminopeptidase during protein expression (verified by NMR and mass spectrometry). The final apo PFA structure is well defined to 1.70 Å resolution with an R-factor of 17.6% and a free R of 21.5%. 98.9% and 100% of all residues are located in the favored and allowed regions of the Ramachandran plot, respectively, and no residues are in the disallowed region as evaluated by MOLPROBITY (31). Similarly, the final apo MBHA structure was well defined to 1.60 Å resolution with an R-factor of 17.8% and a free R of 21.2%. 97.4% and 100% of all residues are located in the favored and allowed regions of the Ramachandran plot, respectively, and no residues are in the disallowed region as evaluated by MOLPROBITY (31). In the MBHA-\(\alpha_3\alpha_6\)-mannopentaose complex crystals, additional electron density was present only in the two carbohydrate binding sites of the second barrel of MBHA. The extra density in both sites permitted fitting of a \(\alpha_3\alpha_6\)-mannopentaose molecule into each binding site. The final complex structure is well refined with an R-factor of 18.6% and a free R of 22.4%. 98.5%, and 100% of all residues lie in the favored and allowed regions of the Ramachandran plot, respectively. A summary of the data collection parameters as well as pertinent structural statistics for all structures is provided in Table 1. All structural figures were generated with Chimera (32) or PyMOL (33). The atomic coordinates and diffraction data for apo PFA, apo MBHA, and the \(\alpha_3\alpha_6\)-mannopentaose-bound MBHA structures have been deposited in the RCSB Protein Data Bank under accession codes 4FBO, 4FBR, and 4FBV, respectively.

Carbohydrate binding studies by NMR Spectroscopy - 3D NMR HNCACB and CBCA(CO)NH spectra (34) were recorded for complete backbone chemical shift assignment of apo and \(\alpha_3\alpha_6\)-mannopentaose-bound PFA at 25°C on a Bruker AVANCE800 spectrometer, equipped with a 5-mm triple-resonance, z-axis gradient cryoprobe. For apo and \(\alpha_3\alpha_6\)-mannopentaose-bound MBHA, partial backbone chemical shift assignment were obtained using 3D NMR TROSY-HNCACB and TROSY-CBCA(CO)NH spectra on a Bruker AVANCE900 spectrometer, also equipped with a 5-mm triple-resonance, z-axis gradient cryoprobe. All protein spectra were recorded on a 13C/15N-labeled sample in 20 mM NaAcetate, 20 mM NaCl, 3 mM NaN\(_3\), 90/10% H\(_2\)O/D\(_2\)O (pH 5.0). The protein concentration (~1.0 mM) was similar in concentration to the one used for crystallization (~40 mg/mL). All spectra were processed with NMRPipe (35) and analyzed using NMRView (36).

Binding of \(\alpha_3\alpha_6\)-mannopentaose (Sigma Aldrich) was investigated at 25°C using 0.040 mM \(^{15}\text{N}\)-labeled PFA or MBHA or OPA in 20 mM NaAcetate, 20 mM NaCl, 3 mM NaN\(_3\), 90/10% H\(_2\)O/D\(_2\)O (pH 5.0) by \(^1\text{H}/^{15}\text{N}\) HSQC spectroscopy on a Bruker AVANCE600 or AVANCE900 spectrometer. 2D \(^1\text{H}/^{15}\text{N}\) HSQC spectra were recorded after each addition of carbohydrate with
final molar ratios of PFA:α3,α6-mannopentaose of 1:0, 1:0.25, 1:0.5, 1:1, 1:1.5, 1:2, and 1:3 (equivalent to 1:0, 1:0.125, 1:0.25, 1:0.5, 1:0.75, 1:1, and 1:1.5 for an individual glycan binding site), MBHA:α3,α6-mannopentaose of 1:0, 1:1, 1:2, 1:4, and 1:6 (equivalent to 1:0, 1:0.5, 1:1, and 1:1.5 for an individual glycan binding site), and OPA:α3,α6-mannopentaose of 1:0, 1:1, 1:2, 1:4 (equivalent to 1:0, 1:0.5, and 1:1 for an individual glycan binding site).

Anti-HIV assay - HIV viral activity assays were performed using TZM-bl (37) cells cultured in Dulbecco’s modified Eagle’s medium containing fetal bovine serum (10%), penicillin, and streptomycin. Cells were seeded at 10,000 per well in 100 µl volumes of media 16 h before inoculation. Wild type HIV-1 (1 ng of p24 per well), produced by transfection of 293T cells with the R9 proviral clone, was used for infection. This dose of virus was selected to keep the infection below saturation, ensuring linearity in the assay. 16 h after of seeding of TZM-bl cells, the infection below saturation, ensuring linearity in the assay. 16 h after of seeding of TZM-bl cells, the medium was aspirated, and 100 µl of medium containing virus and antiviral proteins was added to the cultures, supplemented with DEAE-dextran (20 µg/ml). All samples were carried out as duplicates. Cultures were incubated for 48 h at 37°C, after which 100 µl of Steady-Glo luciferase substrate (Promega) was added to each well. Plates were incubated at room temperature for 5 min, and luminescence was quantified by TopCount NXT™ (Packard Bioscience). The readout is relative luminescence units (RLU), which is directly proportional to infectivity. Values are expressed relative to the control cultures lacking antiviral protein. Antiviral potency was determined as the concentration of the inhibitor required for 50% inhibition of infection (IC50), and was extracted manually based on the nonlinear regression algorithm produced in KaleidaGraph.

RESULT AND DISCUSSION

Structures of OAAHs: PFA and MBHA - The amino acid sequences of the P. fluorescens and M. Xanthus lectins, designated as PFA and MBHA throughout the manuscript, possess extensive sequence similarity to OAA, with ~62% identity for pairwise alignment (Figure 1A). Interestingly, the majority of the amino acid conservation resides in the carbohydrate binding regions, delineated previously in OAA (Figure 1B). This region encompasses the loops between β1-β2, β7-β8, and β9-β10 in the first binding site and between β6-β7, β2-β3, and β4-β5 in the second binding site. In addition, notable sequence conservation is seen throughout the secondary structure elements, with residues Y4 and V6 in β1, W17, G20, G21, and W23 in β2, V34 and A35 in β3, G49, T50, M51, T52, and Y53 in β4, I59, G60, and F61 in β5, Y71, V73, E74, N75, and Q76 in β6, W84, G87, G88, and W90 in β7, A102 in β8, L114, G116, T117, T119, and Y120 in β9, and F128 in β10 being invariant in all three sequences (all numbering refers to OAA).

PFA assembles into a single, compact, β-barrel-like domain as previously observed for OAA (Figure 1C). Each sequence repeat folds up into five β-strands, denoted as β1 to β5 (colored in white) and β6 to β10 (colored in blue) for the first and second repeats of the first molecule in the asymmetric unit, respectively. For the second molecule in the asymmetric unit, the first and second five β-strands are colored in grey and light blue, respectively. A very short linker, comprising residues G67-N69 connects the two-sequence repeats (colored in orange) in both molecules. Note that the two individual PFA molecules in the asymmetric unit are essentially identical with root mean square deviations (r.m.s.d) for backbone and heavy atoms (residues S2 to I132) of 0.49 and 0.90 Å, respectively.

Different to OAA or PFA, MBHA contains four sequence repeats and in its crystal structure each two-sequence repeat folds into a β-barrel, resulting in a tandem arrangement of two barrels (Figure 1D). The first barrel is composed of the first ten β-strands, colored in white and purple, and the second barrel is made up by the second ten β-strands, colored in grey and light magenta, respectively. Three short linkers comprising residues S67-N69, T133-G135, and S201-D203 connect the first and second five β-strands, the first and second barrels, and the third and fourth five β-strands, respectively. The structures of the first (residues A2 to V132) and second (residues D136 to L266) barrels of MBHA are alike, with a
backbone atom r.m.s.d value of 0.63 Å. In comparing the relative orientations of the two molecules of PFA and the two tandem domains of MBHA it should be noted that they are quite different (Figure 1C and 1D), with an almost 180° flip of the second molecule/domain (right side in Figure 1C and 1D) relative to the first one (left side in Figure 1C and 1D).

Comparison of the overall structures of these new family members with that of the founding member, OAA, also revealed a close resemblance. The overall r.m.s.d values for backbone atoms between OAA (residues A2 to L132) and PFA (residues S2 to I132) is 0.50 Å (Figure 2A), between OAA (residues A2 to L132) and the first domain of MBHA (residues A2 to V132) it is 0.79 Å (Figure 2B), and between OAA (residues A2 to L132) and the second domain of MBHA (residues D136 to L266) it is 0.70 Å (Figure 2C). Similarly, PFA is also similar to each of the MBHA domains with backbone atom r.m.s.d. values of 0.74 Å between PFA and the first domain (residues S2 to I132 of PFA and residues A2 to V132 of MBHA) and 0.60 Å between PFA and the second domain of MBHA (residues S2 to I132 of PFA and residues D136 to L266 of MBHA). This extensive structural similarity parallels the significant degree of amino acid identity throughout the protein sequences (Figure 1A). It is worth pointing out, however, that some conformational differences are observed in the loop regions connecting β1-β2, β5-β6, and β6-β7 for OAA and PFA, and the loop regions connecting β1-β2, β4-β5, β5-β6, and β6-β7 for OAA and the first barrel of MBHA, and the loop regions connecting strands β1-β2, β2-β3, β4-β5, β5-β6, β6-β7 and β7-β8 for OAA and the second barrel of MBHA (Figure 2). Given that the amino acids around these loops are highly conserved (Figure 1A) and previous relaxation studies on OAA revealed that these regions are flexible, these minor conformational differences most likely arise from differences in crystal packing among these proteins.

**PFA and MBHA strongly and specifically bind to α3,α6-mannopentaose in solution** - OAA’s anti-HIV activity is associated with its binding to N-linked high mannose glycans on the viral envelope glycoprotein gp120 (23,38). The epitope that is recognized by OAA on Man-8/9 is α3,α6-mannopentaose (Manα(1-3)[Manα(1-3)[Manα(1-6)]Manα(1-6)] Man), the branched core unit of the triantennary high mannose structures (24). Note that the binding affinity for each binding site in OAA for this ligand is very similar (~<10 µM). In order to evaluate whether the same or different carbohydrate specificities are exhibited by the two new OAAH family members, we carried out titrations with α3,α6-mannopentaose and monitored the chemical shift perturbations by 1H-15N HSQC spectroscopy.

The 2D 1H-15N HSQC spectra of apo and glycan-bound PFA exhibit well-dispersed and narrow resonances, as expected for a well-folded protein (Figure 3A). All expected amide backbone resonances are observed and complete backbone assignments for free and bound PFA were obtained from 3D HNCACB and HN(CO)CACB spectra. Similar to OAA, α3,α6-mannopentaose binding to PFA is in slow exchange on the NMR chemical shift scale, suggesting a relatively tight interaction.

Chemical shift differences between free PFA and sugar-bound resonances at saturation (protein:sugar molar ratio of 1:3) are displayed along the polypeptide chain in Figure 3B. As can be appreciated, the most strongly affected residues are located in the loops, connecting strands β1-β2 and β9-β10 (site 1 – black bars) and between strands β4-β5 and β6-β7 (site 2 – red bars), respectively. Smaller changes are seen for the loops connecting β7-β8 (site 1 – black bars) and β2-β3 (site 2 – red bars). Resonances with chemical shift differences of > 0.08 ppm (~> 1x average value) include residues N8-S14, W17-H18, I101, N118-Y120, E123-G124, I126-G127, and G130 in the first binding site and residues M51, Y53, E56-G57, I59-G60, Q76-A82, and W84-H85 in the second binding site. These amino acids are equivalent to those in OAA that were perturbed by addition of the same sugar (24). Therefore, the binding sites in both proteins are located in corresponding regions of the structures, consistent with the extensive sequence conservation in these sites.

Interestingly, our titration experiments revealed a clear difference in carbohydrate affinity between
the two binding sites in PFA. As can be observed in the titration data, for increasing protein-α3,α6-mannopentaose molar ratios, starting with 1:0 (free protein - black), via 1:0.25 (red), 1:0.5 (green), 1:1 (blue), 1:1.5 (yellow), 1:2 (magenta), to 1:3 (cyan), the amide resonances of residues Q9 and A15 (Figure 3C), residue I126 (Figure 3D), and residues G11 and G127 (Figure 3E) that are all located in binding site 1, are perturbed earlier than the corresponding resonances of residues in binding site 2 (residues Q76 and A82 (Figure 3C), residue I59 (Figure 3D), and residues G78 and G60 (Figure 3E)). However, both binding sites are completely saturated at a protein-to-α3,α6-mannopentaose molar ratio of 1:3, suggesting that the affinity for binding site 1 is ~1.5 fold tighter than that for binding site 2. This is different from our findings with OAA, where both binding sites exhibited essentially indistinguishable affinity (24). It may be tempting to speculate that the difference in affinity is related to the slight difference in conformation, in particular considering the loop regions connecting strands β1-β2 in site 1 and β6-β7 in site 2. However, given that these loop regions are flexible in the apo form in solution and influenced by crystal packing, we cannot with any certainty draw this conclusion. We believe that it is more likely that the amino acid differences between the two sites in PFA (24 differences in 65 amino acids), compared to that in OAA (14 differences in 65 amino acids) play a role. Note that the sugar titrations for both proteins were carried out with identical protein concentrations (40 µM), using the same buffer conditions, the same stock of glycan solution, the same spectrometer, and performed on the same day. Therefore, even small differences can be reliably quantified.

We also carried out carbohydrate binding studies by NMR titration experiments for MBHA. The 2D 1H-15N HSQC spectra of apo (black) and α3,α6-mannopentaose-bound (magenta) MBHA are well dispersed, indicative of a well-folded structure (Figure 4A). Since MBHA contains twice the number of amino acids, i.e. double the number of resonances, overlap of backbone amide resonances and Cα and Cβ chemical shift degeneracy permitted only partial backbone assignment using 3D TROSY-HNCACB and TROSY-HN(CO)CACB spectra. We were, however, successful in obtaining unambiguous assignments for the important loop regions, since enough differences in sequences are present. For example, the stretch S147Q148A149T150 (repeat 1 in domain 1) corresponds to T80S81A82P83 (repeat 2 in domain 1), S147A148A149P150 (repeat 3 in domain 2), and S214T215A216P217 (repeat 4 in domain 2). In addition, our previous experience with OAA’s and PFA’s carbohydrate binding also aided in the assignment process for this 28 kDa protein.

Like for OAA and PFA, binding between MBHA and α3,α6-mannopentaose is in slow exchange on the NMR chemical shift scale and complete saturation of all MBHA’s resonances was obtained at a protein:sugar molar ratio of 1:6 (Figure 4A), consistent with four sugar binding sites per polypeptide chain. 2D 1H-15N HSQC spectra of the protein with increasing amounts of α3,α6-mannopentaose at molar ratios of 1:0 (free protein - black), 1:1 (red), 1:2 (green), 1:4 (blue), and 1:6 (magenta) are displayed in Figure 4. Again, the titration experiments resulted in several interesting findings. First, as expected, MBHA contains four binding sites for α3,α6-mannopentaose, as evidenced by the chemical shift changes of the amide resonances of residues A15, A82, A149, and A216, that are located in sequence repeats 1, 2, 3, and 4, respectively (Figure 4A and 4B). Second, among the four binding sites, two sites exhibit tighter binding than the other two, since the A149 and A216 resonances are already perturbed at the first titration point (protein:sugar molar ratio of 1:1, red), while the A15 and A82 resonances are not yet affected (Figure 4B). More importantly, at a molar ratio of 1:2 (green), where each binding site should be half saturated if identical sugar affinities were present, higher intensity for bound resonances of residues A149 and A216 compared to those of A15 and A82 are observed, clearly suggesting tighter affinities for sites 3 and 4. Third, both higher affinity sites are located in the second barrel of MBHA (Figure 1D). Similar observations hold for the Y53, Y120, Y187, and Y254 (Figure 4C), S13, T80, S147, and S214 (Figure 4D) and G60, G127, G194, and G261 (Figure 4E) resonances. Resonances of Y187/Y254/G194 and S147/S214/G261 that belong to amino acids in sequence repeats 3 and 4,
respectively, were affected earlier in the titration than their counterparts in sequence repeats 1 and 2.

Therefore, our data show that of the four carbohydrate binding sites in MBHA, those in the second domain bind carbohydrate more tightly than those in the first domain. This was somewhat unexpected, given the slight differences in affinity for the sites in a single barrel OAAH lectin. However, since at most a factor of 1.5-2 is observed, this variation in affinities most likely just reflects minor variations in amino acid sequence.

Structural basis for carbohydrate specificity of PFA and MBHA - Given that both, PFA and MBHA, strongly and specifically bind to \(\alpha_3,\alpha_6\)-mannopentaose, we aimed to crystallize the PFA- and MBHA-\(\alpha_3,\alpha_6\)-mannopentaose complexes by soaking or co-crystallization. Our attempts to soak sugar into the apo crystals were unsuccessful, and even the addition of small amounts of glycan resulted in the immediate dissolution of the PFA or MBHA crystals. This can be explained by the tight packing of the proteins in the crystal lattice, occluding the sugar binding sites. We therefore pursued co-crystallization at protein:sugar molar ratios of 1:3 and 1:6, with protein concentrations of 40 mg/ml and 30 mg/ml for PFA and MBHA, respectively, in 400 different conditions.

Crystals of \(\alpha_3,\alpha_6\)-mannopentaose-bound MBHA were obtained in 0.05 M KH\(_2\)PO\(_4\) and 20 % w/v polyethylene glycol 3,350. The protein crystallized in space group \(P2_12_12_1\), with one protein molecule per asymmetric unit. The complex structure was solved by molecular replacement using the apo structure of MBHA, comprising residues 2 to 267 as the structural probe in PHASER (28). The final model was refined to 1.76 Å resolution with R=18.6 % and R\(_{free}\)= 21.4 %.

Interestingly, the structure of \(\alpha_3,\alpha_6\)-mannopentaose-bound MBHA revealed that only the two higher affinity sites, i.e binding sites 3 and 4 that are both located in the second domain of MBHA, are occupied by the glycan (Figure 5A). Clear electron density at two opposing ends of this second barrel allowed for an excellent fit of the atomic structures of two \(\alpha_3,\alpha_6\)-mannopentaose molecules, one per site. No equivalent density was observed in the other two binding sites in the first domain. Further inspection of the complex structure in the crystal revealed that the carbohydrate binding regions in the first barrel are close to the bound \(\alpha_3,\alpha_6\)-mannopentaose from two neighboring molecules (Figure 5B). Indeed, one mannose unit at the end of the site 3-bound \(\alpha_3,\alpha_6\)-mannopentaose is in close contact to residues in binding site 2 of a neighboring protein molecule, while the bound sugar in site 4 reaches over to site 1 of another neighboring molecule. Therefore, adjacent protein molecules in the crystal are bridged by the glycans that simultaneously contact sites 3 and 4 and 1 and 2, respectively, facilitating crystal formation. In solution, on the other hand, we know from our NMR titration data, that sites 1 and 2 can bind \(\alpha_3,\alpha_6\)-mannopentaose similar to binding sites 3 and 4 (Figure 4), albeit with slightly reduced affinity.

Not surprisingly, the structures of the apo- and \(\alpha_3,\alpha_6\)-mannopentaose-bound MBHA are very similar with backbone and heavy atom r.m.s.d. values of 0.57 and 0.93 Å, respectively (Figure 5C), with only a minor difference in the binding site 2 loops. Therefore, within the error of the coordinates, no significant difference in the conformation of three of the four binding sites between the free and carbohydrate-bound MBHA structure can be seen.

The structure of the MBHA-\(\alpha_3,\alpha_6\)-mannopentaose complex reveals specific contacts between MBHA side chains and the carbohydrate. As shown in Figures 5D and 5E, the aromatic side chain of W144 in site 3 and W211 in site 4 play critical roles, providing hydrophobic contacts for the pyranose ring of M3. In addition, several polar interactions are also observed. These include hydrogen bonds between the hydroxyl groups of the sugar and main chain amide groups, as well as several side chains. In binding site 3 (Figure 5E), hydrogen bonds are formed between the backbone amide of G145 and the C5 hydroxyl group of M5', the backbone amide of G146 and the C6 hydroxyl group of M5', and the backbone amide of G258 and the C5 hydroxyl group of M4'. Side chain interactions in binding site 3 include hydrogen bonds between the C4 hydroxyl group of M3 and
the side chain carboxyl group of E257 and between the C4 hydroxyl group of M3 and the terminal guanidinium group of R229. Equivalent hydrogen bonds are found in binding site 4 (Figure 5E), such as hydrogen bonds between the backbone amide of G212 and the C5 hydroxyl group of M5', the backbone amide of G213 and the C6 hydroxyl group of M5', and the backbone amide of G191 and the C5 hydroxyl group of M4', with side chain hydrogen bonding between the C4 hydroxyl group of M3 and the side chain carboxyl group of E190 and between the C4 hydroxyl group of M3 and the terminal guanidinium group of R162.

Given the high sequence conservation in the carbohydrate binding regions between MBHA and OAA (Figure 1A), it is not surprising that all specific contacts that are observed in the MBHA-α3,α6-mannopentaose complex are identical to those observed previously in the OAA-α3,α6-mannopentaose complex (24).

Unlike for OAA and MBHA, we were not successful in obtaining crystals of the PFA-α3,α6-mannopentaose complex. Nevertheless, given the high amino acid conservation in the carbohydrate binding loops among PFA, OAA, and MBHA (Figure 1A), the chemical shift perturbation profile of PFA upon α3,α6-mannopentaose titration (Figures 1B and 3B), and the identical interactions observed in the complexes of MBHA-α3,α6-mannopentaose and OAA-α3,α6-mannopentaose (Figure 5D and 5E), we are confident that a reliable model for PFA-α3,α6-mannopentaose complex can be derived. This is provided in Figure 6, and was obtained by exchanging OAA residues in the structure of the OAA-α3,α6-mannopentaose complex to the corresponding PFA amino acids. The residue variability in the hydrophobic core of both proteins is very minor, and only six residues are different, namely I25L, V38L, L47F, L92I, I103L, and M118N (Figure 6A). The remaining non-identical amino acids are located on the surface of the proteins, as illustrated in Figure 6B. Interestingly, all residues that play major roles in contacting the carbohydrate in OAA and MBHA are completely conserved in PFA. These include residues W10-G12, R95, and E123-G124 in binding site 1 (Figure 6C) and W77-G79, R28, and E56-G57 in binding site 2 (Figure 6D). Between OAA and PFA, almost all residues in the carbohydrate binding sites 1 and 2, encompassing loop regions between strands β1-β2, β7-β8, and β9-β10 (site 1) and between strands β6-β7, β2-β3, and β4-β5 (site 2) are also conserved. Even though the interaction model between PFA and α3,α6-mannopentaose could potentially be validated using a docking program, our manual model provides compelling evidence that the interaction between PFA and α3,α6-mannopentaose in all likelihood is the same as in the OAA-α3,α6-mannopentaose or the MBHA-α3,α6-mannopentaose complexes.

Anti-HIV activities of PFA and MBHA - OAA’s anti-HIV activity is mediated by the specific recognition of α3,α6-mannopentaose, the branched core unit of Man-8 and Man-9 (23,38), major high-mannose sugars on the HIV-1 envelope glycoprotein gp120 (39). Here, we show that both PFA and MBHA also interact with this glycan structure, suggesting possible similar activities for these lectins. We therefore tested these molecules in HIV assays, using monomeric P51G CV-N and OAA as controls.

The HIV assay data (Figure 7) clearly shows that PFA and MBHA possess anti-HIV activity. All three OAAHs display very similar IC₅₀ values between 12 ± 1.0 nM (OAA and MBHA) and 15 ± 1.0 nM (PFA), values approximately 30 fold higher than that obtained for CV-N at the same time (0.4 ± 0.1 nM).

If solely avidity considerations were significant, one would expect MBHA to exhibit higher anti-HIV activity than OAA and PFA, given that it possesses four sugar binding sites, compared to only two sites in latter lectins. We interpret the comparable activity for all OAAH family members such that only one or two binding sites can engage the sugars on gp120 and that potentially a single interaction is important for anti-HIV activity.

Importantly, CV-N is more active (~30 fold improved IC₅₀ value) than any of the three OAAHs. Considering that CV-N, OAA, and PFA share the same number of binding sites for carbohydrate and the dissociation constants of these lectins for the individual sugar ligands are
relatively similar (in the low micromolar range) (23,40), one would expect comparable inhibition with similar IC₅₀ values. This clearly is not the case. We therefore propose the following explanation: OAA, PFA, and MBHA recognize only a single epitope of Man-8/9, namely the branched core unit (24), while CV-N binds to Manα(1-2)Man units on the D1 or D3 arms (11,12), i.e. interacts with two alternative epitopes on the glycan. Therefore, multi-site and multi-epitope recognition seem essential for the most potent HIV inactivation (a schematic model for HIV inactivation is provided in Figure 9).

In addition to the native three OAAH family members, we also created a hybrid protein from OAA and PFA (named OPA), in which the two parental proteins are connected by a two-residue N-G linker between the C-terminus of OAA and the N-terminus of PFA. As shown in Figures 8A and 8B, the 2D ¹H-¹⁵N HSQC spectrum of OPA exhibits all the characteristics of a well-folded structure and is essentially identical to the sum of the ¹H-¹⁵N HSQC spectra of OAA (Figure 8A - green) and PFA (Figure 8B - magenta) with only minor perturbations involving the last C-terminal residues of OAA and first N-terminal residues of PFA (Figure 8C). Therefore, the overall structure of OPA is essentially a tandem, double barrel protein in which OAA and PFA are linked.

Interestingly, in the HIV assays OPA inhibited HIV replication slightly better than OAA or PFA, displaying an IC₅₀ of 7.5 ± 0.5 nM (Figure 7). Although this difference is small, it was nevertheless surprising, since only very minor conformational changes from the original OAA and PFA structures were expected in the hybrid OPA. We therefore tested whether, fortuitously, the apparent glycan affinity was altered.

The data in Figure 8D shows that OPA specifically interacts with α₃,α₆-mannopentaose and that binding is in slow exchange on the NMR chemical shift scale. OPA’s resonances were completely saturated at a protein:sugar molar ratio of 1:4 (colored in blue - Figure 8D), indicating that the four binding sites in OPA are competent for sugar binding. For OAA, PFA, and MBHA saturation was reached at a protein:sugar molar ratio of 1:3 (24), 1:3 (Figures 3C – 3E), and 1:6 (Figures 4B – 4E), respectively, suggesting that each binding site in OPA exhibits a slightly higher affinity for the sugar than observed in the parental OAA and PFA lectins or the double domain MBHA protein.

In addition to the delineation of the four binding sites in OPA (see the affected resonances of A15 and A82, Figure 8E and G60 and G127, Figure 8F) derived from OAA and PFA, respectively (colored in cyan and magenta), each binding site in OPA exhibits essentially identical affinity towards α₃,α₆-mannopentaose, with resonances of all interacting residues equally perturbed at the 1:2 molar ratio, where each binding site should be half saturated for identical sugar affinities (Figure 8E and 8F - green). This is different from PFA where slightly different affinities are observed for the two binding sites (Figures 3C – 3E). The equal affinity for each binding site in the hybrid protein, may be the result of a minor structural change caused by the linker, or a small positive avidity effect. This slightly tighter glycan binding may contribute to the somewhat higher apparent anti-HIV potency noted in the HIV assays. It is also worth pointing out that the α₃,α₆-mannopentaose-bound resonances for equivalent residues in OPA, OAA and PFA exhibit identical frequencies (Figure 8G and 8H; OPA, OAA, and PFA are colored blue, cyan, and magenta, respectively), suggesting very similar conformations of OPA and its parental proteins in the bound state.

CONCLUSION

Crystal structures for two members of the OAAH lectin family, PFA and MBHA, were determined at resolutions of 1.70 Å and 1.60 Å, respectively, as well as for the α₃,α₆-mannopentaose-bound MBHA at 1.76 Å. Structural comparison between these two new members and OAA, the founding member of the OAAH lectin family, revealed a very similar overall fold, with only minor conformational differences in the loop regions, connecting strands β₁-β₂, β₂-β₃, β₄-β₅, β₅-β₆, β₆-β₇ and β₇-β₈, that are caused by different crystal packing in these proteins. These new members bind tightly and specifically to α₃,α₆-mannopentaose, exhibiting identical intermolecular interactions to those previously observed in the OAA-α₃,α₆-mannopentaose...
complex structure. HIV assays revealed that all OAAH proteins and the designer hybrid OAAH (OPA) display anti-HIV activity and all are slightly less potent than CV-N. Our results provide further details towards a structural understanding of the protein-carbohydrate interaction in this novel lectin family, as well as insights into the molecular basis of its HIV inactivation properties. Such data can potentially be used in the development of lectins as antiviral reagents that target the gp120 glycans in the quest to combat HIV transmission.

REFERENCES


FOOTNOTES

The authors thank Dr. William Furey for helpful discussion and Mike Delk for NMR technical support. This work was supported by an NIH grant to A.M.G (GM080642). The following reagent was obtained through the NIH AIDS Research and Reference Reagent Program, Division of AIDS, NIAID, NIH: TZM-bl from Dr. John C. Kappes, Dr. Xiaoyun Wu and Tranzyme Inc.
Figure 1: Sequence, carbohydrate binding, and structures of OAAHs. (A) Amino acid sequence alignment of OAA, PFA, and MBHA. The four sequence repeats in MBHA are displayed as two separate two-sequence stretches. Conserved and similar residues are colored in blue and slate, respectively. The two-residue linker connecting the two domains of MBHA is colored in orange. (B) Secondary structure and chemical shift perturbation profile (combined amide chemical shift changes) of OAA at the final titration point (1:3 molar ratio of OAA to α3, α6-mannopentaose). Values were calculated using the equation; Δppm = [(Δppm (1H_N))² + (Δppm (15N) / 5)²]¹/². (C) and (D) Ribbon representation of the backbone structures of PFA and MBHA, respectively. The five β-strands in the first and second sequence repeats in PFA are colored in white and blue and in grey and light blue for molecule one and two, respectively, in the asymmetric unit. For MBHA, the five β-strands in the first and second sequence repeats in the first and second domains are colored in white and purple and in grey and light magenta, respectively. The connecting residues between the repeats (between strands β5 and β6) in both proteins and between the first and second domains in MBHA are colored in orange.

Figure S2: Conservation of the OAAH fold. (A), (B), and (C) Stereoviews of best-fit Cα superpositions of OAA and PFA, OAA and the first domain of MBHA, and OAA and the second domain of MBHA, respectively. OAA, PFA, and the first and second domains of MBHA are colored in white, blue, purple, and light magenta, respectively.

Figure 3: NMR titrations of PFA with α3, α6-mannopentaose. (A) Superposition of the 2D 1H-15N HSQC spectra of free (black) and α3, α6-mannopentaose-bound (cyan) PFA at 1:3 molar ratio of PFA to sugar. Selected resonances that exhibit large chemical shift changes are labeled by amino acid type and number. Note that the amide resonances of G26 and G93 are significantly upfield shifted in their proton frequencies and not within the current spectral boundaries. (B) Secondary structure and chemical shift perturbation profile of combined amide chemical shift changes for PFA at the final titration point (PFA:α3, α6-mannopentaose = 1:3). Values were calculated using the equation; Δppm = [(Δppm (1H_N))² + (Δppm (15N) / 5)²]¹/². (C), (D) and (E) Superposition of selected regions in the 2D 1H-15N HSQC spectra of free (black) and α3, α6-mannopentaose bound PFA at 1:0.25 (red), 1:0.50 (green), 1:1 (blue), 1:1.5 (yellow), 1:2 (magenta), and at 1:3 (cyan) molar ratios of PFA to sugar for residues A15/A82 and Q9/Q76 (A), residues I59/I126 (D), and residues G60/G127 and G11/G78. All spectra are plotted at the same contour level and all titrations were carried out using 0.040 mM PFA in 20 mM NaAcetate, 20 mM NaCl, 3 mM NaN3, 90/10% H2O/D2O (pH 5.0), 25°C.

Figure 4: NMR titrations of MBHA with α3, α6-mannopentaose. (A) Superposition of the 2D 1H-15N HSQC spectra of free (black) and α3, α6-mannopentaose-bound MBHA at 1:6 molar ratio (magenta) of MBHA to α3, α6-mannopentaose. The four amide resonances of MBHA that correspond to G26 and G93 in OAA or PFA are displayed in the bottom inset. Three selected regions that are discussed in the main text are indicated by dashed boxes. (B), (C), (D), and (E) Superposition of selected regions of the 2D 1H-15N HSQC spectra of free (black) and
\( \alpha_3, \alpha_6 \)-mannopentaose bound MBHA at 1:1 (red), 1:2 (green), 1:4 (blue), and at 1:6 (magenta) molar ratios of MBHA to sugar for residues A15/A82/A149/A216 (B), residues Y53/Y120/Y187/Y257 (C), residues S13/T80/S147/S214 (D), and residues G60/G127/G194/G261 (E). All spectra are plotted at the same contour level and all titrations were carried out using 0.100 mM MBHA in 20 mM NaAcetate, 20 mM NaCl, 3 mM NaN₃, 90/10% H₂O/D₂O (pH 5.0), 25°C.

**Figure 5:** Crystal structures of \( \alpha_3, \alpha_6 \)-mannopentaose-bound MBHA. (A) Ribbon representation of \( \alpha_3, \alpha_6 \)-mannopentaose-bound MBHA structure. The five \( \beta \)-strands from the first and second sequence repeat in the first and second domain of MBHA are colored in white and purple, and in grey and light magenta, respectively. The connecting amino acids between the sequence repeats (between strands \( \beta 5 \) and \( \beta 6 \) and strands \( \beta 15 \) and \( \beta 16 \)) and between the first and second domain in MBHA are colored orange. The two bound \( \alpha_3, \alpha_6 \)-mannopentaose molecules in binding sites three and four, located in the second domain, are depicted in green stick representation, with oxygen atoms shown in red. The electron density contoured at 1.0\( \sigma \) enclosing the molecular model of \( \alpha_3, \alpha_6 \)-mannopentaose (or Man\( \alpha \)(1-3)[Man\( \alpha \)(1-3)]Man\( \alpha \)(1-6)]Man\( \alpha \)(1-6)]Man pentasaccharide) in stick representation in binding sites 3 and 4 of MBHA is colored blue. (B) Arrangement of \( \alpha_3, \alpha_6 \)-mannopentaose-bound MBHA depicting two neighboring molecules in the crystal. The color scheme and secondary structure depiction is similar to that in (A). (C) Superposition of ligand-free and \( \alpha_3, \alpha_6 \)-mannopentaose-bound MBHA structures in ribbon representation. The bound \( \alpha_3, \alpha_6 \)-mannopentaose binding sites 3 and 4 are depicted in cyan and light green stick representation, with oxygen atoms shown in red. (D) and (E) Details of the interaction network between \( \alpha_3, \alpha_6 \)-mannopentaose and MBHA in binding sites 3 and 4. The protein is depicted in both surface and stick representations, with residues that are in direct contact with the carbohydrate are the following: W144, G145, G146, R229, E257, and G258 in binding site 3 (D) and W211, G212, G213, R162, E190, and G191 in binding site 4 (E). Inter-molecular hydrogen bonds are indicated by black dashed lines. The same color scheme as in (C) is used for the carbohydrate. Protein residues are labeled by single-letter code and the sugar rings of the carbohydrate are labeled according to standard nomenclature.

**Figure 6:** Structural model of \( \alpha_3, \alpha_6 \)-mannopentaose-bound PFA based on the OAA-\( \alpha_3, \alpha_6 \)-mannopentaose complex. (A) and (B) Stereoviews of PFA’s carbohydrate binding sites 1 and 2, highlighting amino acid differences between OAA and PFA in the hydrophobic (A) and solvent exposed regions (B). (C) and (D) Predicted inter-molecular hydrogen bonds between PFA and \( \alpha_3, \alpha_6 \)-mannopentaose (black dashed lines) in binding sites 1 (C) and 2 (D). Residues that are in direct contact with the carbohydrate are identical in PFA and OAA: W10, G11, G12, R95, E123, and G124 in binding site 1 (C) and W77, G78, G78, R28, E56, and G57 in binding site 2 (D). The protein is depicted in both ribbon and stick representation, with bound \( \alpha_3, \alpha_6 \)-mannopentaose in stick representation only. All carbon, oxygen, and nitrogen atoms of the conserved residues in PFA and OAA and in the carbohydrates are shown in white, red, and blue, respectively. The carbon atoms of differing residues in PFA and OAA are colored in cyan. Amino acids are labeled by single-letter code and the sugar rings of the carbohydrate are labeled according to standard nomenclature.
Figure 7: Anti-HIV activity assays for OAAHs. Single-cycle HIV-1 infectivity assays were performed using a luciferase reporter cell line to quantify the antiviral activity of the lectins. Experimental data and the lines for CV-N (square), OAA (circle), PFA (diamond), MBHA (up triangle), and OPA (down triangle) are colored in black, blue, green, red, and orange, respectively. Each data point was obtained from 6 duplicate measurements. The dashed black lines were used for extracting the IC_{50} concentration.

Figure 8: Comparison of the structures and NMR titrations of OPA with α3,α6-mannopentaose by $^1$H-$^{15}$N HSQC spectroscopy. (A) and (B) Superposition of the 2D $^1$H-$^{15}$N HSQC spectra of OPA (black) and OAA (green) (A) and of OPA (black) and PFA (magenta) (B). Residues of OAA and PFA that exhibit small chemical shift changes are labeled by amino acid type and number. The four amide resonances of the OPA glycines that correspond to G26 and G93 in OAA or PFA are shown in the bottom inset. (C) Structural mapping of OAA and PFA residues that exhibit chemical shift differences between the two proteins. The C-terminal residue of OAA and the N-terminal residue of PFA are colored in red and blue, respectively. The short two-residue linker connecting OAA and PFA is depicted by a dashed orange line. (D) Superposition of the 2D $^1$H-$^{15}$N HSQC spectra of free (black) and α3,α6-mannopentaose bound (blue) OPA at molar ratio of OPA to α3,α6-mannopentaose of 1:4. (E) and (F) Superposition of selected regions of the 2D $^1$H-$^{15}$N HSQC spectra of free (black) and α3, α6-mannopentaose bound OPA at 1:1 (red), 1:2 (green), and 1:4 (blue) molar ratios of OPA to sugar for residues A15/A82 (E; cyan) and residues G60/G127 (F; magenta). (G) and (H) Superposition of selected regions of the 2D $^1$H-$^{15}$N HSQC spectra of OAA (cyan), PFA (magenta) and OPA (blue) at the final titrations point. All spectra are plotted at the same contour level and all titrations were carried out using 0.100 mM OPA in 20 mM NaAcetate, 20 mM NaCl, 3 mM NaN₃, 90/10% H₂O/D₂O (pH 5.0), 25°C.

Figure 9: Schematic depiction of HIV inactivation by CVNH and OAAH lectins. In the absence of antiviral lectins, the interaction between gp120 and CD4 introduces a conformational change that allows the fusion peptide of gp41 to penetrate the cell membrane, leading to viral-cell membrane fusion and HIV capsid deposition into the cell (A). In the presence of antiviral lectins they bind to the high-mannose glycans on gp120/41, preventing the required conformational change, thereby blocking infection (B). The glycan epitopes of Man-9 that are recognized by CV-N (highlighted in magenta) and OAA (highlighted in blue) as well as their detailed interactions at the atomic level are provided at the right hand side.
Figure 1, Koharudin et al.
Figure 2, Koharudin et al.
Figure 3, Koharudin et al.
Figure 4, Koharudin et al.
Figure 5, Koharudin et al.
Figure 6, Koharudin et al.
Figure 7, Koharudin et al.
Figure 9, Koharudin et al.
Table 1: PFA and MBHA data collection, refinement and Ramachandran statistics.

<table>
<thead>
<tr>
<th></th>
<th>apo PFA&lt;sup&gt;a,b,c&lt;/sup&gt;</th>
<th>apo MBHA&lt;sup&gt;a,b,c&lt;/sup&gt;</th>
<th>MBHA complex&lt;sup&gt;c&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Data collection</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Space group</td>
<td>C2</td>
<td>P&lt;sub&gt;3&lt;/sub&gt;1</td>
<td>P&lt;sub&gt;2&lt;/sub&gt;1/2/1</td>
</tr>
<tr>
<td><strong>Cell dimensions</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a, b, c (Å)</td>
<td>58.95 / 70.31 / 62.29</td>
<td>76.10 / 76.10 / 37.67</td>
<td>45.59 / 57.32 / 105.97</td>
</tr>
<tr>
<td>α, β, γ (°)</td>
<td>90 / 111.73 / 90</td>
<td>90 / 90 / 120</td>
<td>90 / 90 / 90</td>
</tr>
<tr>
<td>Wavelength (Å)</td>
<td>1.5418</td>
<td>1.5418</td>
<td>1.5418</td>
</tr>
<tr>
<td>Resolution (Å)</td>
<td>28.14 - 1.70 (1.76 – 1.70)&lt;sup&gt;d&lt;/sup&gt;</td>
<td>32.95 - 1.60 (1.66 – 1.60)&lt;sup&gt;d&lt;/sup&gt;</td>
<td>33.82 - 1.76 (1.82 – 1.76)&lt;sup&gt;d&lt;/sup&gt;</td>
</tr>
<tr>
<td>R&lt;sub&gt;merge&lt;/sub&gt;</td>
<td>0.038 (0.294)&lt;sup&gt;d&lt;/sup&gt;</td>
<td>0.116 (0.430)&lt;sup&gt;d&lt;/sup&gt;</td>
<td>0.058 (0.206)&lt;sup&gt;d&lt;/sup&gt;</td>
</tr>
<tr>
<td>R&lt;sub&gt;meas&lt;/sub&gt;</td>
<td>0.052 (0.403)&lt;sup&gt;d&lt;/sup&gt;</td>
<td>0.124 (0.491)&lt;sup&gt;d&lt;/sup&gt;</td>
<td>0.057 (0.253)&lt;sup&gt;d&lt;/sup&gt;</td>
</tr>
<tr>
<td>I / σI</td>
<td>12.6 (2.4)&lt;sup&gt;d&lt;/sup&gt;</td>
<td>10.6 (1.7)&lt;sup&gt;d&lt;/sup&gt;</td>
<td>22.1 (3.8)&lt;sup&gt;d&lt;/sup&gt;</td>
</tr>
<tr>
<td>Completeness (%)</td>
<td>93.8 (90.6)&lt;sup&gt;d&lt;/sup&gt;</td>
<td>97.0 (88.8)&lt;sup&gt;d&lt;/sup&gt;</td>
<td>96.8 (81.5)&lt;sup&gt;d&lt;/sup&gt;</td>
</tr>
<tr>
<td>Redundancy</td>
<td>1.91 (1.90)&lt;sup&gt;d&lt;/sup&gt;</td>
<td>6.49 (3.65)&lt;sup&gt;d&lt;/sup&gt;</td>
<td>6.63 (2.53)&lt;sup&gt;d&lt;/sup&gt;</td>
</tr>
<tr>
<td><strong>Refinement</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Resolution (Å)</td>
<td>28.14 – 1.70 (1.74 – 1.70)&lt;sup&gt;d&lt;/sup&gt;</td>
<td>32.95 – 1.60 (1.64 – 1.60)&lt;sup&gt;d&lt;/sup&gt;</td>
<td>33.82 – 1.76 (1.81 – 1.76)&lt;sup&gt;d&lt;/sup&gt;</td>
</tr>
<tr>
<td>No. reflections</td>
<td>22,452</td>
<td>28,106</td>
<td>25,993</td>
</tr>
<tr>
<td>R&lt;sub&gt;work&lt;/sub&gt; / R&lt;sub&gt;free&lt;/sub&gt;</td>
<td>0.176 / 0.215 (0.276 / 0.327)&lt;sup&gt;d&lt;/sup&gt;</td>
<td>0.178 / 0.212 (0.240 / 0.287)&lt;sup&gt;d&lt;/sup&gt;</td>
<td>0.186 / 0.214 (0.224 / 0.315)&lt;sup&gt;d&lt;/sup&gt;</td>
</tr>
<tr>
<td>No. atoms</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Protein</td>
<td>1,970</td>
<td>1,990</td>
<td>1,998</td>
</tr>
<tr>
<td>Ligand/ion</td>
<td>-</td>
<td>-</td>
<td>128</td>
</tr>
<tr>
<td>Water</td>
<td>312</td>
<td>243</td>
<td>197</td>
</tr>
<tr>
<td>B-factors</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Protein</td>
<td>37.61</td>
<td>28.96</td>
<td>23.30</td>
</tr>
<tr>
<td>Ligand/ion</td>
<td>-</td>
<td>-</td>
<td>28.45</td>
</tr>
<tr>
<td>Water</td>
<td>45.66</td>
<td>36.95</td>
<td>31.97</td>
</tr>
<tr>
<td>R.m.s deviations</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bond lengths (Å)</td>
<td>0.013</td>
<td>0.013</td>
<td>0.012</td>
</tr>
<tr>
<td>Bond angles (°)</td>
<td>1.343</td>
<td>1.396</td>
<td>1.400</td>
</tr>
<tr>
<td><strong>Ramachandran</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>favored region</td>
<td>98.9</td>
<td>97.4</td>
<td>98.5</td>
</tr>
<tr>
<td>allowed region</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
</tr>
</tbody>
</table>

<sup>a</sup>Data were obtained from the best diffracting crystal according to crystallization condition (see text for details).
<sup>b</sup>Crystallographic phases were obtained by molecular replacement using the previously determined structure of OAA in orthorhombic P<sub>2</sub>1<sub>1</sub>2<sub>1</sub>2<sub>1</sub> space group (PDB:3S5V).
<sup>c</sup>Refinement was carried out using Refmac5.
<sup>d</sup>Values in parentheses are for highest-resolution shell.
<sup>e</sup>Values are obtained by MOLPROBITY (31).