GLYCO-2004-00030
Submitted 2-28-04

Arg343 in Human Surfactant Protein D Governs Discrimination between Glucose and *N*-Acetylglucosamine Ligands

Martin J. Allen[1,7], Alain Laederach[2,3,8], Peter J. Reilly[2], Robert J. Mason[1,4], and Dennis R. Voelker[1,5,6]

[1]Department of Medicine, National Jewish Medical and Research Center, Denver, CO 80206, USA

[2]Department of Chemical Engineering and [3]Bioinformatics and Computational Biology Program, Iowa State University, Ames, IA 50011, USA

Departments of [4]Medicine and [5]Biochemistry and Molecular Genetics, University of Colorado Health Sciences Center, Denver, CO 80262, USA.

[6] To whom correspondence should be addressed: Dennis R. Voelker, Department of Medicine, National Jewish Medical and Research Center, 1400 Jackson Street, Denver, CO 80206. Tel.: +1-303-398-1300; Fax: +1-303-398-1806. *E-mail address*: voelkerd@njc.org

[7] Present address: Department of Cell Sciences, Amgen Inc., 51 University Street, Seattle, WA 98101

[8] Present address: Department of Genetics, Stanford University, 300 Pasteur Drive, Stanford, CA 94305

Running Title: Mechanism of ligand affinity by SP-D

Keywords: C-type lectin/Glucose/Ligand binding/*N*-acetyl-D-glucosamine/Surfactant protein D

Abstract

Surfactant protein D (SP-D), one of the members of the collectin family of C-type lectins, is an important component of pulmonary innate immunity. SP-D binds carbohydrates in a calcium-dependent manner, but the mechanisms governing its ligand recognition specificity are not well understood. SP-D binds glucose (Glc) stronger than *N*-acetylglucosamine (GlcNAc). Structural superimposition of hSP-D with mannose binding protein C (MBP-C) complexed with GlcNAc reveals steric clashes between the ligand and the side chain of Arg^{343} in hSP-D. To test whether Arg^{343} contributes to Glc > GlcNAc recognition specificity, we constructed a computational model of $Arg^{343}\rightarrow$Val (R343V) mutant hSP-D based on homology with MBP-C. Automated docking of α-Me-Glc and α-Me-GlcNAc into wild-type hSP-D and the R343V mutant of hSP-D suggests that Arg^{343} is critical in determining ligand-binding specificity by sterically prohibiting one binding orientation. To empirically test the docking predictions, an R343V mutant recombinant hSP-D was constructed. Inhibition analysis shows that the R343V mutant binds both Glc and GlcNAc with higher affinity than the wild-type protein, and that the R343V mutant binds Glc and GlcNAc equally well. These data demonstrate that Arg^{343} is critical for hSP-D recognition specificity and plays a key role in defining ligand specificity differences between MBP and SP-D. Additionally, our results suggest that the number of binding orientations contributes to monosaccharide binding affinity.

Introduction

Pulmonary surfactant is a complex mixture of lipids and proteins that is present in the alveolar compartment of the lungs. In addition to preventing alveolar collapse during expiration, there is increasing evidence that the protein components of surfactant play a significant role in pulmonary host defense. Specifically, numerous in vitro (Crouch 1998;Crouch 2000;Mason et al., 1998;Wright 1997) and more recently in vivo studies that include the use of gene-targeted mice deficient in the surfactant proteins, SP-A or SP-D (Hickling et al., 1999;LeVine et al., 1997;LeVine et al., 1998;LeVine et al., 1999b;LeVine et al., 1999a;LeVine et al., 2000;Madan et al., 2001), have demonstrated a host defense role for these two proteins.

SP-A and SP-D are members of the collectin family of C-type lectins that also includes the mannose binding proteins (MBPs). C-type lectins have a 115-120 amino acid calcium-dependent carbohydrate recognition domain (CRD). SP-D also has an *N*-terminal region involved in interchain disulfide bonding followed by a collagen-like domain and a neck region that connects the collagen-like domain to the *C*-terminal CRD (Kuroki and Voelker 1994). SP-D oligomerizes through trimeric intermediates to form cruciform-like dodecamers. Both the noncovalent and covalent oligomerization of SP-D function to amplify its binding affinity for multivalent ligands.

Among the C-type lectins, selectins function in cell adhesion and the collectins SP-A, SP-D, and the MBPs are effectors of innate immunity. Recognition of specific carbohydrate structures is crucial to these functions, and therefore the mechanisms of carbohydrate ligation by C-type lectins have been the subject of extensive study (Allen et al., 2001;Burrows et al., 1997;Feinberg et al., 2001;Graves et al., 1994;Hitchen et al., 1998;Lee et al., 1991;Ng et al., 1996;Ng et al., 2002;Simanek et al., 1998;Somers et al., 2000;Weis et al., 1992). Previous work has included engineering the MBPs, selectins, SP-A, and SP-D for altered ligand binding specificity (Blanck

et al., 1996;Drickamer 1992;Kogan et al., 1995;McCormack et al., 1994;Ng and Weis 1997;Ogasawara and Voelker 1995a). The three-dimensional structure of the trimeric neck-CRD regions of hSP-D has been reported in both unligated (Hakansson et al., 1999;Shrive et al., 2003) and maltose-ligated forms (Shrive et al., 2003). The corresponding three-dimensional structure for unligated SP-A has also recently been reported (Head et al., 2003).

We previously used automated computational docking and inhibition analysis to examine and define the glycosidic bond configurations required for nonterminal sugar unit recognition by hSP-D (Allen et al., 2001). One advantage of automated docking is that it allows the visualization of multiple ligand binding modes, some of which may not be the most energetically favorable and might not be detected by other techniques. However, given that collectin-ligand interactions are multivalent with respect to both the receptor and the ligand, it is important to consider suboptimal binding modes when using a monovalent model system such as competitive inhibition with monosaccharides. In fact, our previous docking efforts suggested a novel orientation for glucose binding by hSP-D (Allen et al., 2001). Although this orientation was not the lowest energy structure, its existence was supported using competitive inhibition analysis (Allen et al., 2001). Thus, when combined with experimental work, automated docking is a powerful tool for examining complex protein-ligand interactions.

Previous work by others examined the mono-, di-, and trisaccharide specificity of rat SP-D (rSP-D) by competitive inhibition analysis (Persson et al., 1990), and assigned monosaccharide binding affinity in the order α-Me-Glc > Glc > β-Me-Glc > GlcNAc. We found the specificity Glc >> GlcNAc intriguing since the highly homologous MBP-A shows recognition affinity in the order Glc ≈ GlcNAc (Drickamer 1992;Lee et al., 1991).

This study was designed to more completely define the mechanisms governing carbohydrate

recognition specificity by hSP-D. We used automated docking, mutagenesis, and inhibition analysis to examine monosaccharide recognition by hSP-D. Our findings provide a clear molecular explanation for the differences in binding specificity between SP-D and MBP for GlcNAc ligands. In addition, these results suggest a rational mechanism for genetic engineering of C-type lectin recognition specificity.

Results and discussion

hSP-D recognizes carbohydrates in the order Glc ≈ α-Me-Glc ≈ β-Me-Glc > GlcNAc

We first tested Glc, α-Me-Glc, β-Me-Glc, and α-Me-GlcNAc for their ability to inhibit carbohydrate recognition by hSP-D. Binding of wild-type hSP-D to mannose-Sepharose beads is inhibited by monosaccharides in the order Glc ≈ α-Me-Glc ≈ β-Me-Glc > α-Me-GlcNAc. The K_i for α-Me-GlcNAc is approximately three times that found for Glc, α-Me-Glc or β-Me-Glc (Table I). In work performed previously and in this project, Glc is approximately three times more effective than GlcNAc at inhibiting hSP-D binding to *A. fumigatus* conidia [IC_{50} for Glc = 17.8 ± 1.7 mM (Allen et al., 2001); IC_{50} for GlcNAc = 52.0 ± 4.3 mM; data are average ± standard error of three experiments], showing that this recognition profile is not unique to a specific ligand. Furthermore, the Glc > GlcNAc affinity difference for SP-D found here agrees with findings from other laboratories (Kuan et al., 1992;Persson et al., 1990;Vuk-Pavlovic et al., 2001). Specifically, rSP-D showed a tenfold preference for Glc over GlcNAc when the monosaccharides were used as inhibitors of rSP-D binding to maltosyl-BSA (Persson et al., 1990).

The observation that SP-D recognizes Glc more readily than GlcNAc is interesting, since the only structural difference between the sugars is a 2-*N*-linked acetyl group, while the predominant

binding interactions occur between the protein and the 3- and 4-OH of the ligand (Shrive et al., 2003). Furthermore, the highly homologous MBP-A recognizes Glc and GlcNAc equally (Drickamer 1992;Lee et al., 1991).

The lack of a significant difference among Glc, α-Me-Glc, and β-Me-Glc as inhibitors of hSP-D binding to mannose-Sepharose beads differs from the results with rSP-D, where the order is α-Me-Glc > Glc > β-Me-Glc (Persson et al., 1990). The reasons for the differences between hSP-D and rSP-D are unclear, but they may relate to subtle changes in amino acid residues distant from the primary binding site. Residue 343, which is near the binding site and is the subject of computational modeling and mutagenesis in this study, is Lys in rSP-D and Arg in hSP-D.

Superimposition of GlcNAc in the hSP-D binding site reveals steric clashes with Arg343

Since MPB-A and SP-D differ in their relative affinity for Glc and GlcNAc, we wished to elucidate the mechanisms governing this specificity in hSP-D. Fig. 1 shows the structure of α-Me-GlcNAc complexed with amino acid residues Glu^{190}, Asn^{192}, Glu^{198}, Asn^{210}, Asp^{211}, and Val^{212} of MBP-C (Ng et al., 1996). The first five residues comprise the primary carbohydrate and calcium-binding site and are conserved in the MBPs and SP-D. The ligand is bound primarily by hydrogen bonds between the protein and the 3- and 4-OH groups on the sugar ring, as seen for a variety of carbohydrate ligands bound by both MBP-C and MBP-A (Ng et al., 1996;Ng et al., 2002). When the extracyclic C-6 is on the left as shown in Fig. 1, the orientation is designated ECL (Extracyclic Carbon Left) (Allen et al., 2001). The ECL orientation is also seen in MBP-C complexed with α-Me-Man (Ng et al., 1996). Significantly, in the structure of MBP-A complexed with a terminal mannose unit, the ligand is rotated 180° relative to Fig. 1 so that the

extracyclic carbon is on the right in an orientation we refer to as ECR (Extracyclic Carbon Right and similar to the orientation shown in Fig. 2b).

Superimposition of the SP-D (Hakansson et al., 1999) and MBP-C binding sites reveals a steric clash between the ligand and hSP-D residue Arg^{343}, corresponding to Val^{212} in MBP-C (Fig. 1). The residue corresponding to Val^{212} is conserved in the MBPs as either Val or Ile, while it is conserved as either Arg or Lys in all published SP-D sequences. In the recently reported hSP-D structure complexed with maltose, a strained relationship between the Arg^{343} side chain and the carbohydrate ligand was noted (Shrive et al., 2003), leading the authors to speculate that this residue plays a critical role in ligand recognition by the protein. Additionally, although the Arg^{343} side chain was noted to adopt different conformations in the six CRD structures reported (Shrive et al., 2003), all conformations displayed steric clashes similar to those represented in figure 1 upon superimposition of GlcNAc as described above. Moreover, previous studies of monosaccharide binding by the macrophage mannose receptor C-type lectin and chicken hepatic lectin led to the conclusion that the homologous residue in those proteins (Ile^{749} and Val^{191} respectively) participates in GlcNAc binding specificity (Burrows et al., 1997;Hitchen et al., 1998). Taken together these observations strongly suggest that Arg343 in hSP-D plays a significant, but as yet undefined, role in carbohydrate recognition by the protein.

α-Me-GlcNAc binding by MBP

To begin our theoretical analysis of carbohydrate recognition specificity by hSP-D we first docked α-Me-GlcNAc into the MBP-C carbohydrate-binding site (not shown; docking energies presented in Table II). Using these methods we reproduced the known orientation of α-Me-GlcNAc complexed with MBP-C (ligand in the ECL orientation (Ng et al., 1996)). The rmsd for the ring atoms of the docked ligand was 1.0 Å relative to the known crystal structure. Our

docking simulations also predicted ligand binding in the ECR orientation. This suggests that both binding ECR and ECL modes may exist in solution. The crystal structure of highly homologous MBP-A complexed with α-Me-GlcNAc was recently reported (Ng et al., 2002). In that work the ligand was bound in both ECR and ECL orientations (called "orientation I" and "orientation II" respectively). Additionally, in the structure of the disaccharide Manα1-3Man complexed with MBP-C the ligand was reported to be bound in both ECL and ECR orientations (Ng et al., 2002). Thus, multi-orientation ligand binding is possible for C-type lectins including MBP-C.

Automated docking suggests that wild-type hSP-D binds α-Me-Glc in two orientations but binds α-Me-GlcNAc in only one orientation

Automated docking was used to model α-Me-Glc and α-Me-GlcNAc binding by hSP-D. It predicts binding of α-Me-Glc in both ECR and ECL orientations (Fig. 2), similar to previous results for docking of β-Glc with hSP-D (Allen et al., 2001). The docked energies of both orientations are similar (Table II). However, only the ECR orientation was permissible when α-Me-GlcNAc was docked into hSP-D (Fig. 3). Given the superimposition shown in Fig. 1, it seems reasonable to conclude that the ECL orientation is not allowed for α-Me-GlcNAc binding due to steric hindrance between the bulky group at C-2 of the GlcNAc ring and the Arg^{343} side chain.

We hypothesize that the number of low energy binding orientations contributes to the overall affinity for a given ligand. The fact that α-Me-GlcNAc is a weak ligand for SP-D compared to α-Me-Glc likely reflects binding in the ECR orientation only. In support of this idea, previous work has shown that *myo*-Inositol is a more potent SP-D inhibitor than Glc in competition

studies with complex ligands (Allen et al., 1999;de Wetering et al., 2001). *myo*-Inositol has five equatorial OH groups whereas α-Glc and β-Glc have three and four, respectively. Since vicinal equatorial OH groups are critical to carbohydrate recognition by SP-D, the likely explanation for the *myo*-Inositol > Glc affinity profile is that the number of possible binding orientations is greater for *myo*-Inositol than for Glc.

Automated docking suggests that R343V mutant hSP-D binds α-Me-Glc and α-Me-GlcNAc in two orientations

To probe the role of Arg^{343} in monosaccharide recognition, a computational model of R343V hSP-D was constructed, and both α-Me-Glc and α-Me-GlcNAc were docked into it. Their docking energies are presented in Table II. These docking simulations suggest that R343V hSP-D binds α-Me-Glc in ECR and ECL orientations, similar to wild-type hSP-D. However, unlike the wild-type protein, the docking studies suggest that R343V hSP-D also binds α-Me-GlcNAc in both ECR and ECL orientations (Fig. 4). These calculations support the idea that Arg^{343} prohibits the ECL orientation for α-Me-GlcNAc in wild-type hSP-D.

R343V mutant hSP-D recognizes Glc and GlcNAc equally

To test the hypothesis that steric clashes between the *N*-linked acetyl group at the 2-position on the GlcNAc sugar ring and the side chain of Arg343 are responsible for the Glc > GlcNAc recognition specificity, an R343V hSP-D was constructed and purified by mannose-Sepharose affinity chromatography and elution with EDTA, identical to the method used for wild-type hSP-D purification. The successful purification of the R343V mutant demonstrates that it retains calcium-dependent carbohydrate binding, and suggests proper folding of the mutant protein's

CRD.

Analysis of the purified R343V mutant by SDS-PAGE and Western blotting showed the presence of a significant amount of mutant SP-D monomer and dimer (in addition to the expected trimer) under denaturing, nonreducing conditions (not shown). Wild-type hSP-D migrated predominately as trimer with only minor amounts of monomer and dimer under similar conditions. The reasons for the unexpected migration are most likely related to the altered carbohydrate recognition profile discussed below. It is known that single CRD units of SP-D have only weak affinity for carbohydrate ligands (Ogasawara and Voelker 1995b) while fully oligomerized SP-D binds the same ligands with much higher affinity due to increased avidity of the multimeric molecule. The R343V mutant hSP-D displays higher affinity for monosaccharide binding than its wild-type counterpart that results in successful affinity purification of the incompletely oligomerized protein.

α-Me-Glc and α-Me-GlcNAc were tested for their ability to inhibit R343V hSP-D binding to mannose-Sepharose beads. R343V hSP-D has increased apparent affinity for both α-Me-Glc and α-Me-GlcNAc (Table I). These findings are consistent with the R343V mutant having a less restricted binding site due to the substitution of a smaller side chain for a larger one, thus relieving the strain between the protein and ligand (Shrive et al., 2003). Significantly, the mutant showed K_is reduced by a factor of three for α-Me-Glc, and reduced by a factor of nine for α-Me-GlcNAc relative to the wild-type hSP-D. Thus, the R343V mutation not only increases affinity for monosaccharides generally, but specifically increases affinity for α-Me-GlcNAc. Therefore, the K_i data in Table I agree with the structural superimposition and automated docking observations that suggest restricted recognition of α-Me-GlcNAc by wild-type hSP-D.

Future work

Given the physical and chemical nature of Arg^{343}, and its close proximity to the carbohydrate-binding site, it will be interesting to test Arg^{343} mutants for recognition of other targets. For example, SP-D binds negatively-charged ligands such as phosphatidylinositol (Ogasawara et al., 1992) and bacterial lipopolysaccharide (Kuan et al., 1992). It is possible that electrostatic interactions play a role in recognition of these ligands, and that the positive charge associated with Arg^{343} contributes to binding specificity. Additionally, the conservation of this residue is noteworthy. The residue corresponding to Arg^{343} is conserved as either Arg or Lys in all SP-D and SP-A sequences reported, as Val or Ile in the MBPs and as Glu or Asp in the selectins. Clearly, this site has been the subject of selective pressure. Understanding the role of this residue in C-type lectin ligand recognition specificity will be crucial to understanding the functions of these proteins.

Conclusion

We have demonstrated that Arg^{343} plays a key role in dictating the affinity of hSP-D for substituted monosaccharides and that R343V is a gain-of-function mutation for α-Me-GlcNAc binding by the protein. Additionally, we conclude that the identity of the residue corresponding to Arg^{343} is critical in determining the ligand recognition specificity differences between SP-D and MBP. This residue likely plays a key role in the ligand recognition profile of other C-type lectins as well.

The recently reported structures of MBP-A and MBP-C bound to single ligands in multiple orientations (Ng et al., 2002) demonstrate that multi-orientation binding is possible for C-type lectins. Although not all ligands may be bound in this manner, multi-orientation binding may be a general property of C-type lectins and be important for microbial recognition by the proteins.

For example, multivalent binding of C-type lectins is thought to be required for productive microbial recognition. Multivalent binding is a result of multiple CRDs on a single lectin molecule recognizing multiple ligand sites on the cell surface. This multivalent binding contributes to the affinity and specificity of the recognition event. If some of the ligands are presented in suboptimal orientations we propose that they could still be bound by the lectin and contribute to the overall binding affinity.

Materials and methods

Materials

The FITC-conjugated $F(ab')_2$ fragment of donkey anti-rabbit IgG was from Jackson ImmunoResearch (West Grove, PA). Unless otherwise noted, all other materials were from Sigma (St. Louis, MO). *A. fumigatus* conidia, recombinant hSP-D expressed in CHO K1 cells, and polyclonal rabbit anti-hSP-D antibody were prepared exactly as described previously (Allen et al., 1999).

Mutagenesis of hSP-D

For site-directed mutagenesis, the cDNA encoding hSP-D was cloned into the pGEM-7Zf plasmid (Promega, Madison, WI) at 5' *Hin*dIII and 3' *Eco*RI sites. Mutations coding for the Arg343→Val (R343V) substitution were introduced using the QuickChange site-directed mutagenesis kit (Stratagene, La Jolla, CA) and the mutagenic primers TGGCAAGTGGAATGACGTGGCTTGTGGAGAAAAGCGTC and GACGCTTTTCTCCACAAGCCACGTCATTCCACTTGCCA, where the underlined nucleotides indicate DNA mismatches coding for the substitution. The presence of the mutation

was verified by DNA sequencing. The cDNA encoding the mutant hSP-D gene was then cloned into pEE14 on 5' *Hin*dIII and 3' *Eco*RI sites and transfected into CHO K1 cells using LipofectAMINE (GibcoBRL, Rockville, MD). Clones were selected using cloning cylinders. A single high-expressing clone was selected and the mutant hSP-D was purified as previously described using mannose-Sepharose affinity chromatography (Allen et al., 1999). The wild-type and mutant proteins used in this work were judged to be pure by SDS-PAGE, Coomassie blue staining, and Western blotting.

Binding of hSP-D to mannose-Sepharose beads

Varying concentrations of hSP-D were incubated with 0.52 mg dry weight (the pellet from 20 μL of a 50% aqueous suspension) mannose-Sepharose beads for 1 h at 25°C in CBB (130 mM NaCl, 13 mM NaN_3, 5 mM KCl, 3 mM sodium phosphate buffer, 10 mM HEPES, 2 mM $CaCl_2$, and 1 mM $MgSO_4$ at pH 7.4) containing 1% heat-inactivated and dialyzed fetal bovine serum. The total binding volume was 0.1 mL. The mannose-Sepharose beads were then washed four times with CBB to remove unbound hSP-D (centrifugation at 510 x *g* for 2 min) and incubated for 1 h at 25°C with 0.3 mL total volume of 20 μg/mL rabbit polyclonal anti-hSP-D IgG in CBB. The beads were again washed four times with CBB to remove unbound anti-hSP-D (centrifugation at 510 X *g* for 2 min) and incubated for 1 h at 25°C with 0.24 mL total volume of 20 μg/mL FITC conjugated $F(ab')_2$ fragment of donkey anti-rabbit IgG in CBB. The beads were washed three times with CBB to remove unbound secondary antibody (centrifugations at 510 x *g* for 2 min) and suspended in 1 mL CBB. For analysis, 0.8 mL of the mannose-Sepharose bead suspension was examined for FITC fluorescence by using a Hitachi F-2000 fluorescence spectrophotometer with an excitation wavelength of 492 nm and an emission wavelength of 520 nm. Since the mannose-Sepharose beads sedimented rapidly in the cuvette, three readings were

taken for each sample 10 s after resuspension by repeated pipetting.

Inhibition of hSP-D binding to Aspergillus fumigatus *or mannose-Sepharose*

For inhibition of binding to mannose-Sepharose we used methylated derivatives of the anomeric carbon to be consistent with previous MBP-C structural work (Ng et al., 1996). The inhibitor concentration yielding 50% inhibition (IC_{50}) of hSP-D binding to *A. fumigatus* was determined exactly as before (Allen et al., 1999), using 20 μg/mL hSP-D. For inhibition of binding to mannose-Sepharose beads, we found the maximal protein binding concentration ([C_{max}]) by double-reciprocal analysis of the binding data for each protein between 0.3 and 17 μg/ml. In subsequent inhibition experiments we used [C_{max}]/7, which were 1.7 and 2.7 μg/mL for wild-type and R343V hSPD, respectively. The dissociation constant of binding between protein and mannose-Sepharose beads (K_d) was determined by global least-squares fitting of equation 1 to the concentration dependence of the fluorescence intensity.

$$F = \frac{F_{\max} K_a [\text{hSP-D}]}{1 + K_a [\text{hSP-D}]} \qquad (1)$$

A molecular weight of 43 kDa was used for both WT hSPD and the R343V mutant. F_{max} is a constant equal to the maximal fluorescence extrapolated for infinite hSPD concentration, and $K_a = 1/K_d$. To determine the IC_{50} values for the various inhibitors, the proteins were preincubated with each inhibitor for 15 min at 25°C in CBB in a total volume of 0.1 mL. The hSP-D plus inhibitor mixture was then added to mannose-Sepharose beads, and binding was performed as described above. The inhibition constant K_i was calculated from the IC_{50} using equation 2 (Cheng and Prusoff 1973):

$$K_i = \frac{IC_{50}}{1 + \frac{[\text{hSP} - \text{D}]}{K_d}} \quad (2)$$

Errors in K_i were computed using standard propagation of error analysis (Bevington 1969), assuming a 10% error for the protein concentration.

Computational methods

Automated docking simulations were performed using the AutoDock 3.06 suite of programs (Scripps Research Institute, La Jolla, CA) and the Lamarkian genetic algorithm (Morris et al., 1998) as described previously (Allen et al., 2001). Grid files were prepared as described (Allen et al., 2001) except that hydrogen atoms were added to the receptor protein files for MBP-C [PDB accession code 1rdl chain 2 (Ng et al., 1996)] and hSP-D [PDB accession code 1b08 chain A (Hakansson et al., 1999)] using the Insight 2000/Biopolymer module (Accelrys, San Diego, CA). The ligands were placed near the binding site for each receptor using the corresponding Man 9 coordinates from PDB accession code 2msb (Weis et al., 1992) as before (Laederach et al., 1999).

Ligand files were built by using PC-MODEL (Serena Software, Bloomington, IN) and minimized by using the MM3 force field (Allinger et al., 1990). Internal coordinates were then exported to MOPAC (Stewart 1990) and used as starting points for the geometrical optimization procedure. For ligands lacking nitrogen (*e.g.* α-Me-Glc), the AM3 hamiltonian was used. For ligands containing nitrogen (*e.g.* α-Me-GlcNAc), the MNDO hamiltonian was used along with a molecular mechanics correction (MMOK) that produced better agreement with known carbon-nitrogen bond geometry. The final optimized structure and the corresponding partial atomic charges determined by MOPAC were subsequently used in docking simulations.

The R343V hSP-D model was constructed by using the BioPolymer and CHARMM (Harvard University, Cambridge, MA) modules of Insight 2000. Side-chain torsions of Arg^{343} were taken from the wild-type structure [PDB accession code 1b08, chain A (Hakansson et al., 1999)]. The residue was mutated using the MUTATE command in the Biopolymer module, which automatically selected the optimal rotamer from a standard library. The mutated residue torsions were then adjusted to correspond to those measured in the wild-type protein. Harmonic positional constraints (K = 1000 kcal/mol) were placed on all protein atoms except those on the side chain of the mutated residue. The structure was then minimized using the conjugate-gradient algorithm implemented in CHARMM (Brooks et al., 1983). Generalized Born (GBORN) electrostatics (Dominy and Brooks 1999) were used to approximate solvent effects, since Arg^{343} is a surface-exposed residue. The resulting structure was used to generate grid files for AutoDock input as described above.

As previously discussed, the binding interaction is mediated primarily by four hydrogen bonds between the protein and the ligand (Allen et al., 2001). Thus many nonspecific binding modes have only slight differences in binding energies. Therefore, energetic and hydroxyl group distance criteria was used for monosaccharides as before (Allen et al., 2001). Established molecular modeling force fields such as CHARMM (Brooks et al., 1983) or AMBER (Cornell et al., 1995) are excellent for optimizing bound conformations. However, their ability to rank the binding energies of ligands docked in significantly different modes is limited (Morris et al., 1998). The latest AutoDock release uses a new empirical binding free energy function to correct this problem. However, this new function is not parameterized for metal atoms, and therefore it was not possible to apply it to our docking calculations. Instead, we used the AutoDock 1.0 parameters for the intra- and inter-molecular potential. The calculated docking energies

therefore represent in vacuo enthalpies of formation and their correlation with free energies of formation (or K_i) is limited (Morris et al., 1998). A docked conformer represents a sterically and electrostatically optimal conformation, but the determination of the free energy of formation of the complex must be carried out experimentally. For example, in docking α-Me-Glc, we systematically obtained a minimal energy structure in an orientation where the 2- and 3-OH groups of the ligand were positioned near the protein and occupied the corresponding positions of the 3- and 4-OH groups previously noted for the binding of mannose by MBP-A and MBP-C. This orientation resembles that previously obtained for α-D-glucopyranose (Fig. 3 of ref. (Allen et al., 2001)). Our previously reported experimental results using glucosyl polysaccharides as SP-D inhibitors suggest that this orientation is possible but is not as favorable as the 3-OH/4-OH group binding orientations (Allen et al., 2001). Thus, although we obtained 2-OH/3-OH bound ligand structures with docking energies similar to the 3- and 4-OH structures shown, they were not considered.

Other methods

Protein concentrations were determined with the bicinchoninic acid assay using bovine serum albumin standard according to the manufacturer's instructions (Pierce, Rockford, IL). Structural superpositions were performed using Swiss-Pdb viewer (Guex and Peitsch 1997). Figures were prepared using MolMol (Koradi et al., 1996).

Acknowledgements

We thank Erika Crouch for providing the CHO-K1 cells expressing wild-type hSP-D and Amanda Evans for excellent technical assistance. The experimental part of this work was performed in the Lord and Taylor Laboratory for Lung Biochemistry at National Jewish Medical and Research Center. This work was supported by grants from the Environmental Protection

Agency (R825702 to R. J. M.) and the National Institutes of Health (HL-29891 to R. J. M., and HL-45286 and HL-66235 to D. R. V.). M. J. A. was funded by the Andrew Goodman Fellowship at National Jewish Medical and Research Center and A. L. was mainly supported by National Science Foundation IGERT Program Award (9972653).

Abbreviations

BSA, bovine serum albumin; CBB, calcium binding buffer; CRD, carbohydrate recognition domain; ECL, extracyclic carbon left; ECR, extracyclic carbon right; EDTA, ethylenediaminetetraacetic acid; FITC, fluorescein isothiocyanate; Glc, D-glucopyranose; GlcNAc, 2-acetamido-2-deoxy-D-glucose (*N*-acetyl-D-glucosamine); hSP-D, human surfactant protein D; IC_{50}, 50% inhibitory concentration; IgG, immunoglubulin G; MBP-A, mannose binding protein A; MBP-C, mannose binding protein C; α-Me-Glc, methyl α-D-glucopyranoside; β-Me-Glc, methyl β-D-glucopyranoside; α-Me-GlcNAc, methyl 2-acetamido-2-deoxy-α-D-glucopyranoside; α-Me-Man, methyl α-D-mannopyranoside; R343V, $Arg^{343} \rightarrow Val$ mutation; rmsd, root-mean-square deviation; rSP-D, rat surfactant protein D; SDS-PAGE, sodium dodecyl sulfate-polyacrylamide gel electrophoresis; SEM, standard error of the mean; SP-A, surfactant protein A; SP-D, surfactant protein D

References

Allen,M.J., Harbeck,R., Smith,B., Voelker,D.R., and Mason,R.J. (1999) Binding of rat and human surfactant proteins A and D to Aspergillus fumigatus conidia. *Infect.Immun.*, **67**,4563-4569.

Allen,M.J., Laederach,A., Reilly,P.J., and Mason,R.J. (2001) Polysaccharide recognition by surfactant protein D: novel interactions of a C-type lectin with nonterminal glucosyl residues. *Biochemistry*, **40**,7789-7798.

Allinger,N.L., Rahman,M., and Li,J.-H. (1990) A molecular mechanics force field (MM3) for alcohols and ethers. *J.Am.Chem.Soc.*, **112**,8293-8307.

Bevington,P.R. (1969) Data reduction and error analysis for physical sciences. McGraw-Hill, New York, NY.

Blanck,O., Iobst,S.T., Gabel,C., and Drickamer,K. (1996) Introduction of selectin-like binding specificity into a homologous mannose-binding protein. *J.Biol.Chem.*, **271**,7289-7292.

Brooks,B.R., Bruccoleri,R.E., Olafson,B.D., States,D.J., Swaminathan,S., and Karplus,M. (1983) CHARMM: A program for macromolecular energy minimization and dynamics calculations. *J.Comp.Chem.*, **4**,187-217.

Burrows,L., Iobst,S.T., and Drickamer,K. (1997) Selective binding of N-acetylglucosamine to the chicken hepatic lectin. *Biochem.J.*, **324**,673-680.

Cheng,Y., Prusoff,W.H. (1973) Relationship between the inhibition constant (KI) and the concentration of inhibitor which causes 50 percent inhibition (IC50) of an enzymatic reaction. *Biochem.Pharmacol.*, **22**,3099-3108.

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) A second generation force field for the simulation of proteins and nucleic acids. *J.Am.Chem.Soc.*, **117**,5179-5197.

Crouch,E.C. (1998) Collectins and pulmonary host defense. *Am.J.Respir.Cell Mol.Biol.*, **19**,177-201.

Crouch,E.C. (2000) Surfactant protein-D and pulmonary host defense. *Respir.Res.*, **1**,93-108.

de Wetering,J.K., van Eijk,M., van Golde,L.M., Hartung,T., van Strijp,J.A., and Batenburg,J.J. (2001) Characteristics of surfactant protein a and d binding to lipoteichoic acid and peptidoglycan, 2 major cell wall components of gram-positive bacteria. *J.Infect.Dis.*, **184**,1143-1151.

Dominy,B., Brooks,C.L.I. (1999) Development of a generalized born model parameterization for proteins and nucleic acids. *J.Phys.Chem.*, **103**,3765-3773.

Drickamer,K. (1992) Engineering galactose-binding activity into a C-type mannose-binding protein. *Nature*, **360**,183-186.

Feinberg,H., Mitchell,D.A., Drickamer,K., and Weis,W.I. (2001) Structural basis for selective recognition of oligosaccharides by DC-SIGN and DC-SIGNR. *Science*, **294**,2163-2166.

Graves,B.J., Crowther,R.L., Chandran,C., Rumberger,J.M., Li,S., Huang,K.S., Presky,D.H., Familletti,P.C., Wolitzky,B.A., and Burns,D.K. (1994) Insight into E-selectin/ligand interaction from the crystal structure and mutagenesis of the lec/EGF domains. *Nature*, **367**,532-538.

Guex,N., Peitsch,M.C. (1997) SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis*, **18**,2714-2723.

Hakansson,K., Lim,N.K., Hoppe,H.J., and Reid,K.B. (1999) Crystal structure of the trimeric alpha-helical coiled-coil and the three lectin domains of human lung surfactant protein D. *Structure.Fold.Des*, **7**,255-264.

Head,J.F., Mealy,T.R., McCormack,F.X., and Seaton,B.A. (2003) Crystal structure of trimeric carbohydrate recognition and neck domains of surfactant protein A. *J.Biol.Chem.*, **278**,43254-43260.

Hickling,T.P., Bright,H., Wing,K., Gower,D., Martin,S.L., Sim,R.B., and Malhotra,R. (1999) A recombinant trimeric surfactant protein D carbohydrate recognition domain inhibits respiratory syncytial virus infection in vitro and in vivo. *Eur.J.Immunol.*, **29**,3478-3484.

Hitchen,P.G., Mullin,N.P., and Taylor,M.E. (1998) Orientation of sugars bound to the principal C-type carbohydrate-recognition domain of the macrophage mannose receptor. *Biochem.J.*, **333**,601-608.

Kogan,T.P., Revelle,B.M., Tapp,S., Scott,D., and Beck,P.J. (1995) A single amino acid residue can determine the ligand specificity of E- selectin. *J.Biol.Chem.*, **270**,14047-14055.

Koradi,R., Billeter,M., and Wuthrich,K. (1996) MOLMOL: a program for display and analysis of macromolecular structures. *J.Mol.Graph.*, **14**,51-32.

Kuan,S.F., Rust,K., and Crouch,E. (1992) Interactions of surfactant protein D with bacterial lipopolysaccharides. Surfactant protein D is an Escherichia coli- binding protein in bronchoalveolar lavage. *J.Clin.Invest*, **90**,97-106.

Kuroki,Y., Voelker,D.R. (1994) Pulmonary surfactant proteins. *J.Biol.Chem.*, **269**,25943-25946.

Laederach,A., Dowd,M.K., Coutinho,P.M., and Reilly,P.J. (1999) Automated docking of maltose, 2-deoxymaltose, and maltotetraose into the soybean beta-amylase active site. *Proteins*, **37**,166-175.

Lee,R.T., Ichikawa,Y., Fay,M., Drickamer,K., Shao,M.C., and Lee,Y.C. (1991) Ligand-binding characteristics of rat serum-type mannose-binding protein (MBP-A). Homology of binding site architecture with mammalian and chicken hepatic lectins. *J.Biol.Chem.*, **266**,4810-4815.

LeVine,A.M., Bruno,M.D., Huelsman,K.M., Ross,G.F., Whitsett,J.A., and Korfhagen,T.R. (1997) Surfactant protein A-deficient mice are susceptible to group B streptococcal infection. *J.Immunol.*, **158**,4336-4340.

LeVine,A.M., Gwozdz,J., Stark,J., Bruno,M., Whitsett,J., and Korfhagen,T. (1999a) Surfactant protein-A enhances respiratory syncytial virus clearance in vivo. *J.Clin.Invest*, **103**,1015-1021.

LeVine,A.M., Kurak,K.E., Bruno,M.D., Stark,J.M., Whitsett,J.A., and Korfhagen,T.R. (1998) Surfactant protein-A-deficient mice are susceptible to Pseudomonas aeruginosa infection. *Am.J.Respir.Cell Mol.Biol.*, **19**,700-708.

LeVine,A.M., Kurak,K.E., Wright,J.R., Watford,W.T., Bruno,M.D., Ross,G.F., Whitsett,J.A., and Korfhagen,T.R. (1999b) Surfactant protein-A binds group B streptococcus enhancing phagocytosis and clearance from lungs of surfactant protein-A-deficient mice. *Am.J.Respir.Cell Mol.Biol.*, **20**,279-286.

LeVine,A.M., Whitsett,J.A., Gwozdz,J.A., Richardson,T.R., Fisher,J.H., Burhans,M.S., and Korfhagen,T.R. (2000) Distinct effects of surfactant protein A or D deficiency during bacterial infection on the lung. *J.Immunol.*, **165**,3934-3940.

Madan,T., Kishore,U., Singh,M., Strong,P., Hussain,E.M., Reid,K.B., and Sarma,P.U. (2001) Protective role of lung surfactant protein D in a murine model of invasive pulmonary aspergillosis. *Infect.Immun.*, **69**,2728-2731.

Mason,R.J., Greene,K., and Voelker,D.R. (1998) Surfactant protein A and surfactant protein D in health and disease. *Am.J.Physiol*, **275**,L1-13.

McCormack,F.X., Kuroki,Y., Stewart,J.J., Mason,R.J., and Voelker,D.R. (1994) Surfactant protein A amino acids Glu195 and Arg197 are essential for receptor binding, phospholipid aggregation, regulation of secretion, and the facilitated uptake of phospholipid by type II cells. *J.Biol.Chem.*, **269**,29801-29807.

Morris,G.M., Goodsell,D.S., Halliday,R.S., Huey,R., Hart,W.E., Belew,R.K., and Olson,A.J. (1998) Automated docking using a Lamarkian genetic algorithm and an emperical binding free energy function. *J.Comp.Chem.*, **19**,1639-1662.

Ng,K.K., Drickamer,K., and Weis,W.I. (1996) Structural analysis of monosaccharide recognition by rat liver mannose- binding protein. *J.Biol.Chem.*, **271**,663-674.

Ng,K.K., Kolatkar,A.R., Park-Snyder,S., Feinberg,H., Clark,D.A., Drickamer,K., and Weis,W.I. (2002) Orientation of Bound Ligands in Mannose-binding Proteins. *J.Biol.Chem*, **277**,16088-16095.

Ng,K.K., Weis,W.I. (1997) Structure of a selectin-like mutant of mannose-binding protein complexed with sialylated and sulfated Lewis(x) oligosaccharides. *Biochemistry*, **36**,979-988.

Ogasawara,Y., Kuroki,Y., and Akino,T. (1992) Pulmonary surfactant protein D specifically binds to phosphatidylinositol. *J.Biol.Chem.*, **267**,21244-21249.

Ogasawara,Y., Voelker,D.R. (1995a) Altered carbohydrate recognition specificity engineered into surfactant protein D reveals different binding mechanisms for phosphatidylinositol and glucosylceramide. *J.Biol.Chem.*, **270**,14725-14732.

Ogasawara,Y., Voelker,D.R. (1995b) The role of the animo-terminal domain and collagenous region in the structure and function of rat surfactant protein D. *J.Biol.Chem*, **270**,19052-19058.

Persson,A., Chang,D., and Crouch,E. (1990) Surfactant protein D is a divalent cation-dependent carbohydrate- binding protein. *J.Biol.Chem.*, **265**,5755-5760.

Shrive,A.K., Tharia,H.A., Strong,P., Kishore,U., Burns,I., Rizkallah,P.J., Reid,K.B., and Greenhough,T.J. (2003) High-resolution structural insights into ligand binding and immune cell recognition by human lung surfactant protein D. *J.Mol.Biol.*, **331**,509-523.

Simanek,E.E., McGarvey,G.J., Jablonowski,J.A., and Wong,C.-H. (1998) Selectin-carbohydrate interactions: from natural ligands to designed mimics. *Chem.Rev.*, **98**,833-862.

Somers,W.S., Tang,J., Shaw,G.D., and Camphausen,R.T. (2000) Insights into the molecular basis of leukocyte tethering and rolling revealed by structures of P- and E-selectin bound to SLex and PSGL-1. *Cell*, **103**,467-479.

Stewart,J.J. (1990) MOPAC: a semiempirical molecular orbital program. *J.Comput.Aided Mol.Des*, **4**,1-105.

Vuk-Pavlovic,Z., Standing,J.E., Crouch,E.C., and Limper,A.H. (2001) Carbohydrate recognition domain of surfactant protein D mediates interactions with Pneumocystis carinii glycoprotein A. *Am.J.Respir.Cell Mol.Biol.*, **24**,475-484.

Weis,W.I., Drickamer,K., and Hendrickson,W.A. (1992) Structure of a C-type mannose-binding protein complexed with an oligosaccharide. *Nature*, **360**,127-134.

Wright,J.R. (1997) Immunomodulatory functions of surfactant. *Physiol Rev.*, **77**,931-962.

Legends to figures

Fig. 1. hSP-D crystal structure (green) superimposed on the crystal structure of MBP-C (red) complexed with α-Me-GlcNAc. SP-D residue numbering: no parentheses; MBP-C residue numbering: parentheses. The methyl group of the bound ligand is not visible. Carbon atom numbering for the ligand is indicated. Coordinates for MBP-C and SP-D are from PDB accession codes 1rdn (Ng et al., 1996) and 1b08 (Hakansson et al., 1999), respectively.

Fig. 2. Structures of α-Me-Glc docked into wild-type hSP-D (Hakansson et al., 1999) A) ECL orientation; B) ECR orientation. Oxygen: light gray spheres. Calcium: dark gray sphere. Nitrogen: black spheres. Dashed lines represent hydrogen bonds.

Fig. 3. Structure of α-Me-GlcNAc docked into wild-type hSP-D (Hakansson et al., 1999). Atomic color scheme as in Fig. 2.

Fig. 4. Structures of α-Me-GlcNAc docked into R343V hSP-D. A) ECR orientation; B) ECL orientation. Atomic color scheme as in Fig. 2.

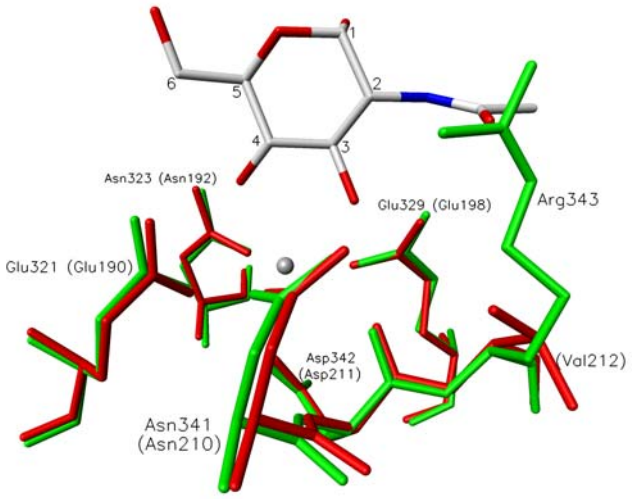
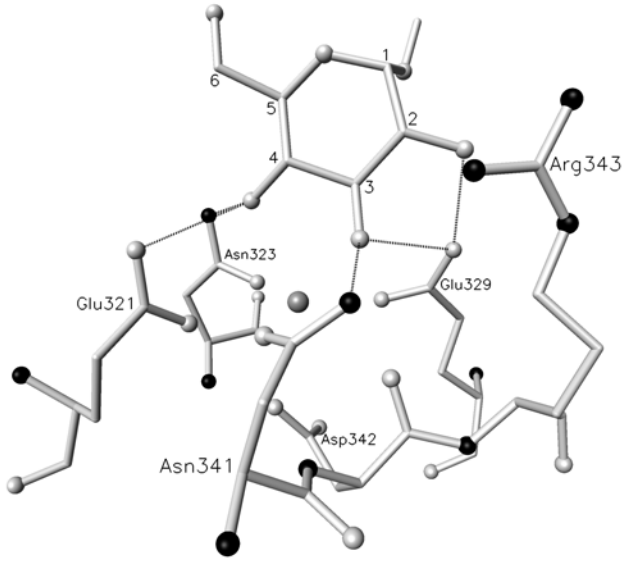

1
2
3
4
5
6
Asn323 (Asn192)
Glu329 (Glu198)
Arg343
Glu321 (Glu190)
Asp342
(Asp211)
(Val212)
Asn341
(Asn210)

Fig. 1

A


1
2
3
4
5
6
Arg343
Asn323
Glu321
Glu329
Asp342
Asn341

B

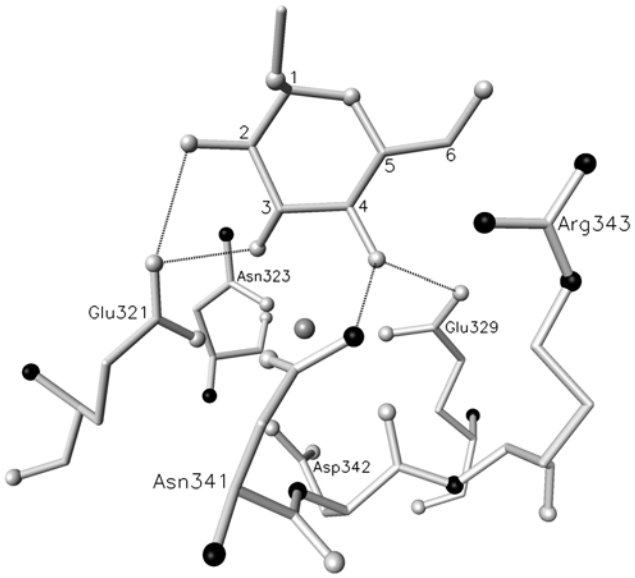

1
2
3
4
5
6
Arg343
Asn323
Glu321
Glu329
Asp342
Asn341

Fig. 2

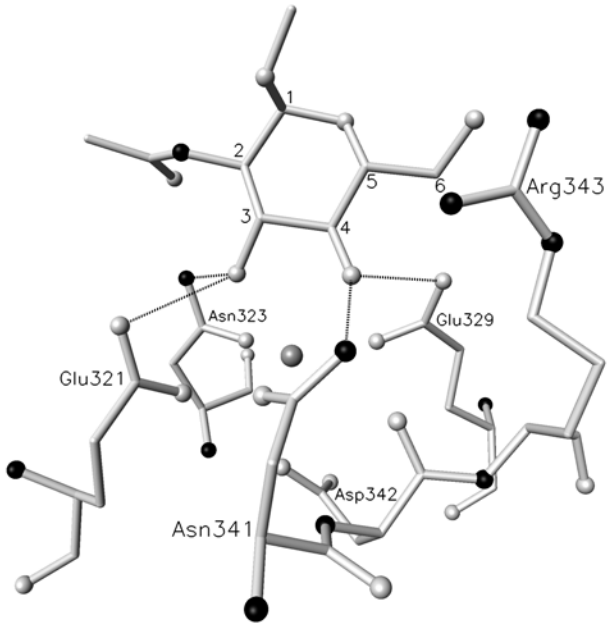
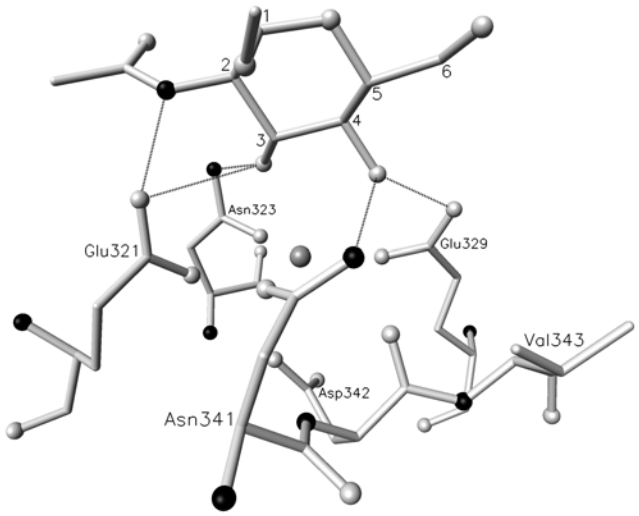


Fig. 3

A


1
2
3
4
5
6
Asn323
Glu321
Glu329
Val343
Asp342
Asn341

B

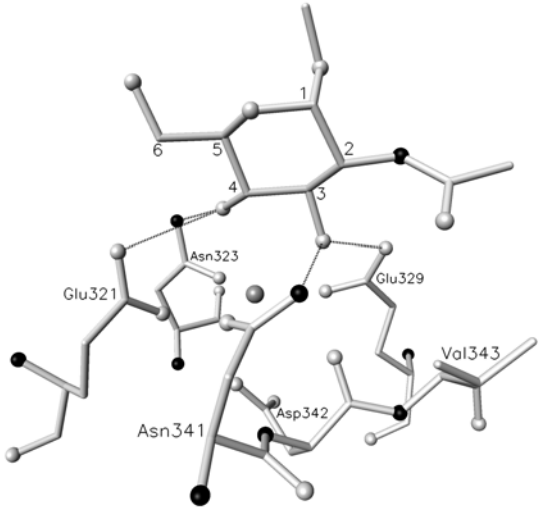

1
2
3
4
5
6
Asn323
Glu321
Glu329
Val343
Asp342
Asn341

Fig. 4

Table I

Wild-type and R343V hSP-D binding to mannose-Sepharose and inhibition of binding by monosaccharides. Dissociation constants (K_d) and inhibition constants (K_i) are shown.

SP-D form	K_d (nM)	K_i (mM)			
		Glc	β-Me-Glc	α-Me-Glc	α-Me-GlcNAc
Wild-type	120 ± 30[a]	29 ± 5	37 ± 4	30 ± 2	89 ± 14[b]
R343V	150 ± 50	ND[c]	ND	9 ± 2	10 ± 2

[a] Data are the average ± standard error of three experiments unless otherwise indicated.

[b] Datum is the average ± range of two experiments.

[c] Not determined.

Table II

Calculated energies for docked monosaccharides

Ligand	Receptor protein	Optimal docked energy (kcal/mol)		
		ECL	ECR	Figure
α-Me-GlcNAc	MBP-C	–71.12	–69.64	Not Shown
α-Me-Glc	Wild-type hSP-D	–65.22	–63.32	2
α-Me-GlcNAc	Wild-type hSP-D		–66.39	3
α-Me-Glc	R343V hSP-D	–57.41	–57.79	Not Shown
α-Me-GlcNAc	R343V hSP-D	–60.33	–63.87	4