Polysaccharide Recognition by Surfactant Protein D: Novel Interactions of a C-Type Lectin with Nonterminal Glucosyl Residues†

Martin J. Allen,*,‡ Alain Laederach,§,|| Peter J. Reilly,§ and Robert J. Mason‡,⊥

Department of Medicine, National Jewish Medical and Research Center, 1400 Jackson Street, Denver, Colorado 80206-2761, Department of Chemical Engineering, 2114 Sweeney Hall, Iowa State University, Ames, Iowa 50011-2230, Bioinformatics and Computational Biology Program, 2014 Molecular Biology Building, Iowa State University, Ames, Iowa 50011-3260, and Department of Medicine, University of Colorado Health Sciences Center, 4200 East Ninth Street, Denver, Colorado 80262

Received December 21, 2000; Revised Manuscript Received May 7, 2001

ABSTRACT: Surfactant protein D (SP-D), a C-type lectin, is an important pulmonary host defense molecule. Carbohydrate binding is critical to its host defense properties, but the precise polysaccharide structures recognized by the protein are unknown. SP-D binding to *Aspergillus fumigatus* is strongly inhibited by a soluble β-(1→6)-linked but not by a soluble β-(1→3)-linked glucosyl homopolysaccharide (pustulan and laminarin, respectively), suggesting that SP-D recognizes only certain polysaccharide configurations, likely through differential binding to nonterminal glucosyl residues. In this study we have computationally docked α/β-D-glucopyranose and α/β-(1→2)-, α/β-(1→3)-, α/β-(1→4)-, and α/β-(1→6)-linked glucosyl trisaccharides into the SP-D carbohydrate recognition domain. As with the mannose-binding proteins, we found significant hydrogen bonding between the protein and the vicinal, equatorial OH groups at the 3 and 4 positions on the sugar ring. Our docking studies predict that α/β-(1→2)-, α-(1→4)-, and α/β-(1→6)-linked but not α/β-(1→3)-linked glucosyl trisaccharides can be bound by their internal glucosyl residues and that binding also occurs through interactions of the protein with the 2- and 3-equatorial OH groups on the glucosyl ring. By using various soluble glucosyl homopolysaccharides as inhibitors of SP-D carbohydrate binding, we confirmed the interactions predicted by our modeling studies. Given the sequence and structural similarity between SP-D and other C-type lectins, many of the predicted interactions should be applicable to this protein family.

Pulmonary SP-D[1] is a member of the collectin family of C-type lectins. This protein binds carbohydrates in a calcium-dependent manner and shows monosaccharide specificity in the order glucose/mannose > galactose. SP-D monomers are 355 amino acids (aa) in length and contain a short (25 aa) N-terminal region involved in interchain disulfide bonding, a collagen-like domain (177 aa), a neck region (38 aa), and a C-terminal CRD (115 aa) (*1*). SP-D monomers oligomerize through trimeric intermediates to form a cruciform-like dodecamer (*1*).

SP-D binds many microorganisms in vitro (*2*−*6*) and is thought to be an important component of the innate immune system. It is thought that SP-D host defense functions are mediated through recognition of specific carbohydrates on the surface of invading microorganisms, but the precise carbohydrate structures recognized by the protein are not well understood. Although monosaccharide specificity for SP-D has been studied (*7*), the specificity of the protein for long-chain polysaccharides has not. Since diverse long-chain carbohydrates are present on the surface of numerous microorganisms, knowledge of the mechanisms governing polysaccharide recognition by SP-D is crucial to our understanding of the host defense functions of this protein.

In addition to SP-D, other C-type lectins include SP-A, MBP-A, MBP-C, and the selectins. Like SP-D, the monosaccharide recognition specificities for SP-A (*8*, *9*) and MBP (*10*, *11*) have been examined extensively. Several MBP−carbohydrate complex structures, including complexes of MBP with several monosaccharides (*11*) and in one case with a polysaccharide (*12*), have been determined by X-ray crystallography. These studies have shown that MBP specifically recognizes vicinal, equatorial OH groups on monosaccharides equivalent to those present at the 3 and 4 positions of sugars such as mannose and glucose (*13*). It appears that MBP recognizes similar OH groups present on the nonreducing terminal carbohydrate residues of polysaccharides (*12*). Given the sequence and structural similarities between

† This work was supported by grants from NIH (HL-29891) and EPA (R825702). M.J.A. was funded by Great West Life Assurance and Andrew Goodman Fellowships at the National Jewish Medical and Research Center, and A.L. was supported by National Science Foundation IGERT Program Award 9972653.

* Corresponding author. Telephone: +1-303-398-1226. Fax: +1-303-398-1806. E-mail: allenm@njc.org.

‡ National Jewish Medical and Research Center.

§ Department of Chemical Engineering, Iowa State University.

|| Bioinformatics and Computational Biology Program, Iowa State University.

⊥ University of Colorado Health Sciences Center.

[1] Abbreviations: aa, amino acids; CRD, carbohydrate recognition domain; CMC, carboxymethylcellulose; ECL, extracyclic carbon left; ECR, extracyclic carbon right; FITC, fluorescein isothiocyanate; hSP-D, human surfactant protein D; IgG, immunoglobulin G; LGA, Lamarkian genetic algorithm; MPB-A, mannose-binding protein A; MBP-C, mannose-binding protein C; α-Me-GlcNAc, α-methyl-*N*-acetylglucosamine; PDB, Protein Data Bank; SP-A, surfactant protein A; SP-D, surfactant protein D; rmsd, root-mean-square deviation.

10.1021/bi002901q CCC: $20.00
Published on Web 06/08/2001

the CRDs from SP-D and MBP, mutagenesis data (*14*, *15*), and studies using various sugars as competitive inhibitors of carbohydrate recognition (*7*, *10*), it seems reasonable to conclude that SP-D recognizes carbohydrates by a mechanism similar to that of MBP.

We have recently demonstrated that pustulan [a soluble β-(1→6)-linked glucosyl homopolysaccharide] but not laminarin [a soluble β-(1→3)-linked glucosyl homopolysaccharide] inhibits SP-D binding to *Aspergillus fumigatus* and aggregation of *Saccharomyces cerevisiae* (*16*). Since the nonreducing terminal positions of these polysaccharides are identical, we concluded that the observed specificity for pustulan is a result of differential internal glucosyl residue recognition by SP-D. Recognition of internal carbohydrate residues has not been previously suggested for C-type lectins.

The three-dimensional structure of hSP-D was recently reported (*17*). To more completely define polysaccharide recognition by SP-D, we used this structure and automated computational docking to model SP-D interactions with glucose and several glucosyl trisaccharides. We were specifically interested in examining nonterminal glucosyl residue recognition. In addition, several soluble glucosyl polysaccharides were tested for their ability to inhibit carbohydrate binding by SP-D. These studies greatly extend our knowledge of carbohydrate recognition by SP-D and are likely applicable to other C-type lectins.

EXPERIMENTAL PROCEDURES

Materials. Dextran (average molecular weight 74 200) from Sigma (St. Louis, MO), CMC (average molecular weight 250 000 and ds 0.7) from Aldrich (Milwaukee, WI), soluble starch from Fisher (Pittsburgh, PA), and the FITC-conjugated $F(ab')_2$ fragment of donkey anti-rabbit IgG from Jackson ImmunoResearch (West Grove, PA) were used. *A. fumigatus* conidia, recombinant hSP-D expressed in CHO K1 cells, and polyclonal rabbit anti-hSP-D antibody were prepared exactly as described previously (*2*).

Computational Methods. Automated docking simulations were conducted with the AutoDock 3.06 suite of programs (Scripps Research Institute, La Jolla, CA) (*18*, *19*), using chain A of the hSP-D structure (PDB accession code 1b08) (*17*) as the model for the CRD. All water and heteroatoms were removed except for the three calcium ions. Hydrogen atoms were added to the crystal structure using the HBUILD command in CHARMm (Harvard University, Cambridge, MA) (*20*). Atomic partial charges were assigned to the protein atoms according to Cornell et al. (*21*). Intermolecular interaction energy grids were calculated using atomic probes corresponding to each atom type found in the ligand. Grid spacing was set to 0.375 Å, with 70 grid points centered on calcium ion 1. The electrostatic interaction energy grid used a sigmoidal distance-dependent dielectric function (*22*) to account for the solvent-screening effect. Self-consistent 12-6 Lennard-Jones coefficients (*19*) were used along with a distance criterion with sinusoidal directional attenuation to account for hydrogen bonding (*21*). Lennard-Jones parameters for calcium were taken from Åqvist (*23*). Internal energies (electrostatic and van der Waals) of the ligand were computed according to Glycam_93 (*24*). Glycosidic bond torsional constraints were fitted to Gaussian distributions for ease of implementation in AutoDock.

The α-carbon backbones of MBP-A and SP-D CRDs were superimposed using SwissPDB viewer (*25*). The relative coordinates of the six ring atoms of the terminal mannose (Man9 in PDB accession code 2msb) (*12*) bound to MBP were used to place the structures near the binding site as previously described (*26*). Three docking runs were carried out for each of the trisaccharides docked to SP-D. The reducing, internal, and nonreducing rings were each initially placed near the SP-D carbohydrate-binding pocket. All docking simulations were carried out using the LGA in AutoDock. For the monosaccharides, 1000 LGA runs were executed, with a population size of 50, a mutation rate of 0.2, and a crossover rate of 0.8. For trisaccharides, since three initial conformations were used, a total of 3000 LGA runs were executed with the same population size and mutation/crossover rates as for the monosaccharides.

The use of the LGA in combination with a pseudo Solis and Wets local minimizer (*18*, *19*) provides a thorough search of the conformational space around the global energy minimum. All output structures are local minima of the energy potential. For trisaccharides capable of docking via internal units, three structurally different clusters of local minima exist, in which either the reducing, nonreducing, or internal sugar unit is bound to the putative hSP-D carbohydrate-binding pocket. These clusters can be analyzed independently, since they represent different binding modes. The use of a 5 kcal/mol energy cutoff (see below) ensures that all of the accepted structures have a favorable binding energy. Therefore, any saccharide conformation with a final docked energy below the cutoff is considered a productive binding mode. For the case of trisaccharides, if there are major steric clashes with adjacent glucosyl units, the binding energy will not be favorable, and the energy cutoff will discriminate against these structures.

We were interested in determining binding conformations where the glucosyl residue interacted with SP-D via a mechanism similar to that observed in MBP. Since this interaction is mediated through only four hydrogen bonds between the ligand and the protein (*11*, *12*), many nonspecific binding modes are possible with only slightly less favorable energies. Thus we were unable to establish a consistent energy cutoff that alone would differentiate between specific and nonspecific binding.

Water molecules occupy putative sugar ligand OH positions in the SP-D structure (*17*). When the structures of wild-type MBP−ligand complexes (PDB accession codes 2msb and 1rdi−1rdn) (*11*, *12*) are superimposed on the SP-D chain A structure, the greatest sum distance separating the ligand oxygens from the water oxygen atoms (superimposition of 1b08 chain A with 1rdj chain 2) is 1.38 Å. Accordingly, all ligand structures with two hydroxyl oxygen atoms within 1.38 Å of the water oxygen atoms in hSP-D were considered specifically bound. We also considered using a separate distance criterion for each of the two ligand OH groups. However, the results of that analysis were not consistent with experimental data and were therefore judged unsatisfactory. Terminal ring binding was observed for all the trisaccharides, but these structures were not considered since the focus of our study was nonterminal glucosyl residue recognition.

For the monosaccharides, only those structures within 5 kcal/mol of the best energy structure were considered. For the trisaccharides, we considered only ligand structures

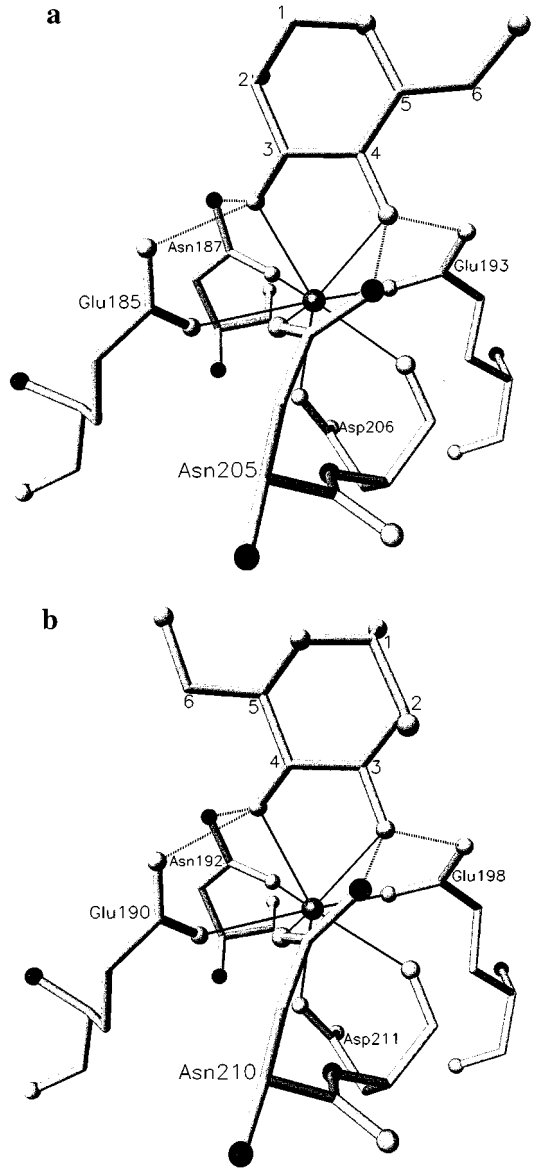


FIGURE 1: Crystal structures of sugars bound to MBP: a) terminal mannose unit complexed with MBP-A; (b) α-methyl-D-mannopyranoside complexed with MBP-C. Oxygen: lightly shaded spheres. Nitrogen: black spheres. Ca^{2+}: larger gray sphere. Carbon atom numbering is shown for the bound sugar. Solid lines: Ca^{2+} coordination bonds. Dashed lines: hydrogen bonds. The methyl group in the structure shown in (b) is not visible. The coordinates for (a) and (b) were taken from PDB accession codes 2msb (*12*) and 1rdl (*11*), respectively.

within 5 kcal/mol of the β-(1→6)-linked trisaccharide, which we used as our benchmark since we had experimental evidence supporting the existence of an internally bound β-(1→6)-linked glucosyl homopolysaccharide (*16*).

To confirm the validity of our docking methodology, we performed docking simulations using MBP-C. The structure of this protein is known when complexed with numerous monosaccharides including α-methyl-D-mannopyranoside (Figure 1b, PDB accession code 1rdl) (*11*). Using the parameters established for SP-D, we docked α-D-mannopyranose into MBP-C. We were able to reproduce the crystallographically determined ligand orientation; superimposition of the known ligand structure with the docked ligand yields an rmsd of 0.9 Å for the sugar ring atoms (not shown). Thus we feel our docking methods are reasonable.

Table 1: Computationally Docked Saccharides

α-D-Glc*p* (α-D-glucopyranose)	α-glucose
β-D-Glc*p* (β-D-glucopyranose)	β-glucose
α-D-Glc*p*-(1→2)-α-D-Glc*p*-(1→2)-β-D-Glc*p*	β-kojitriose
β-D-Glc*p*-(1→2)-β-D-Glc*p*-(1→2)-β-D-Glc*p*	β-sophorotriose
α-D-Glc*p*-(1→3)-α-D-Glc*p*-(1→3)-β-D-Glc*p*	β-nigerotriose
β-D-Glc*p*-(1→3)-β-D-Glc*p*-(1→3)-β-D-Glc*p*	β-laminarotriose
α-D-Glc*p*-(1→4)-α-D-Glc*p*-(1→4)-β-D-Glc*p*	β-maltotriose
β-D-Glc*p*-(1→4)-β-D-Glc*p*-(1→4)-β-D-Glc*p*	β-cellotriose
α-D-Glc*p*-(1→6)-α-D-Glc*p*-(1→6)-β-D-Glc*p*	β-isomaltotriose
β-D-Glc*p*-(1→6)-β-D-Glc*p*-(1→6)-β-D-Glc*p*	β-gentiotriose

Inhibition of SP-D Binding to A. fumigatus. The inhibition experiments were performed as previously described (*2*). Briefly, 20 μg/mL purified recombinant hSP-D was incubated with the appropriate inhibitor for 15 min at 25 °C with mild shaking. The final reaction volume was 100 μL. Following preincubation, the SP-D−inhibitor mixture was added to 2×10^6 paraformaldehyde-fixed *A. fumigatus* conidia. The binding progressed for 1 h at 25 °C, and then the conidia were washed three times to remove unbound SP-D and were incubated with 10 μg/mL polyclonal rabbit anti-hSP-D antibody for 1 h at 25 °C. The conidia were again washed three times to remove unbound primary antibody and were then incubated with 10 μg/mL FITC-conjugated $F(ab')_2$ fragment of donkey anti-rabbit IgG for 1 h at 25 °C. The conidia were then washed twice and analyzed for FITC fluorescence using a Becton Dickinson FACSCalibur flow cytometer and CELLQuest software.

Other Methods. All protein concentrations were determined using the bicinchoninic acid assay according to the manufacturer's instructions (Pierce, Rockford, IL). All figures appearing in this paper were prepared using MolMol (*27*).

RESULTS

We were particularly interested in identifying homopolysaccharides that can be bound by SP-D via interactions with nonterminal carbohydrate residues. We presumed these polysaccharides would be the best ligands for the lectin on the basis of our previous finding that pustulan but not laminarin is a powerful SP-D inhibitor (*16*). One potential concern with this idea is that although the nonreducing terminal glucosyl unit is identical on long-chain polysaccharides, the reducing terminal position is somewhat different due to differences in glycosidic bonding. However, in long-chain polymers the reducing terminal glucosyl unit represents a very small portion of the total molecule. For example, on the basis of average molecular weight, the ratio of reducing terminal glucosyl units to internal units is approximately 0.018 for laminarin and 0.008 for pustulan. Thus we feel it is highly unlikely that differences at the reducing terminal position could account for large differences in polysaccharide inhibitory ability.

To begin our studies, we docked not only glucose but also the glucosyl trisaccharides shown in Table 1. We chose trisaccharides since they are the smallest molecules possessing a nonterminal glucosyl residue, and we tested the eight possible glucosyl trisaccharides in which both glycosidic bonds are identical. For simplicity the reducing terminal glucosyl residue of all of the trisaccharides was maintained in the β-anomeric configuration.

Several structures of MBP complexed with carbohydrate ligands have been reported (*11*, *12*). On the basis of these

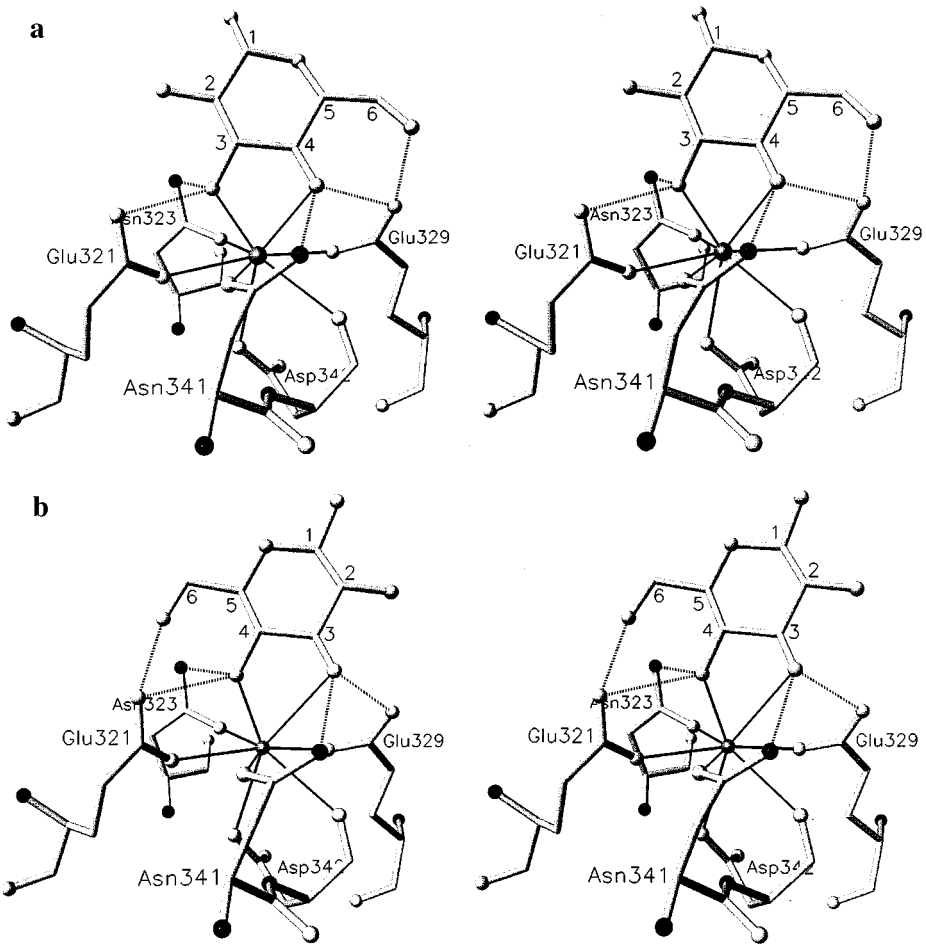


FIGURE 2: Stereoviews of two different β-glucose structures docked into SP-D. Only those amino acids making hydrogen bonds with the ligand or involved in Ca^{2+} coordination are shown. Ca^{2+} coordination bonds inferred from previous work with MBP (*11*, *12*). Other notes as in Figure 1.

findings and additional mutagenic analysis (*15*), five amino acids (Glu185, Asn187, Glu193, Asn205, and Asp206, shown in Figure 1a for MBP-A and referred to as the carbohydrate-binding pocket) are critical for ligation of Ca^{2+} and carbohydrate. These amino acids are conserved in MBP-A and MBP-C, and their structures in this region are essentially identical.

Two alternate carbohydrate ligand orientations have been noted for MBP-A and MBP-C (Figure 1), but the basic structural interactions between the protein and ligand are the same for both orientations. The ligand orientation shown in Figure 1b is a 180° rotation of the orientation in Figure 1a. For simplicity we refer to these alternate orientations as extracyclic carbon right (ECR) as in Figure 1a or extracyclic carbon left (ECL) as in Figure 1b.

Docking of Glucose into SP-D. For β-glucose we found three docked structures that satisfied the criteria described in Experimental Procedures. Two of these structures, one ECR (Figure 2a) and the other ECL (Figure 2b), are shown, and their docked energies are presented in Table 2. The third structure is an ECR orientation, similar to that shown in Figure 2a (rmsd for ring atoms is 1.0 Å). Our structures are similar to the known MBP structures except that we predict one additional hydrogen bond between the 6-OH of the ligand and either Glu321 or Glu329 of SP-D.

There are four satisfactory docked structures of α-glucose. The best energetic structure is shown in Figure 3, and its docking energies are shown in Table 2. The other structures (not shown) are in the ECR orientation similar to β-glucose in Figure 2a. Figure 3 shows an unexpected orientation for α-glucose, given the strong similarity between MBP and SP-D. The glucosyl residue shown in Figure 3 is bound primarily through hydrogen bonding between the protein and the 2 and 3 ligand OH groups and not the 3- and 4-OH groups as previously shown. Although equatorial OH groups occupy the 2 and 3 positions on the glucose ring, they cannot be superimposed on the 3 and 4 equatorial OH groups. This has led previous investigators to speculate that interactions similar to those shown in Figure 3 may not occur for MBP (*12*). However, there are no steric factors precluding this binding arrangement, and as will be discussed below, we

Table 2: Calculated Energies for Docked Mono- and Trisaccharides

	energy (kcal/mol)		
ligand	intermolecular	internal ligand	final docked
α-glucose	−63.60	−3.30	−66.91
β-glucose[a]	−62.14	−10.17	−72.31
β-glucose[b]	−58.81	−9.35	−68.16
β-kojitriose	−82.56	−5.22	−87.78
β-sophorotriose	−69.37	−10.73	−80.10
β-maltotriose	−74.24	−8.61	−82.84
β-isomaltotriose	−90.97	+13.34	−77.63
β-gentiotriose	−83.30	+1.18	−82.12

[a] Data for the structure shown in Figure 2a. [b] Data for the structure shown in Figure 2b.

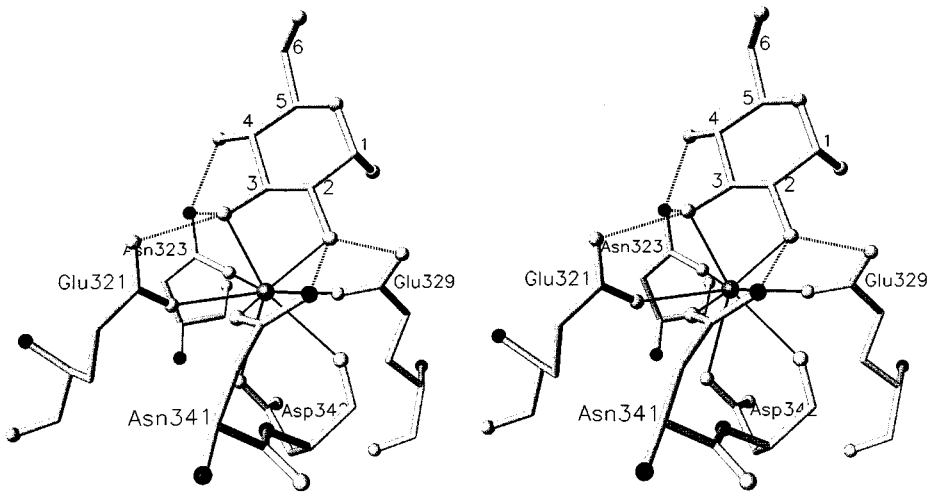


FIGURE 3: Stereoview of an α-glucose structure docked into SP-D. Notes as in Figure 1.

feel that this structure is reasonable and is likely relevant to carbohydrate recognition by SP-D.

Docking of Glucosyl Trisaccharides into SP-D. Both α- and β-(1→2)-linked glucosyl trisaccharides were docked into SP-D. We found five satisfactory structures for β-kojitriose, which is α-(1→2)-linked. The internal glucosyl residues for all of the structures are in the ECR orientation. The top energy structure is shown in Figure 4a, and its docking energies are shown in Table 2. The other structures differ from the structure shown in Figure 4a only by the positions of the two terminal glucosyl residues; the positions of the internal glucosyl residues are very similar (rmsd values for the ring atoms of the internal sugars not shown ranged from 0.3 to 0.9 Å compared with that shown in Figure 4a). The structure shown in Figure 4a demonstrates recognition of the 3- and 4-OH groups of the internal glucosyl residue of an α-(1→2)-linked glucosyl trisaccharide. In addition, the two terminal glucosyl residues are hydrogen-bonded to the side chains of Glu321, Asn323, Asp325, and Arg343.

One satisfactory structure of the β-(1→2)-linked trisaccharide β-sophorotriose appeared. This structure is shown in Figure 4b, and its docking energies are shown in Table 2. It exhibits recognition of the 3- and 4-OH groups of the internal glucosyl residue of a β-(1→2)-linked glucosyl trisaccharide. The internal sugar ring is in the ECR orientation. The rmsd for the ring atoms of the internal ring shown in Figure 4b compared with β-glucose shown in Figure 2a is 0.6 Å. Additionally, as shown in Figure 4b, the terminal glucosyl residues are hydrogen-bonded to the side chain of Asp325 and the main-chain carbonyl oxygen of Pro319.

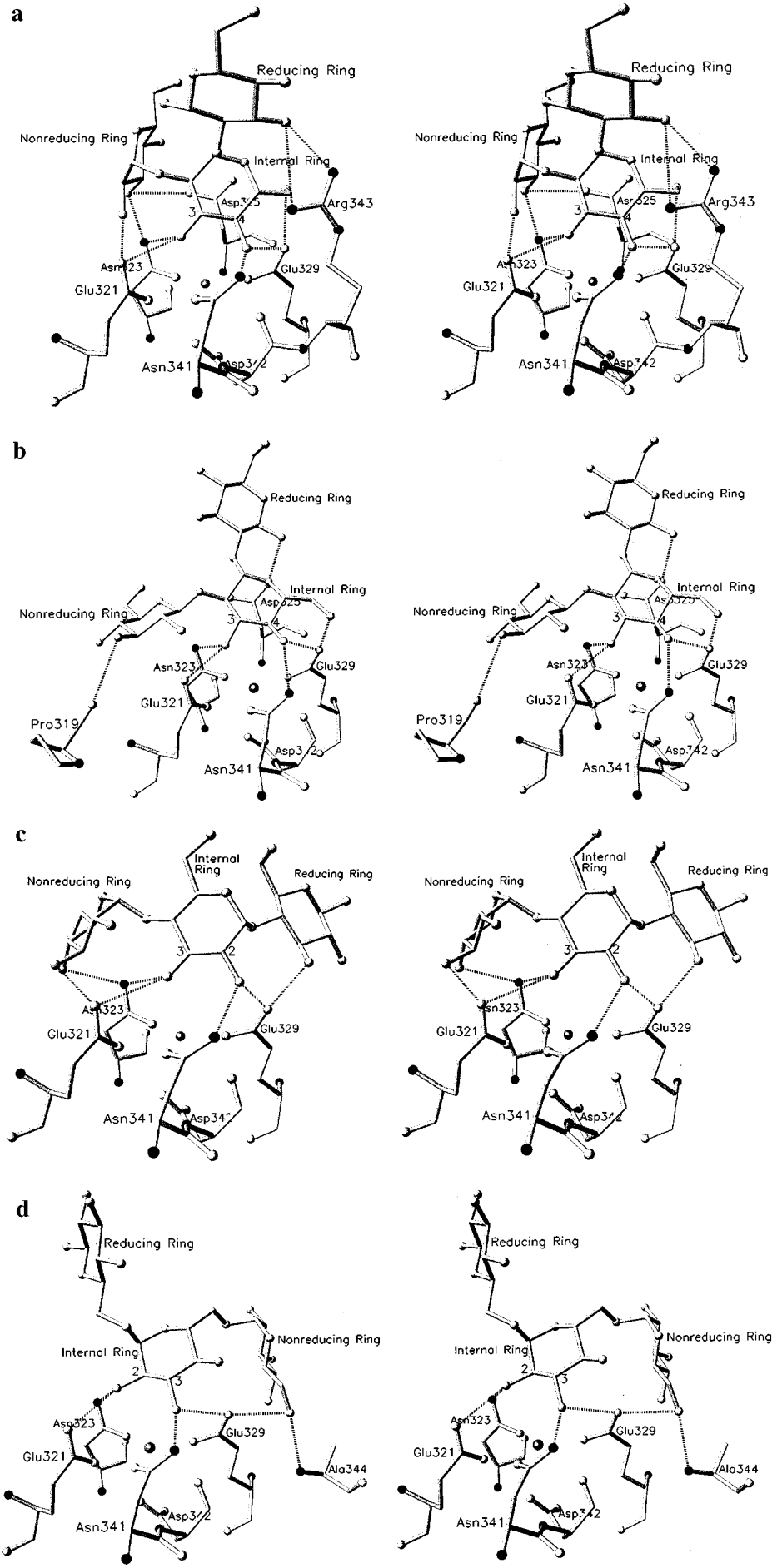
Neither the α- nor the β-(1→3)-linked glucosyl trisaccharides β-nigerotriose and β-laminarotriose, respectively, gave satisfactory structures when docked into SP-D, which is not surprising since no vicinal, equatorial OH groups are present on the internal glucosyl residue of these molecules. This occurs because the 3 position is involved in a glycosidic bond with a neighboring glucosyl ring, and thus it is not possible to form the hydrogen bonds believed to be essential for carbohydrate binding. These observations confirm our previous work where laminarin, a soluble β-(1→3)-linked glucosyl polysaccharide, failed to inhibit SP-D binding to *A. fumigatus* conidia (Table 3) (*16*).

One satisfactory structure was found for the α-(1→4)-linked trisaccharide β-maltotriose. This structure is shown in Figure 4c, and its docking energies are presented in Table 2. It demonstrates recognition of the 2- and 3-OH groups on the internal glucosyl residue of an α-(1→4)-linked glucosyl trisaccharide. The orientation of the internal glucosyl residue is the same as shown for α-glucose in Figure 3; the rmsd for the ring atoms of the internal glucosyl residue compared to the docked α-glucose structure is 0.8 Å. The structure shown in Figure 4c suggests that SP-D can bind long-chain α-(1→4)-linked glucosyl polysaccharides by interactions with the 2- and 3-OH groups on internal glucosyl residues. Below, we present inhibition data supporting this idea.

We found no structures satisfying our criteria for the β-(1→4)-linked glucosyl trisaccharide β-cellobiotriose. A possible explanation for this finding will be discussed below.

One satisfactory docked structure emerged for the α-(1→6)-linked molecule β-isomaltotriose. It is shown in Figure 4d, and its docking energies are presented in Table 2. Interestingly, the interactions between the internal glucosyl residue and the protein are limited to the 2- and 3-OH groups, and the internal glucosyl residue is rotated 180° relative to the orientation of α-glucose in Figure 3. There are additional hydrogen bonds between the 6-OH group of the nonreducing glucosyl residue and the main-chain amide nitrogen of Ala344 and the side chain of Glu329.

One satisfactory structure for the β-(1→6)-linked trisaccharide β-gentiotriose was found. This structure is shown in Figure 4e, and its docking energies are presented in Table 2. It is bound by interactions between the 3- and 4-OH groups of the internal glucosyl residue and the carbohydrate-binding pocket. The internal glucosyl residue is in the ECR orientation. The rmsd for the ring atoms of the internal glucosyl residue is 0.5 Å compared to the docked β-glucose structure shown in Figure 2a. In addition to the interactions between the internal glucosyl residue and the protein, two additional hydrogen bonds are predicted, one from the 1-OH group of the reducing ring and Asn321 and the other between the 2-OH group and the main-chain carbonyl oxygen of Asp325. Figure 4e is in agreement with the conclusion that β-(1→6)-linked glucosyl polysaccharides can be bound by SP-D via interactions primarily with nonterminal glucosyl residues of the ligand. These observations also agree with our previous work showing that a soluble β-(1→6)-linked glucosyl polysaccharide strongly inhibits SP-D (Table 3) (*16*).


a
Reducing Ring
Nonreducing Ring
Internal Ring
Asp325
Arg343
Asn323
Glu329
Glu321
Asn341
Asp342
b
Reducing Ring
Internal Ring
Nonreducing Ring
Asp325
Asn323
Glu329
Glu321
Pro319
Asn341
Asp342
c
Internal Ring
Nonreducing Ring
Reducing Ring
Asn323
Glu321
Glu329
Asn341
Asp342
d
Reducing Ring
Nonreducing Ring
Internal Ring
Asp323
Glu329
Glu321
Ala344
Asn341
Asp342

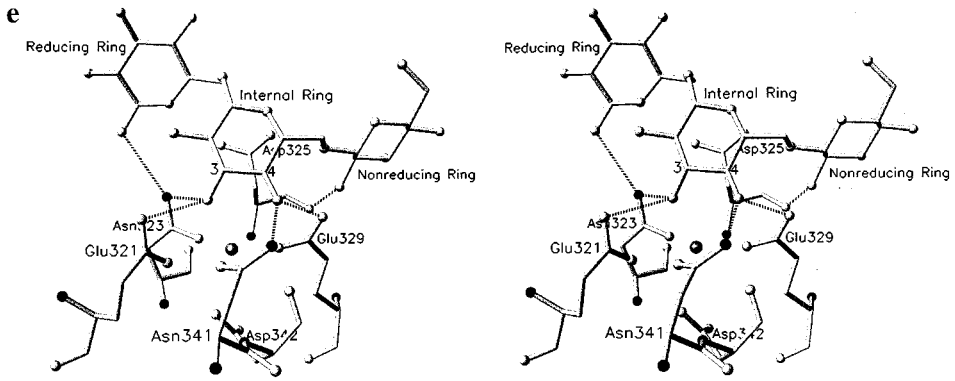


FIGURE 4: Stereoviews of trisaccharides docked into SP-D: (a) β-kojitriose, (b) β-sophorotriose, (c) β-maltotriose, (d) β-isomaltotriose, and (e) β-gentiotriose. Other notes as in Figure 1.

Table 3: IC_{50} Values for Inhibition of Recombinant hSP-D Binding to *A. fumigatus* Conidia by Various Glucosyl Homopolysaccharides

inhibitor	predominant glycosidic bond configuration	IC_{50} (μM glucose equiv)[a]
laminarin[b]	β-(1→3)	$>10^5$
starch	α-(1→4)	63.6 ± 11.8
carboxymethylcellulose	β-(1→4)	$>10^3$
dextran	α-(1→6)	4.8 ± 2.4
pustulan[b]	β-(1→6)	1.0 ± 0.3

[a] Values are mean ± standard error of three experiments. [b] Values taken from Allen et al. (*16*).

Inhibition of SP-D Binding to A. fumigatus Conidia. We have previously investigated the binding of SP-D to *A. fumigatus* conidia (*2*, *16*). Use of carbohydrate inhibitors and mutated surfactant proteins suggested that the binding resulted from the recognition of carbohydrate structures on the surface of the organism. Additionally, we reported that pustulan [a β-(1→6)-linked glucosyl homopolysaccharide] but not laminarin [a β-(1→3)-linked glucosyl homopolysaccharide] strongly inhibited SP-D binding to *A. fumigatus* conidia (*16*). To experimentally examine the carbohydrate recognition specificity of SP-D, we tested the soluble glucosyl homopolysaccharides starch, dextran, and CMC as inhibitors of SP-D binding to *A. fumigatus* conidia. The heterogeneous nature of other long-chain glucosyl polysaccharides such as nigeran would greatly complicate the interpretation of inhibition data, and therefore they were not tested. We have reported the inhibitor concentrations as glucose equivalents since the polymorphic nature of the long-chain polysaccharides makes direct comparisons less reliable. As can be seen, starch and dextran inhibited binding, but CMC failed to do so (Table 3). These data and the data for pustulan and laminarin are consistent with the results of our glucosyl trisaccharide docking analysis (Table 2). A summary of our findings is shown in Table 4. It is also noteworthy that the IC_{50} for glucose inhibition of SP-D binding to *A. fumigatus* conidia is $(17.8 \pm 1.7) \times 10^3$ μM (mean ± standard error of three experiments). The fact that glucose showed an IC_{50} value between 300 and 18 000 times that of the long-chain glucosyl polymers on a glucose equivalent basis demonstrates a cooperative interaction between the protein and the long-chain polysaccharides. This is likely due to binding of the extended ligand to multiple CRDs. As shown in the crystal structure (*17*), the distance between adjacent carbohydrate binding sites is 51 Å. The long-chain polysaccharides tested could easily span this distance.

Table 4: Study Summary

ligand	does the trisaccharide dock via the internal glucosyl residue?	is the related polysaccharide an SP-D inhibitor?
β-kojitriose	yes	not determined
β-sophorotriose	yes	not determined
β-nigerotriose	no	not determined
β-laminarotriose	no	no
β-maltotriose	yes	yes
β-cellotriose	no	no
β-isomaltotriose	yes	yes
β-gentiotriose	yes	yes

DISCUSSION

Carbohydrate recognition is a critical function of C-type lectins (*28*). For example, the selectins mediate cell adhesion by binding specific cell surface carbohydrates, whereas MBP and SP-D mediate pathogen clearance through recognition of carbohydrate structures on the surface of invading microorganisms. Most of the structural knowledge of carbohydrate recognition by C-type lectins comes from studies using MBP. These studies demonstrated that carbohydrate binding is accomplished by specific interactions between the protein and two adjacent equatorial OH groups on monosaccharides or the nonreducing terminal carbohydrate residue of polysaccharides (*11*−*13*). Given the structural similarity in the CRDs of the C-type lectins, it appears that other members of this protein family bind carbohydrates through similar interactions. Although the C-type lectins generally show weak affinity for monosaccharides (*7*, *10*), tight binding may be achieved through multivalent interactions of the oligomeric protein with numerous properly spaced carbohydrate ligands on the cell surface (*12*, *28*). Thus, it appears that the geometric arrangement of ligands on the cell surface contributes to binding specificity (*28*).

Long-chain polysaccharides are present on the surface of many microorganisms. We have previously investigated the interactions of SP-D with the fungi *A. fumigatus* and *S. cerevisiae* (*2*, *16*). The surface of these organisms contains mannosyl and glucosyl polysaccharides, and we showed that β-(1→6) but not β-(1→3) glucan is a fungal ligand for SP-D (*16*). Since the nonreducing terminal glucosyl residues of these polysaccharides are essentially the same, we reasoned that the ability of the β-(1→6)-linked molecule to inhibit SP-D was a function of differential binding to nonterminal glucosyl residues. Additionally, since the majority of carbohydrate residues in long-chain polysaccharides are non-

terminal, we reasoned that only soluble polysaccharides that can be bound via their internal carbohydrate residues will be strong SP-D inhibitors. This raised the possibility that binding to nonterminal sugar residues contributes significantly to carbohydrate recognition by SP-D and other C-type lectins.

In the present study we used automated docking and inhibition analysis to examine polysaccharide recognition by SP-D. We were specifically interested in examining the recognition of nonterminal carbohydrate residues. While C-type lectin binding to nonterminal carbohydrate residues has not been excluded, most investigations into carbohydrate recognition by these proteins have focused on binding to monosaccharides or terminal carbohydrate residues of polysaccharides. Recently, one group concluded that the binding of MBP to different *Salmonella enterica* and *Neisseria gonorrhoeae* mutants could not always be predicted by the identity of the terminal lipopolysaccharide carbohydrate residue present on the cell surface (*29*), but binding of the lectin to nonterminal residues was not investigated.

The docking of β-glucose into the SP-D receptor site was essentially as expected based on the known crystallographic structures of MBP-A complexed with an oligosaccharide (*12*) and MBP-C complexed with various monosaccharides (*11*). The structures shown in Figure 2 closely resemble the MBP structures where the predominant binding interactions are between the protein and vicinal, equatorial OH groups at the 3 and 4 positions on the sugar ring. We found both saccharide orientations previously seen for MBP. In agreement with earlier observations (*30*), we feel the most reasonable explanation for these findings is that both orientations represent possible binding modes.

Our findings with α-glucose were unexpected. Although we found ligand orientations resembling that shown in Figure 2a, the best energetic structure (Figure 3) was bound primarily through interactions of the protein with the 2- and 3-OH groups of the ligand. For glucose the 2-, 3-, and 4-OH groups are all in the equatorial orientation. However, due to ring twist the 3- and 4-OH groups cannot be superimposed on the 2- and 3-OH groups. This has been used to explain why MBP binds glucose and mannose but not galactose (*12*). On the basis of previous structural analysis, it was not considered that the lectin could bind a ligand through interactions equivalent to those shown in Figure 3. Upon examination of the structures, it was noted that, with only slight changes in the orientation of the sugar ring relative to the known crystallographic structure, the 2- and 3-OH groups of glucose can be placed in positions equivalent to those of the 3- and 4-OH groups of mannose (*12*). As shown in Figure 5, the 4-carbon position of the docked α-glucose is displaced 1.1 Å from the structurally related 2-carbon position of mannose. The displacement is toward Asn323 and allows a hydrogen bond between the side chain of Asn323 and the 4-OH group of the ligand (Figure 3). Although this orientation is not as favorable as the 3,4-OH orientations seen for β-glucose (Table 2), there are no steric clashes for this structure. Given the fact that starch is a SP-D inhibitor, this represents a relevant ligand orientation for SP-D.

We have shown that SP-D can dock to the internal glucosyl residue of the α- and β-(1→2)-linked trisaccharides. Long-chain (1→2)-linked glucosyl homopolysaccharides do not exist and thus could not be tested as SP-D inhibitors. Cyclic

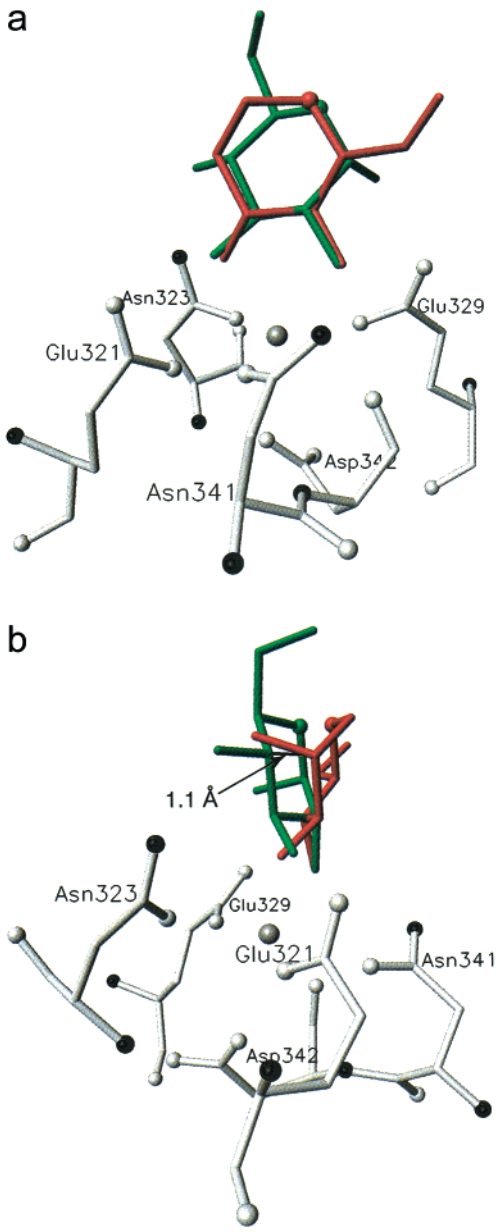


FIGURE 5: Comparison of the orientations of mannose bound to MBP with α-glucose docked into SP-D. (a) The MBP CRD structure (2msb chain B from Figure 1a) (*12*) is superimposed on the SP-D CRD structure (1b08A chain A) (*17*) to give the position of the bound mannose (red) in the SP-D carbohydrate-binding pocket. The docked α-glucose (from Figure 3) is shown in green. (b) 90° rotation of the structure in (a) to show sugar ring separation. Other notes as in Figure 1.

β-(1→2) glucans (*31*) were not tested because they are small (10−40 glucosyl residues per molecule) compared to the long-chain polysaccharides shown in Table 3 and would not yield a direct comparison of inhibition data with them.

It is noteworthy that although a total of six docked α/β-(1→2)-linked trisaccharide structures were identified (Figure 4a,b, some not shown), the internal glucosyl residues for all of the structures are in the ECR orientation. The ECL orientation may be blocked by the presence of a bulky substitution at the 2 position. The structure of MBP complexed with α-Me-GlcNAc (PDB accession code 1rdn), which also has a bulky 2-position substitution, is known. The α-Me-GlcNAc in this structure is in the ECL orientation. However, when this structure is superimposed on the SP-D

structure, steric clashes occur between the substituted 2 position and the side chain of Arg343 (not shown). In MBP the amino acid corresponding to Arg343 is either Val or Ile and does not clash with the substituted 2 position. These observations suggest not only a mechanism for the orientations seen for the (1→2)-linked trisaccharides but also a mechanism for fine-tuning C-type lectin binding specificity. We are currently using mutagenesis to examine this idea further.

Internal glucosyl residues of (1→3)-linked glucosyl polysaccharides do not dock to SP-D, and laminarin is not a SP-D inhibitor (Table 3). Figures 2 and 3 show that the 3-OH groups of all monosaccharides hydrogen-bond with the protein, but this OH group is not available on internal glucosyl residues of (1→3)-linked polysaccharides due to glycosidic bonding.

The α-(1→4)-linked trisaccharide β-maltotriose docks to SP-D via its internal glucosyl residue (Figure 4c) in an orientation similar to that shown for α-glucose (Figure 3). This observation was confirmed in vitro by demonstrating that starch inhibits SP-D binding to *A. fumigatus*. However, the ability of α-(1→4)-linked polysaccharides to bind to SP-D is likely not true for all other similarly linked polysaccharides. For example, since the 2-OH group on mannose is in the axial orientation, the internal glucosyl residues of an α-(1→4)-linked mannosyl polysaccharide may not be bound by SP-D.

The β-(1→4)-linked glucosyl trisaccharide β-cellotriose is not docked via internal glucosyl residues. A likely explanation for this lies in the internal hydrogen bonding of β-(1→4)-linked glucosyl polysaccharides. The 3-OH group forms a hydrogen bond with the cyclic oxygen of the neighboring sugar ring and would not be available for a hydrogen bond with a protein oxygen. In addition to this intrachain hydrogen bond, the 2-OH group of internal glucosyl residues is hydrogen-bonded to 6-OH groups of a neighboring polysaccharide chain in solid cellulose, giving cellulose a lattice structure in which every internal glucosyl residue hydroxyl group is involved in either intrachain or interchain hydrogen bonding. Thus, we feel it is highly unlikely that SP-D can bind internal glucosyl residues of cellulose. However, SP-D shows significant binding to cellulose granules (data not shown), apparently caused by the large number of terminal glucosyl residues on the granule surface. Therefore, we do not exclude the possibility of binding to terminal glucosyl residues. In fact, our docking studies predict binding to the terminal glucosyl residues of all of the trisaccharides examined, including β-cellotriose (data not shown).

To further examine possible binding of β-(1→4)-linked glucosyl polysacharides, we tested CMC for its ability to inhibit *A. fumigatus* binding by SP-D. CMC is a soluble β-(1→4)-linked polysaccharide in which most of the glucosyl subunits are substituted at the 6 position. If binding to β-(1→4)-linked glucosyl polysaccharides can occur in a manner similar to that previously discussed for α-(1→4)-linked glucosyl saccharides, CMC should be an ideal soluble test polymer since both the 2- and 3-OH groups are available. As shown in Table 3, CMC failed to inhibit SP-D binding to *A. fumigatus*. This observation supports the docking results and suggests that intramolecular ligand hydrogen bonding between the 3-OH and the neighboring sugar ring does not allow binding by SP-D.

Docking analysis predicts that both α- and β-(1→6)-linked glucosyl trisaccharides are bound via their internal glucosyl residue. This was not surprising since pustulan is a powerful SP-D inhibitor. Figure 4e provides a structural explanation for the ability of pustulan to inhibit SP-D. We originally expected that docked α-(1→6)-linked trisaccharide structures bound via the 3- and 4-OH groups of the internal glucosyl residue would be identified, but none were found within the lowest 7.5 kcal/mol of the optimal structure. However, an α-(1→6)-linked trisaccharide with a higher energy docked primarily via its 3- and 4-OH groups does occur, and this structure may represent a relevant ligand orientation. As shown in Table 3, the α-(1→6)-linked glucosyl homopolysaccharide dextran strongly inhibits SP-D, supporting the conclusion that this glycosidic bond configuration allows internal glucosyl residue recognition by SP-D.

Structural studies of C-type lectins have previously implicated the five amino acids equivalent to those shown in Figures 1−3 as defining a carbohydrate-binding pocket. This conclusion was based on observations made using monosaccharides or terminal glucosyl residues of a polysaccharide. Although additional sites, such as the side chains of Asp325 and Arg343 (Figure 4), might be important for polysaccharide binding, that role is likely not as significant as that played by the five amino acids mentioned above. As can be seen, the interactions of the internal glucosyl residue with the carbohydrate-binding pocket are basically the same for all the polysaccharides. The flanking glucosyl residues do form hydrogen bonds with residues outside the binding pocket, but these are typically limited to two or three, compared with the four invariant bonds seen between the internal glucosyl residue and the carbohydrate-binding pocket. For all trisaccharides shown, the internal glucosyl residue contributes more to the intermolecular energy than either of the flanking residues. The contribution from the internal residue ranged from 47% of the total for the α-(1→4)-linked trisaccharide to 72% for the α-(1→2)-linked trisaccharide. Additionally, the interactions with side chains outside the binding pocket are different for each trisaccharide examined. Thus, the role of the Asp325 and Arg343 side chains may be limited to fine-tuning binding specificity, consistent with the fact that the amino acid residues at positions equivalent to theirs are not conserved in C-type lectins.

In conclusion, we have used automated docking and inhibition analysis to examine carbohydrate recognition by SP-D. Our data suggest that SP-D can bind glucose and glucosyl polysaccharides through hydrogen bonding primarily between the protein and the 3- and 4-OH groups of the ligand as has been previously reported for MBP. The present study extends the previous work by demonstrating binding to nonterminal glucosyl residues. Additionally, we propose a novel interaction between the protein and the 2- and 3-OH groups of the ligand. Given the similarity between C-type lectin CRDs, we feel these results may be applicable to other members of the protein family and will assist in the prediction of C-type lectin ligands including microbial surface carbohydrates. These results may also be useful in designing inhibitors of carbohydrate recognition by C-type lectins.

ACKNOWLEDGMENT

We thank Dennis Voelker for providing the recombinant hSP-D and anti-hSP-D antibody used in this study, for reviewing the manuscript, and for helpful discussions. We also thank John Robyt for insightful discussions and comments. We acknowledge the bioinformatics and computational biology program at Iowa State University for providing computational resources. The experimental work was performed in the Lord and Taylor Laboratory for Lung Biochemistry at the National Jewish Medical and Research Center.

REFERENCES

1. Kuroki, Y., and Voelker, D. R. (1994) *J. Biol. Chem. 269*, 25943−25946.
2. Allen, M. J., Harbeck, R., Smith, B., Voelker, D. R., and Mason, R. J. (1999) *Infect. Immun. 67*, 4563−4569.
3. Reid, K. B. (1998) *Biochim. Biophys. Acta 1408*, 290−295.
4. Madan, T., Eggleton, P., Kishore, U., Strong, P., Aggrawal, S. S., Sarma, P. U., and Reid, K. B. (1997) *Infect. Immun. 65*, 3171−3179.
5. Kuan, S. F., Rust, K., and Crouch, E. (1992) *J. Clin. Invest. 90*, 97−106.
6. van Rozendaal, B. A., van Spriel, A. B., van De Winkel, J. G., and Haagsman, H. P. (2000) *J. Infect. Dis. 182*, 917−922.
7. Persson, A., Chang, D., and Crouch, E. (1990) *J. Biol. Chem. 265*, 5755−5760.
8. Haagsman, H. P., Hawgood, S., Sargeant, T., Buckley, D., White, R. T., Drickamer, K., and Benson, B. J. (1987) *J. Biol. Chem. 262*, 13877−13880.
9. Haurum, J. S., Thiel, S., Haagsman, H. P., Laursen, S. B., Larsen, B., and Jensenius, J. C. (1993) *Biochem. J. 293*, 873−878.
10. Lee, R. T., Ichikawa, Y., Fay, M., Drickamer, K., Shao, M. C., and Lee, Y. C. (1991) *J. Biol. Chem. 266*, 4810−4815.
11. Ng, K. K., Drickamer, K., and Weis, W. I. (1996) *J. Biol. Chem. 271*, 663−674.
12. Weis, W. I., Drickamer, K., and Hendrickson, W. A. (1992) *Nature 360*, 127−134.
13. Weis, W. I., and Drickamer, K. (1996) *Annu. Rev. Biochem. 65*, 441−473.
14. Ogasawara, Y., and Voelker, D. R. (1995) *J. Biol. Chem. 270*, 14725−14732.
15. Drickamer, K. (1992) *Nature 360*, 183−186.
16. Allen, M. J., Voelker, D. R., and Mason, R. J. (2001) *Infect. Immun. 69*, 2037−2044.
17. Hakansson, K., Lim, N. K., Hoppe, H. J., and Reid, K. B. (1999) *Struct. Folding Des. 7*, 255−264.
18. Morris, G. M., Goodsell, D. S., Halliday, R. S., Huey, R., Hart, W. E., Belew, R. K., and Olson, A. J. (1998) *J. Comput. Chem. 19*, 1639−1662.
19. Goodsell, D. S., and Olson, A. J. (1990) *Proteins 8*, 195−202.
20. Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S., and Karplus, M. (1983) *J. Comput. Chem. 4*, 187−217.
21. Cornell, W. D., Cieplak, P., Bayly, C. I., Gould, I. R., Merz, K., Ferguson, D. M., Spellmeyer, D. C., Fox, T., Caldwell, J. W., and Kollman, P. A. (1995) *J. Am. Chem. Soc. 117*, 5179−5197.
22. Mehler, E. L., and Solmajer, T. (1991) *Protein Eng. 4*, 903−910.
23. Aqvist, J. (1990) *J. Phys. Chem. 94*, 8021−8024.
24. Woods, R. J., Dwek, R. A., Edge, C. J., and Fraser-Reid, B. (1995) *J. Phys. Chem. 99*, 3832−3846.
25. Guex, N., and Peitsch, M. C. (1997) *Electrophoresis 18*, 2714−2723.
26. Laederach, A., Dowd, M. K., Coutinho, P. M., and Reilly, P. J. (1999) *Proteins 37*, 166−175.
27. Koradi, R., Billeter, M., and Wuthrich, K. (1996) *J. Mol. Graphics 14*, 51−55.
28. Weis, W. I., Taylor, M. E., and Drickamer, K. (1998) *Immunol. Rev. 163*, 19−34.
29. Devyatyarova-Johnson, M., Rees, I. H., Robertson, B. D., Turner, M. W., Klein, N. J., and Jack, D. L. (2000) *Infect. Immun. 68*, 3894−3899.
30. Hakansson, K., and Reid, K. B. (2000) *Protein Sci. 9*, 1607−1617.
31. Breedveld, M. W., and Miller, K. J. (1994) *Microbiol. Rev. 58*, 145−161.

BI002901Q