J Comput Aided Mol Des (2010) 24:423–431
DOI 10.1007/s10822-010-9353-5

Computational studies on the interaction of ABO-active saccharides with the norovirus VA387 capsid protein can explain experimental binding data

Chaitanya A. K. Koppisetty · Waqas Nasir · Francesco Strino · Gustaf E. Rydell · Göran Larson · Per-Georg Nyholm

Received: 8 December 2009 / Accepted: 29 March 2010 / Published online: 21 April 2010

Abstract Norovirus strains are known to cause recurring epidemics of winter vomiting disease. The crystal structure of the capsid protein of VA387, a representative of the clinically important GII.4 genocluster, was recently solved in complex with histo-blood group A- and B-trisaccharides. However, the VA387 strain is known to bind also to other natural carbohydrates for which detailed structural information of the complexes is not available. In this study we have computationally explored the fit of the VA387 with a set of naturally occurring carbohydrate ligands containing a terminal α1,2-linked fucose. MD simulations both with explicit and implicit solvent models indicate that type 1 and 3 extensions of the ABO-determinant including ALe^b and BLe^b pentasaccharides can be well accommodated in the site. Scoring with Glide XP indicates that the downstream extensions of the ABO-determinants give an increase in binding strength, although the α1,2-linked fucose is the single strongest interacting residue. An error was discovered in the geometry of the GalNAc-Gal moiety of the published crystal structure of the A-trisaccharide/VA387 complex. The present modeling of the complexes with histo-blood group A-active structures shows some contacts which provide insight into mutational data, explaining the involvement of I389 and Q331. Our results can be applicable in structure-based design of adhesion inhibitors of noroviruses.

Keywords Norovirus · Blood group determinants · ABO-saccharides · Molecular docking · Molecular dynamics

Electronic supplementary material The online version of this article (doi:10.1007/s10822-010-9353-5) contains supplementary material, which is available to authorized users.

C. A. K. Koppisetty · F. Strino · P.-G. Nyholm
Biognos AB, Generatorsgatan 1, 40274 Gothenburg, Sweden

C. A. K. Koppisetty · W. Nasir
Department of Computer Science and Engineering, Chalmers University of Technology, 41296 Gothenburg, Sweden

W. Nasir · F. Strino · P.-G. Nyholm (✉)
Institute of Biomedicine, University of Gothenburg, 40530 Gothenburg, Sweden
e-mail: per-georg.nyholm@medkem.gu.se; per-georg.nyholm@biognos.se

G. E. Rydell · G. Larson
Department of Clinical Chemistry and Transfusion Medicine, Sahlgrenska University Hospital, Gothenburg, Sweden

Introduction

Noroviruses have been estimated to cause about half of all outbreaks of gastroenteritis in developed countries [1]. Volunteer [2, 3], outbreak [4–6] and antibody titer [7] studies have demonstrated that secretor status, controlled by the α1,2-fucosyltransferase *FUT2* gene, is linked to susceptibility to norovirus infection. Furthermore, in vitro binding studies have demonstrated that non-secretors (*FUT2* -/-) lack attachment factors for most clinically important norovirus strains [8–11].

The *FUT2* gene codes for one of two active human α1,2-fucosyltransferases, the other being coded by the *FUT1* (*H*) gene. The *FUT2* gene is mainly expressed in epithelial (mucosal) tissues and the corresponding transferase adds Fuc in an α1,2 position to a terminal Gal of preferentially the type 1 and 3 chain structures [12]. In this study we have focused on norovirus capsid protein interactions with carbohydrate structures specifically dependent on the secretor gene for their biosynthesis.

Recently crystal structures of the outermost part of the capsid protein in complex with glycan ligands have been

solved for the Norwalk [13, 14] and VA387 [15] strains, representing genoclusters GI.1 and GII.4, respectively. Interestingly these two strains were demonstrated to use different binding sites. The VA387 strain bound to histo-blood group A- and B-trisaccharides in similar orientations in the same binding site. No crystal structures of complexes of the capsid protein with other ABO-saccharides containing the secretor gene dependent Fucα1,2Gal structure, have to our knowledge, been reported. Detailed site-directed mutagenesis has, however, demonstrated that the trisaccharide binding site of VA387 is responsible for adhesion to saliva from secretor positive O, A and B individuals [16]. In view of the broad binding specificity of GII.4 strains [8, 17–19] and their clinical importance [20] we have focused our work on the VA387 binding site. With an aim of providing structural insights into VA387 interactions with ABO saccharides, we used molecular mechanics, molecular dynamics (MD) and Glide scoring to investigate the molecular basis for the binding specificity of the VA387 binding site to secretor gene dependent Fucα1,2Gal containing human ABO-active saccharides. The computational results were evaluated in relation to available experimental data.

Materials and methods

Initial structures

The naturally occurring carbohydrates, H type 1 (H-1), H type 2 (H-2), H type 3 (H-3), Lewis b (Le^b), A type 1 (A-1), A type 3 (A-3), A Lewis b (ALe^b), B type 1 (B-1) and B Lewis b (BLe^b) all contain a terminal Fucα1-2Gal moiety (Table 1). Most of them are established ligands for VA387 and some are potential ligands. All these structures were initially prepared from the SWEET2 server [21] and the geometries of the structures were then optimized with MM4 [22] using the GLYGAL program [23]. The optimized initial angles for the glycosidic torsions were in good agreement with the adiabatic maps of the trisaccharide moieties for A- and B- epitopes [24] and disaccharide moieties for other structures (supplementary material). The structures for the B-trisaccharide and the fucose were obtained from the structure of the B-trisaccharide/VA387 complex (2OBT) [15].

Table 1 Carbohydrate structures used in this study

Abbreviation (name)	Structure
A-tri	GalNAcα3(Fucα2)Galβ
B-tri	Galα3(Fucα2)Galβ
H-1 (H type 1 chain)	Fucα2Galβ3GlcNAcβ
H-2 (H type 2 chain)	Fucα2Galβ4GlcNAcβ
H-3 (H type 3 chain)	Fucα2Galβ3GalNAcα
Le^b (Lewis b)	Fucα2Galβ3(Fucα4)GlcNAcβ
A-1 (A type 1 chain)	GalNAcα3(Fucα2)Galβ3GlcNAcβ
A-3 (A type 3 chain)	GalNAcα3(Fucα2)Galβ3GalNAcα
ALe^b (A Lewis b)	GalNAcα3(Fucα2)Galβ3(Fucα4)GlcNAcβ
B-1 (B type 1 chain)	Galα3(Fucα2)Galβ3GlcNAcβ
BLe^b (B Lewis b)	Galα3(Fucα2)Galβ3(Fucα4)GlcNAcβ

Glide dockings

The docking program Glide (Schrödinger LLC, New York) has been successfully applied in the case of carbohydrate docking to antibodies [25] and to lectins [26]. The XP mode of Glide docking version 5.0 was used in this study. The fucose monosaccharide was prepared with Ligprep (version 2.2; Schrödinger LLC, New York, NY) and minimized using the BMIN software of Schrödinger Inc. with the OPLS2001 force field. The biological dimer of the VA387 capsid protein was obtained from the PDB database (2OBT) and the protein was processed using Schrödinger software. Initially all hydrogens were added and the whole structure was minimized in the OPLS2001 force field in order to remove close contacts and clashes using the Impref restrained minimization utility. Later a grid was set up centered on the binding site and a box with a side of 18 Å was defined. The dockings with fucose were performed using Glide XP. The crystallographic pose of the B-trisaccharide in complex with VA387 (2OBT) was minimized in the OPLS2001 force field which has been applied in previous studies on carbohydrates [27]. The Glide XP score was then calculated with the "refine and score" option of Glide. Similarly a Glide XP score was calculated for the A-trisaccharide modeled from the B-trisaccharide in the site. Glide XP dockings for the A- and B-trisaccharides, using the same procedure as for the fucose monosaccharide were unsuccessful, yielding poses with an RMSD > 5Å compared to the crystal structure. This appears to be due to insufficient sampling of the conformations of the carbohydrate structures in Glide or possibly the lack of protein flexibility during docking.

MD simulations

The failure to reproduce the crystallographic poses of the A- and B-trisaccharides using Glide XP led us to try another approach. AMBER provides a good force field for saccharide-protein complexes, but molecular dynamics is generally not suitable for docking studies. Therefore, in view of the well documented fucose binding site of VA387 [15, 16], we attempted a simple superimposition on the fucose followed by minimizations and MD to explore the fit of different histo-blood group ABO and Lewis structures (Table 1). The fucose ring in all the structures with

Fucα1,2Gal linkage was superimposed on the fucose ring of the B-trisaccharide in complex with VA387 using Sybyl (version 8.0; Tripos Inc., St Louis). These protein–sugar complexes were then used as input structures for minimization and MD simulations with AMBER (version 8.0; University of California, San Francisco) with the GLYCAM06 force field for sugars [28] and the ff99SB force field for proteins [29].

Explicit solvent simulations were performed with TIP3P explict water molecules with periodic boundary conditions using a truncated octahedron. The pre-MD minimizations were done in two steps. First, position restraints of 100 kcal mol^{-1} $Å^{-2}$ were applied on all protein and sugar residues whereas the explicit solvent water molecules were allowed to move for a total of 1000 minimization cycles out of which 500 cycles were carried out using steepest descent and the remaining with conjugate gradient algorithm. In the second step of minimization, the whole system was minimized with the relatively weaker position restraints of 50 kcal mol^{-1} $Å^{-2}$ on main chain protein atoms while the rest of the atoms were free to move for a total of 2000 minimization cycles using steepest descent and conjugate gradient methods for 1000 cycles each. The second step was extended for another 2000 minimization cycles with same parameters. MD equilibration was then carried out for 250 ps. Initally, constant volume (NVT) equilibration was carried out for 50 ps with the position restraints of 50 kcal mol^{-1} $Å^{-2}$ on all protein and ligand residues and the temperature was slowly raised from 0 to 300 K. Later, constant pressure (NPT) equilibration with isotropic position scaling was carried out for the rest of 200 ps with the position restraints of 50 kcal mol^{-1} $Å^{-2}$ on the main chain protein atoms only. The temperature was regulated using Langevin dynamics with collision frequency of 1 ps^{-1}. The production MD runs were performed at 300 K and 1 atm for 2.3 ns with the time step of 1 fs. All atoms were allowed to move except the protein main chain for which position restraints of 10 kcal mol^{-1} $Å^{-2}$ were applied. The pressure was kept constant with isotropic position scaling and pressure relaxation time of 2.0 ps. The SHAKE algorithm was applied for bonds involving a hydrogen atom and a 10 Å cut-off was used for calculating non-bonded interactions.

For the implicit solvent simulations, a control MD run with unrestrained Fucα monosaccharide was performed to check the positional stability of fucose without any restraints. The α-fucose retained its orientation in the binding site despite the absence of positional restraints. However, the in absence of explicit solvent during the MD for larger structures, it was considered reasonable to apply positional restraints on the fucose ring. The idea to restrain the α1,2-linked fucose residue is further reinforced by the crystal structure of the B-trisaccharide/VA387 complex in which the α1,2-linked fucose was more clearly discerned from the electron density maps than the other saccharide residues. The generalized Born solvation model [30] was used for both minimization and MD simulations. Position restraints of 50 kcal mol^{-1} $Å^{-2}$ and 500 kcal mol^{-1} $Å^{-2}$ were applied on main chain protein atoms and on the fucose, respectively, in the site. The other atoms were free to move during the minimization cycles. A total of 200–500 cycles of minimization were carried out on each system before the MD runs. The system was then equilibrated gradually from 0 to 300 K using MD for 10 ps with a time step of 1 fs. The position restraints on the fucose ring of the sugars were relaxed from 500 kcal mol^{-1} $Å^{-2}$ to 50 kcal mol^{-1} $Å^{-2}$ during equilibration. This equilibration was then followed by 250 ps of production run at 300 K. During the production run the side chains of protein residues within 10 Å from the ligand were allowed to move, whereas position restraints of 50 kcal mol^{-1} $Å^{-2}$ were applied on the protein main-chain atoms in that area. All other protein atoms beyond 10 Å from the ligand were frozen with position restraints of 50 kcal mol^{-1} $Å^{-2}$. During both the equilibration and the production runs the SHAKE procedure was employed to constrain all solute bonds containing a hydrogen atom. The temperature was regulated by Langevin dynamics with a collision frequency of 1.0 ps^{-1} and a non-bonded cut off of 12 Å was used to truncate the non-bonded pairs. Snapshots of the trajectories were taken every 1 ps.

Scoring of poses from MD simulations

The poses from the MD simulations were scored using the Glide XP scoring function to evaluate the binding energies of the ligands. For each ligand, 6 snapshots were taken with an interval of 10 ps from the last parts of the 2.3 ns and the 250 ps simulations for explicit and implicit solvent MD trajectories, respectively. The Glide parameters used here are based on the OPLS2001 force field. Since a snapshot from the MD simulation might not represent a minimal state in the OPLS2001 force field, each protein–ligand complex was minimized in the OPLS2001 force field prior to the grid calculations. The snapshots from the explicit solvent simulations for A- and B-trisaccharides were minimized (including the water molecules) in the OPLS2001 force field using the PRCG (Polak-Ribiere Conjugate Gradient) method for a maximum of 5000 iterations and a gradient convergence threshold of 0.05. The water mediated interactions were analyzed visually in Maestro (version 9.0; Schrödinger LLC, New York). The water molecules were then removed and a Glide grid was calculated with a grid size of 18 Å and the Glide XP "score in place" was calculated for the carbohydrate structures. For each snapshot from the implicit solvent simulations, a

Glide grid was calculated in a similar way and then the Glide XP "refine and score" for the ligand was calculated. During the refinement process the protein was held rigid while the saccharides were minimized in place. A conjugate gradient minimization for a maximum of 1000 iterations in the OPLS2001 force field was used for the refinement. For each ligand, the resulting scores from Glide XP were averaged. The scores are provided in kcal mol^{-1} by the program, but the values should not be considered as absolute binding energies but rather as a relative measure.

Results and discussion

Since the expression of α1,2-linked fucose in mucosal tissues is crucial for susceptibility to infection with most norovirus GII.4 strains [4–7, 18, 31] it is expected that the α1,2-linked fucose is bound with significant affinity to the viral capsid protein. In the crystal structure of the VA387 strain, the fucose of the B-trisaccharide is seen to make strong contacts in the saccharide binding site [15]. Using Glide XP dockings we could reproduce the crystallographic pose of α1,2-linked fucose in the binding site and it had a Glide score of −6.0 kcal mol^{-1} (Fig. 1). In comparison, the crystallographic pose of the B-trisaccharide after minimization with the OPLS force field in the crystal structure of VA387 had a score of −7.3 kcal mol^{-1}. The structure of the A-trisaccharide from the X-ray structure of the complex (2OBS) could not be used due to geometry errors in its GalNAcα1,3Gal moiety (Fig. 2); in particular the bond angle of the glycosidic linkage of GalNAcα1,3Gal is 165° instead of the normal value of 114° and the tetrahedral geometry of the C3 of Galβ is significantly distorted as C3 is practically coplanar with C2, C4 and O3 (Fig. 2). The A-trisaccharide was therefore modeled from the B-trisaccharide in the site (2OBT) by addition of the NAc-group followed by IMPREF minimization in OPLS force field. The energy of the modelled conformation of the A-trisaccharide is 125 kcal mol^{-1} lower than the energy of the distorted conformation from the crystal structure according to single point MM4 calculations. As stated by Cao et al. [15] the electron density on the GalNAc residue is very weak whereas the Galβ and especially the Fucα show good density. It appears that the placement of the strained conformation of the GalNAc-Gal moiety in the crystal structure is not justified and we conclude that the crystallographic results do not contradict our present model of A-trisaccharide in the site. The NAc-group of the modeled A-trisaccharide made direct contacts with the protein which were not seen in the crystal structure. Specifically, it made contacts with Q331 and K348. This is in agreement with mutational studies using synthetic A- and B-trisaccharide assays that have shown that these residues are essential for the VA387 binding to A-antigens [16]. After minimization the modeled A-trisaccharide gave a Glide score of −7.2 kcal mol^{-1}.

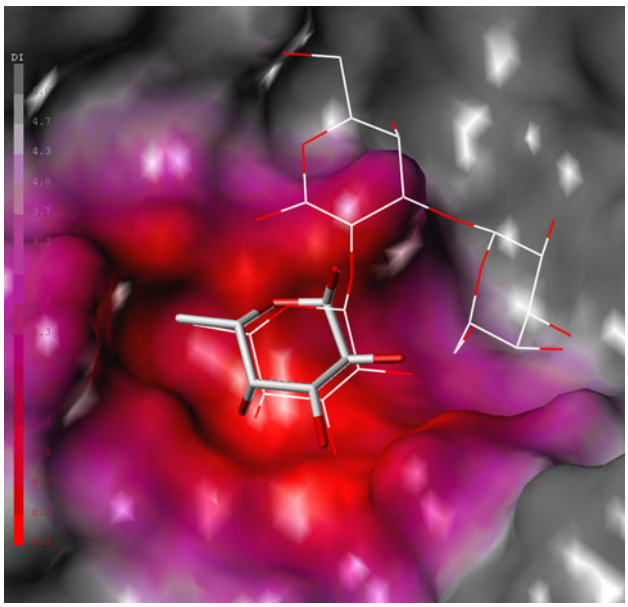

Fig. 1 The α-fucose monosaccharide docked into the crystal structure of VA387. The docked fucose (*stick representation*) was seen in the same orientation as the α1,2-fucose of the B-trisaccharide (*line representation*) in the crystal structure. The surface of the VA387 is color coded with Sybyl according to the distance between the fucose atoms and the surface residues with a cut-off of 3 Å

Fitting and molecular dynamics

The fitting of the extended ABO-structures in Table 1 with the α1,2-linked fucose residue superimposed onto the fucose moiety of the 2OBT structure indicated that all these structures except H-2, could be immediately accommodated in the site. The superimposition of H-2 caused minor clashing between the NAc-group of the

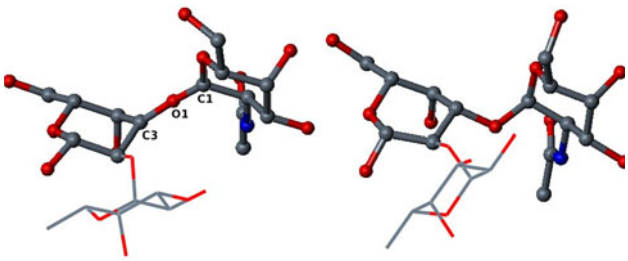

Fig. 2 *Left* The structure of the A-trisaccharide from the complex with VA387 (2OBS). *Right* The modeled A-trisaccharide. The fucose is shown with thin lines. Note in particular, the difference in the geometry of the bond angle at C3–O1–C1 and also the almost planar geometry of C3 in the crystal structure

Table 2 The interactions of capsid protein of VA387 binding site residues with the saccharide residues

	Sugar residues	Protein residues												
		Site I			Site II						N_{393} (B)	H_{395} (B)	S_{441} (B)	Y_{443} (B)
		Q_{390} (B)	D_{391} (B)	G_{392} (B)	A_{346} (A)	H_{347} (A)	Q_{331} (B)	K_{348} (B)	I_{389} (B)	C_{440} (B)				
A-tri	**2**,3Galβ		**4**	**6**									3	
													1	
	GalNAcα				4	6	6(4)	6(1)	6(6)	6			5(3)	
					2	**5**	**4(4)**		**6(6)**				**6(6)**	
							6(O3)							
B-tri	**2**,3Galβ						***1(O4)***						6	
													6	
	Galα				4	5	5	6		5			6	
					3	**6**	**2**	**6**		**3**			**6**	
							3(O3,O2)							
H-1	**2**Galβ												6	
	3GlcNAcβ			5								1		2
H-2	**2**Galβ												5	
	4GlcNAcβ			4							5			
H-3	**2**Galβ												6	
	3GalNAcα			3										3
Le^b	**2**Galβ												6	
	Fucα4	3	6	6										
	3,4GlcNAcβ													
A-1	GalNAcα				4	6	5(5)	6	6(6)	5			5(5)	
	2,3Galβ												4	
	3GlcNAcβ		1	5										1
A-3	GalNAcα				3	3	5(5)	6(1)	6(6)	3			6(5)	
	2,3Galβ												2	
	3GalNAcα			4										1
ALe^b	GalNAcα				1	6	6(5)	6	5(4)	6			6(6)	
	2,3Galβ													
	Fucα4	6	6	6										
	3,4GlcNAcβ													
B-1	Galα				3	4	5	6		4			6	
	2,3Galβ												6	
	3GlcNAcβ			5										3
BLe^b	Galα						1	6		2			6	
	2,3Galβ												3	
	Fucα4	4	6	6										
	3,4GlcNAcβ													2

Site I interacts with the sugar residues at the reducing terminus, whereas interactions with non-reducing terminal sugar residues are seen in site II. The numbers represent the number of snapshots out of 6 in which interactions were observed and interactions with the NAc-group are shown in parenthesis. The bold numbers indicate the contact details from the explicit solvent simulations. The numbers in bold italics indicate water mediated contacts. In this case the contacting sugar atoms are shown in parenthesis. Y443 also interacts with the α1,2-linked fucose residue. Concerning the glycosidic linkages, the attachment point for the α2-linked fucose is indicated in bold

GlcNAc and the region of the residues Q390 and D391. However, initial energy minimization with AMBER relieved these clashes.

MD simulations both with explicit solvent and implicit solvent showed torsion angles of the glycosidic linkages within favorable energy wells according to predictions with

GLYGAL (see supplementary material). The sampled glycosidic torsions superimposed on the adiabatic maps for their corresponding unbound structures are provided as supplementary material. As a benchmark of the simulations, root mean square deviations (RMSD) of the B-trisaccharide structure from the explicit and implicit models were calculated and compared to the crystallographic pose. After the explicit solvent MD simulations the αFuc-βGal moiety had an RMSD of 0.8 Å and the complete B-trisaccharide had an RMSD of 1.5 Å. After the OPLS minimization, the corresponding RMSDs for αFuc-βGal moiety and the complete B-trisaccharide were 1.2 Å and 1.9 Å, respectively. The contacts between the αFuc and the residues T344, R345, G442, and D374 in the crystal structure were retained. The larger deviation due to the αGal residue indicates the flexibility in that particular region, which is supported by the weak electron density in the crystallographic data of 2OBT [15]. The RMSDs are compiled in a table and provided as supplementary material. Despite some differences in positions of the αGal residue the explicit solvent and impicit solvent models showed similar results.

The contacts and the final pose from explicit solvent MD simulations on the A- and B-trisaccharide after OPLS minimization are shown in Table 2 and Fig. 3. The A-trisaccharide results show contacts between the NAc-group and I389 which are consistent with the results from implicit solvent simulations (Table 2 and Fig. 4). The interactions between the NAc-group and I389 are mainly due to hydrophobic interactions between the methyl groups. These interactions were not seen with either of the methods for the B-trisaccharide. This is in full agreement with the reported loss of binding for A-structures but not for B-structures of neo-glycoproteins and saliva samples [16]. In the case of Q331, the explicit solvent simulations indicate weaker contacts with the B-trisaccharide than with the A-trisaccharide, whereas the implicit solvent method gives similar contacts. A deeper analysis of the explicit solvent simulation snapshots indicated that the B-trisaccharide had fewer solvent mediated contacts (occasionally through O4 of Galβ, O2 and O3 of Galα with Nδ of Q331) compared to the A-trisaccharide, for which solvent mediated contacts (through O2 of Galα and O of Q331) are observed in all of the snapshots. Also both in the explicit and implicit solvent simulations, the A-trisaccharide had higher number of direct contacts with Q331 than the αGal of the B-trisaccharide (Table 2). These results are compatible with the reported reduction of binding of A- structures, but not B-structures in neoglycoprotein and saliva assays [16]. In the case of K348A the present results cannot explain the experimentally reported reduced binding of A-structures. This might be due to the inherent flexibility of the lysine residue. Thus our present modelling is partially supported by experimental results. Our model of the fit of the A-determinant appears more realistic than the strained structure in the reported crystal structure.

The agreement between the explicit and implicit solvent simulations for A- and B-trisaccharides (Table 2) justified the use of the less demanding procedure with short implicit solvent simulations for extended structures. The geometries of the complexes after the MD and the OPLS minimization are shown in Fig. 4 and the contacts with specific amino acid residues of the protein are summarized in Table 2. Due to the restraints applied the hydrogen bonds formed by the α1,2-linked fucose residue and the amino acid residues T344, R345, G442, and D374, were retained for every ligand as in the crystal structure [15]. The implicit solvent results of H-1, H-2, A-1 and B-1 shows hydrogen bonding and hydrophobic interactions of the GlcNAcβ residue with G392 and Y443. In case of Le^b, ALe^b and BLe^b the GlcNAcβ residue interacts only with Y443, while the α1-4 linked fucose is involved in interactions with Q390, D391 and G392. Mutational studies performed with saliva glycoconjugates [16] have suggested that residues Q390, D391, G392, and H395 are involved in the interactions with the downstream extensions of the ABO-determinants

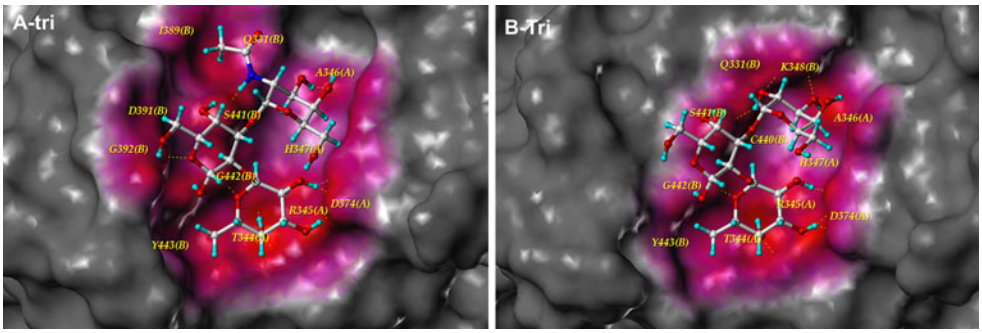

Fig. 3 The MOLCAD view of final explicit solvent MD snapshots after OPLS minimzation showing interacting residues identified with LIGPLOT are indicated with labels. The water molecules were removed prior Glide scoring. The coloring of protein surface shows the ligand proximity with a distance threshold of 3 Å. The coordinates of the complexes are accessible as supplementary material

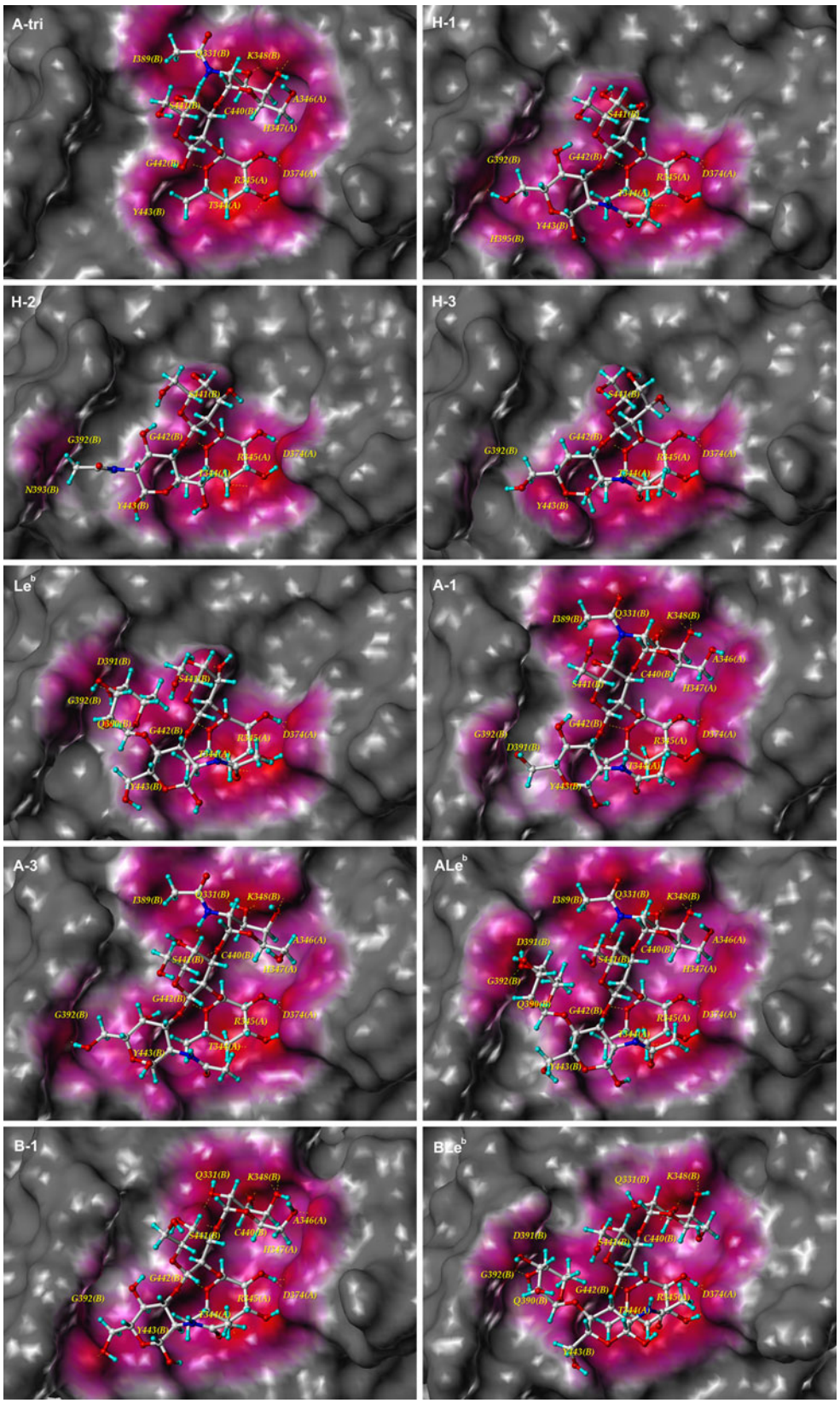

Fig. 4 The MOLCAD view of the final implicit solvent MD snapshots (250 ps) after OPLS minimization showing the interacting residues identified by LIGPLOT. The coloring of protein surface shows the ligand proximity with a distance threshold of 3 Å. The coordinates of the complexes are accessible as supplementary material

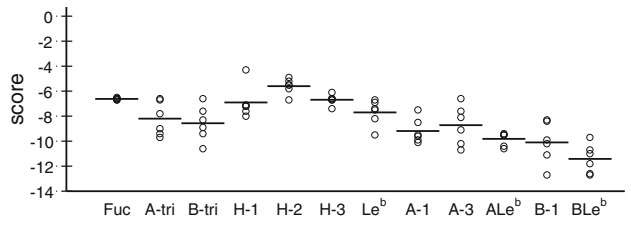

Fig. 5 Interaction scores calculated with Glide XP (expressed as kcal mol^{-1}) for poses obtained with implicit solvent simulations for the different saccharides. Horizontal lines represent average of the data points

in agreement with our modeling data. It can also be noted that the mutation Y443A is known to affect binding of α1,2-linked fucose [16].

The H-2 structure showed clashing after the initial fitting but not after minimization and during the MD simulations. Preliminary studies on extended type 2 structures with unrestrained minimization and MD indicate clashing (data not shown). Further studies on the extended type 2 chain structures are needed.

Glide scoring

The average Glide scores for the A- and B-trisaccharides from the explicit solvent simulations were −8.4 kcal mol^{-1} and −8.3 kcal mol^{-1} respectively. Similar score values for the A- and B-trisaccharides were obtained after the implicit solvent simulations and OPLS minimziation (Fig. 5). Out of the extended carbohydrate ligands, B-1 and BLe^b had the lowest scores. For the corresponding A-structures, a somewhat weaker binding was predicted. The A-3 structure had a score only slightly higher than A-1. H-2, H-3 and H-1 were the weakest in scores while A-, B-trisaccharides and Le^b had intermediate scores. The lower binding affinities for H-1 and H-3 structures, as compared to A- and B-structures, can be explained by the interactions involving the terminal GalNAcα and Galα residues in A- and B-structures that is not present in the H-structures, as stated by Cao et al. [15].

In type 3 structures the O1 of the reducing end 3GalNAcα- points into the bulk medium allowing space for further extensions, whereas in type 1 chain structures the extension of the 3GlcNAcβ- would be along the surface of the protein potentially resulting in steric hindrance. This may explain the weaker binding of the H-1 conjugate as compared to the H-3 conjugate [8, 16]. The favorable interaction scores of Le^b, ALe^b and BLe^b structures compared to the structures without α1,4-linked fucose suggest that this fucose has a significant contribution in the binding of these Lewis structures. This observation is supported by binding studies showing that Le^b binds stronger than H-1 to VA387 [8] and other GII.4 strains [17–19].

Conclusions

The present study shows that a variety of extended ABO-saccharides of type 1 and 3 can be well accommodated into the known fucose binding site of VA387. These results are in agreement with the crystal structure of the B-trisaccharide complex and can explain experimental binding data as well as dependence on the secretor gene. The A-trisaccharide in the crystal structure of the complex was found to have geometric errors. The herein modeled A-structures fitted into the site showed contacts in agreement with earlier mutational data. Energy calculations indicate that the α1,2-linked fucose and also the terminal GalNAcα/Galα1,3 make strong interactions with the protein. It is interesting that larger extended A- and B-active structures are well accommodated in the site. The results should be considered in the structure-based design of inhibitors for the carbohydrate binding site of GII.4 noroviruses.

Acknowledgments Grants from the Swedish research council (8266) and from governmental funds to the Sahlgrenska University Hospital and financial support from Biognos AB, Göteborg, are gratefully acknowledged. We are grateful for valuable discussions with Drs. I. Pascher, S. Sundell and G. Kemp.

References

1. Atmar RL, Estes MK (2006) Gastroenterol Clin North Am 35(2):275
2. Hutson AM, Airaud F, LePendu J, Estes MK, Atmar RL (2005) J Med Virol 77(1):116
3. Lindesmith L, Moe C, Marionneau S, Ruvoen N, Jiang X, Lindblad L, Stewart P, LePendu J, Baric R (2003) Nat Med 9(5):548
4. Kindberg E, Akerlind B, Johnsen C, Knudsen JD, Heltberg O, Larson G, Bottiger B, Svensson L (2007) J Clin Microbiol 45(8):2720
5. Tan M, Jin M, Xie H, Duan Z, Jiang X, Fang Z (2008) J Med Virol 80(7):1296
6. Thorven M, Grahn A, Hedlund KO, Johansson H, Wahlfrid C, Larson G, Svensson L (2005) J Virol 79(24):15351
7. Larsson MM, Rydell GE, Grahn A, Rodriguez-Diaz J, Akerlind B, Hutson AM, Estes MK, Larson G, Svensson L (2006) J Infect Dis 194(10):1422
8. Huang P, Farkas T, Zhong W, Tan M, Thornton S, Morrow AL, Jiang X (2005) J Virol 79(11):6714
9. Marionneau S, Ruvoen N, Le Moullac-Vaidye B, Clement M, Cailleau-Thomas A, Ruiz-Palacois G, Huang P, Jiang X, Le Pendu J (2002) Gastroenterology 122(7):1967
10. Harrington PR, Vinje J, Moe CL, Baric RS (2004) J Virol 78(6):3035
11. Harrington PR, Lindesmith L, Yount B, Moe CL, Baric RS (2002) J Virol 76(23):12335
12. Kelly RJ, Rouquier S, Giorgi D, Lennon GG, Lowe JB (1995) J Biol Chem 270(9):4640
13. Choi JM, Hutson AM, Estes MK, Prasad BV (2008) Proc Natl Acad Sci U S A 105(27):9175
14. Bu W, Mamedova A, Tan M, Xia M, Jiang X, Hegde RS (2008) J Virol 82(11):5340

15. Cao S, Lou Z, Tan M, Chen Y, Liu Y, Zhang Z, Zhang XC, Jiang X, Li X, Rao Z (2007) J Virol 81(11):5949
16. Tan M, Xia M, Cao S, Huang P, Farkas T, Meller J, Hegde RS, Li X, Rao Z, Jiang X (2008) Virology 379(2):324
17. Rydell GE, Nilsson J, Rodriguez-Diaz J, Ruvoen-Clouet N, Svensson L, Le Pendu J, Larson G (2009) Glycobiology 19(3):309
18. Lindesmith LC, Donaldson EF, Lobue AD, Cannon JL, Zheng DP, Vinje J, Baric RS (2008) PLoS Med 5(2):e31
19. Shirato H, Ogawa S, Ito H, Sato T, Kameyama A, Narimatsu H, Xiaofan Z, Miyamura T, Wakita T, Ishii K, Takeda N (2008) J Virol 82(21):10756
20. Siebenga JJ, Vennema H, Zheng DP, Vinje J, Lee BE, Pang XL, Ho EC, Lim W, Choudekar A, Broor S, Halperin T, Rasool NB, Hewitt J, Greening GE, Jin M, Duan ZJ, Lucero Y, O'Ryan M, Hoehne M, Schreier E, Ratcliff RM, White PA, Iritani N, Reuter G, Koopmans M (2009) J Infect Dis 200(5):802
21. Bohne A, Lang E, von der Lieth CW (1999) Bioinformatics 15(9):767
22. Lii J-H, Chen K-H, Allinger NL (2003) J Comput Chem 24(12):1504
23. Nahmany A, Strino F, Rosen J, Kemp GJL, Nyholm P-G (2005) Carbohydr Res 340(5):1059
24. Strino F, Lii J-H, Gabius H-J, Nyholm P-G (2009) J Comput Aided Mol Des 23(12):845
25. Agostino M, Jene C, Boyle T, Ramsland PA, Yuriev E (2009) J Chem Inf Model 49(12):2749
26. Nurisso A, Kozmon S, Imberty A (2008) Mol Simul 34(4):469
27. Stortz CA, Johnson GP, French AD, Csonka GI (2009) Carbohydr Res 344(16):2217
28. Kirschner KN, Yongye AB, Tschampel SM, Gonzalez-Outeirino J, Daniels CR, Foley BL, Woods RJ (2008) J Comput Chem 29(4):622
29. Hornak V, Abel R, Okur A, Strockbine B, Roitberg A, Simmerling C (2006) Proteins 65(3):712
30. Onufriev A, Bashford D, Case DA (2004) Proteins 55(2):383
31. Carlsson B, Kindberg E, Buesa J, Rydell GE, Lidon MF, Montava R, Abu Mallouh R, Grahn A, Rodriguez-Diaz J, Bellido J, Arnedo A, Larson G, Svensson L (2009) PLoS One 4(5):e5593